Conference Proceedings 6 – 8 October 2010
Glenburn Lodge, Gauteng

SAIIE 2010 Conference
Sustaining South Africa’s Global Competitiveness after 2010


24th Annual SAIIE Conference
Preface:

South Africa's Global Competitiveness is obvious in the daily news. Due to the 2010 FIFA Soccer World Cup, South Africa has invested in a number of large capital project developments, contributing to the economic activity in the engineering profession. However, what will happen after 2010? How will we ensure that South Africa remains competitive? The selected theme, “Sustaining South Africa’s Global Competitiveness after 2010” focuses on the role of Industrial Engineers in keeping the South African contribution sustainable. Sustainability takes many forms. The smart use of resources is one form of sustainability, the development of continuing innovations is another; protecting the environment, while maintaining economic growth, is yet another facet.

In 2009 we started off with an academic track in our annual conference, and the 2010 SAIIE Conference is now only the second one in many years, that has a full peer review process for academic papers. The academic environment in South Africa requires academics to publish in journals, or to present academic papers at peer reviewed conferences. This publication opportunity successfully complements our highly regarded accredited academic journal (The South African Journal of Industrial Engineering). Papers published in the 2009 proceedings received high visibility on the academic search engines, and we therefore believe that this is a worthwhile format to continue for the future.

Delegates were therefore offered two tracks for submission of presentation proposals: the applied engineering track and the academic track.

- Presentations submitted for the applied engineering track were approved on submission of the abstract only.
- Presentations submitted for the academic track were provisionally approved on the basis of an abstract, where after the authors were invited to submit a full-length academic paper, which was reviewed by a peer review refereeing process, by a programme committee, consisting of academics and industry experts. Referees were allocated based on their fields of expertise, and also on the basis that a referee may not review a paper by another author at his/her own institution or organisation.

The review process was managed through an on-line conference system, allowing referees to provide on-line feedback, and to ensure that a fully traceable record exists for all editorial decisions taken during the process. Accepted papers are published in the conference proceedings. Over 45 academic track proposals were initially submitted to SAIIE 2010 for consideration, but of these, only 22 academic papers made it through the total process.
This conference has two outputs:

- The printed *Conference Programme* includes an abstract of each Academic paper or Applied engineering presentation, not exceeding 300 words, to enable the delegates to select the sessions to attend.
- The *Conference Proceedings* (this document) is a CD-ROM, containing full-length papers that were submitted, reviewed and approved for the Academic Track only. Its purpose is to give full access to the complete conference material for many years after the conference is over.

We trust that you will enjoy the 2010 SAIIE Conference, and that this publication will serve as an ideal first step for exposing the work of the authors to the world!

2010 SAIIE Conference Organising Team
October 2010
Organising Team

Organising Committee
Corne Schutte, SAlIE President
Brian Munsamy, Thales / SAlIE Council
Hannelie Nel, University of Johannesburg / SAlIE Council
Bernadette Patricia Sunjka, University of Witwatersrand / SAlIE Council
Thereza Botha, Conference Secretariat - Technoscene

Programme Committee
Susan Adendorff, (Prof) University of Pretoria / SAlIE Council
Olufemi Adetunji, University of Pretoria
James Florence Bekker, Stellenbosch University
Alan Brent, (Prof) Stellenbosch University (SOPMP)
Marne De Vries, University of Pretoria
Dimitri Marinov Dimitrov, (Prof) Stellenbosch University
Theuns Dirkse van Schalkwyk, Stellenbosch University
Niek du Preez, (Prof) Indutech / Stellenbosch University
Heinz Erich Essmann, (Dr) Indutech
Neels Fourie, (Prof) Stellenbosch University
Igor Gorlach, (Prof) Nelson Mandela Metropolitan University
Dieter Hartmann, WITS
Grace Mukondeleli Kanakana Kanakana, Tshwane University of Technology
Paul Kruger, (Prof) University of Pretoria
Louis Louw, (Dr) Indutech
Brian Munsamy, Thales / SAlIE Council
Hannelie Nel, University of Johannesburg / SAlIE Council
Rene Pellissier, (Prof) UNISA
Tinus Pretorius, (Prof) University of Pretoria
Francois Retief, Falk Systems Engineering
Corne Schutte, (Dr) Stellenbosch University / SAlIE Council
Ranil Sing, Durban University of Technology
Herman Steyn, (Prof) University of Pretoria
Bernadette Patricia Sunjka, University of Witwatersrand / SAlIE Council
Wilhelm Uys, (Dr) Indutech
Lielz van Dyk, Stellenbosch University
Andre van der Merwe, (Dr) Stellenbosch University
Antonie van Rensburg, (Dr) University of Pretoria
Rias van Wyk, (Dr) Technoscan, United States
Krige Visser, (Prof) University of Pretoria
Tanya (Lane) Visser, Stellenbosch University
Konrad Harald von Leipzig, Stellenbosch University
Sarma Yadavalli, (Prof) University of Pretoria

Conference Organiser
Thereza Botha, Technoscene
<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Performance of South African Low Cost Airlines measured against International Market Leaders</td>
<td>I. Campbell and C. Pratley</td>
</tr>
<tr>
<td>32</td>
<td>Refining the Operating Model Concept to enable systematic growth in Operating Maturity</td>
<td>M de Vries, A vd Merwe, A Gerber and P Kotze</td>
</tr>
<tr>
<td>47</td>
<td>An Empirical study of factors influencing Project Quality in the Design Phase: A case of building projects in the Department of Building and Engineering Services of Botswana</td>
<td>E.N.D. Dodoo, C.C. van Waveren and K.Y. Chan</td>
</tr>
<tr>
<td>75</td>
<td>Improving Loading Times and Truck Rollover Tendency using a Pyramid Stacking Method</td>
<td>F. Kienhöfer, R. Crichton, J. Grobler and T. Dessein</td>
</tr>
<tr>
<td>88</td>
<td>Lean Healthcare: A Casualty of Inefficiency</td>
<td>R. Mandavha and D Hartmann</td>
</tr>
<tr>
<td>101</td>
<td>Micro-material handling through the manipulation of van der Waals forces by the generation of a predetermined surface topography.</td>
<td>S. Matope, A.F. van der Merwe, R. Nemutudi, M. Nkosi, M. Cele and M. Maaza</td>
</tr>
<tr>
<td>111</td>
<td>Warehouse pre-positioning for disaster relief in Southern African development communities</td>
<td>E. Meyer, W.L. Bean and V.S.S. Yadavalli</td>
</tr>
<tr>
<td>125</td>
<td>Exploring ways to reduce the Greenhouse Gas Emissions in the Textile Supply Chain</td>
<td>S. Nagel, C. Schoeman, V.S.S. Yadavalli</td>
</tr>
<tr>
<td>140</td>
<td>Deploying Lean Thinking to improve graduate quality and service delivery in Higher Education</td>
<td>R. Nkumbwa</td>
</tr>
<tr>
<td>152</td>
<td>A systems thinking approach to advancing GTL Technologies</td>
<td>A. Nyamvumba, G Kanakana and J Trimble</td>
</tr>
<tr>
<td>160</td>
<td>A conceptual Framework for the manufacturing of tooling by metal casting in refractory moulds and shells produced by Rapid Prototyping (Rapid Casting for Tooling)</td>
<td>K Nyembwe, D de Beer and S Bhero</td>
</tr>
<tr>
<td>172</td>
<td>Performance Enhancement in the milling of Ti6al4v</td>
<td>G.A. Oosthuizen, G. Akdogan and N. F. Treurnicht</td>
</tr>
<tr>
<td>185</td>
<td>The application of a Resilience Engineering Model for Organizational Innovation</td>
<td>R. Pellissier* and F. de Sousa</td>
</tr>
<tr>
<td>208</td>
<td>The effect of Time Variations in Assembly Line Balancing: Lessons learned in the clothing industry in South Africa</td>
<td>K. Ramdass and D. Kruger</td>
</tr>
<tr>
<td>220</td>
<td>An overview of recent machining improvements in Titanium</td>
<td>N.F. Treurnicht and G.A. Oosthuizen</td>
</tr>
<tr>
<td>232</td>
<td>An effective labour performance measurement aimed at optimizing productivity for a South African company</td>
<td>D.I. van Blommestein, S. Matope, A.D. Swart and G.A. Ruthven</td>
</tr>
<tr>
<td>244</td>
<td>XAP Open Source Communication protocol for Telemedicine</td>
<td>A.F. van der Merwe</td>
</tr>
<tr>
<td>256</td>
<td>Telemedicine - leveraging competitive advantage through the use of ICT Investment</td>
<td>L. van Dyk</td>
</tr>
<tr>
<td>267</td>
<td>Supply Response Analysis of the Namibian Mutton Industry: A Methodology Overview</td>
<td>D.N. van Wyk and N.F. Treurnicht</td>
</tr>
</tbody>
</table>
PERFORMANCE OF SOUTH AFRICAN LOW COST AIRLINES MEASURED AGAINST INTERNATIONAL MARKET LEADERS

I Campbell¹ and C Pratley¹
¹School of Mechanical Industrial and Aeronautical Engineering
University of the Witwatersrand, South Africa
lan.campbell@wits.ac.za

ABSTRACT

Financial and operational performance indicators of South African low cost airlines were determined and compared with international airlines, namely Southwest and Ryanair. Operational measures and financial data from two SA airlines, 1Time and Kulula, were used. Information was gathered from various sources including Internet-published documents and field observations. Operating strategies of the various airlines were evaluated. Of the airlines considered, Ryanair have the best operating profit margins due to the lowest cost structure. Southwest have the best passenger yields because of a differentiated product offering. There is opportunity for the South African low cost airlines to improve performance by modifying operations such as turnaround times and reducing staff costs.
1. INTRODUCTION

The South African (SA) airline industry has been through turbulent times in recent years, most notably indicated by the appearance and growth of low-cost airlines, the demise of Nationwide and Sunair, and the launch of the state-controlled Mango. Many company specific changes have also occurred, such as the public listing of 1-Time [1] on the JSE and the recent management turmoil at SAA [2]. Currently there are three low cost airlines (LCAs) in South Africa, namely Kulula, 1Time and Mango.

Volatility in both costs and demand is problematic for LCAs. The inherent price sensitivity of LCA customers means that price adjustments to account for cost fluctuations can radically affect demand and thereby revenues. Additionally many LCA tickets are sold six or more months in advance and so interim cost fluctuations cannot be recovered. This effect is amplified for SA LCAs due to the volatility of the local currency, since revenues are earned in local currency while a large part of the costs such as fuel and depreciation are based on international currencies.

Airline volumes in SA have grown since the introduction of LCAs. This may be because many LCA offerings can be favourably compared with other, more traditional forms of transport particularly over longer distances. The LCA industry in SA has become an important part of the transportation infrastructure. Further reductions in costs could increase volumes still more.

LCAs are characterised by a strong drive to reduce costs in comparison with traditional or legacy airlines. Additional characteristics include increasing the number of seats per aircraft, flying from uncongested hubs, charging for food and drinks, not assigning seat numbers and using the Internet for sales [3].

There is a growing tendency among LCAs to charge customers separately for non-essential items. This practice ensures a lower cost base offering and thereby widens the target customer base and creates opportunities for alternative revenue streams. Charging for luggage or additional luggage is one such approach. Other complementary sources of revenue include food and drink sales, website advertising, website sales of accommodation and car rental and sale of branding rights on aircraft.

LCAs in North America and Europe are more mature than the relatively new SA LCAs. Southwest launched in 1971. Ryanair was modeled on Southwest and began operations in 1990. Kulula, the first LCA in SA, started operations in 2001 [4][5][6]. Both Southwest and Ryanair have proven to be sustainable businesses. Southwest were the largest airline in the world in terms of passenger numbers until 2008 when the merger of Delta and Northwest displaced them. Ryanair are currently the seventh largest airline after many years of remarkably fast growth [7]. Ryanair are renowned for innovation and consistently return a profit for investors. It is therefore useful to compare performance and operating styles of the SA LCAs with well-established international counterparts. The sustainability of the local industry can be evaluated. Opportunities for improvements in the operations of local LCAs may also be revealed. This study compares the performance and operational strategies of local LCAs with Southwest and Ryanair.
2. LITERATURE REVIEW

2.1 Performance indicators

There are a variety of operational performance indicators specific to airlines in addition to financial performance indicators that can be gleaned from the financial statements. Several useful performance indicators for airline performance comparison are discussed in this section.

Aircraft block time is the time taken for an aircraft to complete a single journey, from the time at which it moves under its own power until it stops moving at its destination. The block aircraft speed is the distance flown divided by the block time.

Available seat miles (ASM) are defined as the distance flown times the number of seats for sale. It is a measure of the production of an airline over a period of time. The two key drivers of ASM are capacity and block speeds [8]. Cost per available seat mile (CASM) equals the airline operating cost divided by the ASM.

Revenue per available seat mile (RASM) is the total airline revenue divided by ASM [9]. The ultimate measure of operating performance is whether RASM exceeds CASM [8]. These measures are, however, inappropriate for comparing airlines with dissimilar operating strategies. Legacy or traditional airlines cannot be compared with LCAs using this approach. In this situation, a preferred measure is the Revex Ratio. This is the revenue from a unit expenditure, and equals operating revenue divided by operating cost [8]. Since the Revex Ratio is dimensionless, it is useful for comparing airlines operating in different currencies.

The operating margin is defined as the operating profit divided by the operating revenue [9]. The operating margin is therefore a dimensionless measure of the operating efficiency.

The passenger load factor (PLF), also known as the average load factor (ALF), is the average of the percentage of seats occupied in any aircraft. It is a commonly-used performance measure indicating to what extent the airline output is being utilized by passengers. It is also a measure of the demand/supply balance at the operating price point [8]. High ALFs mean lower costs per passenger, but they can increase aircraft turnaround times (eg passenger embarkation and disembarkation times) due to greater volumes of passengers and baggage. Revenue passenger miles (RPM) are the number of fare paying passengers multiplied by the average distance flown [10]. Therefore ALF can be calculated by dividing RPM by ASM.

Passenger yield (PY) is a measure of revenue per unit of output sold and is defined as the operating revenue divided by the RPM [9]. According to Holloway [8], decreasing PYs over time indicate either growing competition or a declining market, whereas increasing PYs over time indicate that airline growth is less than the demand.

The breakeven load factor (BELF) is defined by Holloway [8] as the operating cost per ASM divided by the operating revenue per RPM. BELF is sometimes more convenient to use than ALF as a comparative performance indicator since it accounts for flying costs. It is the point at which an airline’s CASM is equal to its RASM.

The employee efficiency ratio is the number of airline employees per aircraft operated and indicates whether there is employee inefficiency or overcapacity within an airline.

An airline’s stage length (average route distance) has a significant impact on its CASM, RASM and PY. This is because long stage lengths allow higher utilization of crew and aircraft combined with lower fuel burn, landing fees and maintenance per unit of distance. The effect also depends on the aircraft used, since CASM is a function of aircraft size, with smaller aircraft having lower CASM over shorter legs than larger aircraft. As a result, when
comparing airlines that operate over different stage lengths, performance indicators need to be adjusted accordingly. MIT [9] suggest multiplying each performance indicator by an adjustment factor defined as the square root of the ratio between the observed stage length and the average stage length. The average stage length is taken as the average of those of all the airlines being compared.

2.2 Airline operational performance

After take-off, the time to arrival at the destination is predictable between flights. There is little that can be done to improve block speeds apart from extending flight distances.

What happens after landing and before take-off for the next scheduled flight can make a difference to airline profitability and customer satisfaction. After landing, the aircraft is prepared for the next flight. The time this takes is referred to as the turnaround time or rotation time. Delays often occur due to the number of activities that must be carried out in the turnaround process and the fact that many different people are involved, including disembarking and embarking passengers. Turnaround time is often taken from the time the aircraft stops (brake or chocks on) to the time it is ready to start moving again (brake or chocks off) [11].

Departure delays and disruptions affect customer satisfaction. For this reason, the differences between the scheduled times and the actual departure times and arrival times are measured. Long taxi-in and taxi-out times can negatively impact arrival and departure times.

2.3 Aircraft types

LCAs commonly use the McDonnell-Douglas MD-80 series (eg MD-82, MD-83, MD-87) and the Boeing 737 series of aircraft. Modern aircraft types use less fuel than older generation types. The actual fuel consumption depends on engine types installed. The older generation MD-82 can use 19% more fuel [12] than a new generation 737-800.

Maintenance costs are normally lower on new aircraft than older aircraft, particularly if the newer aircraft are still under guarantee. However, maintenance costs also depend on variable factors such as labour cost, aircraft age, aircraft mileage and aircraft type, and therefore will be different for each aircraft and airline company.

2.4 Airline fleet size and age

Southwest operates 544 Boeing 737 jets of various sub-types as of July 2009, with an average age of 10 years and seating either 122 or 137 passengers. Ryanair [4] operates 202 new generation 737-800's with an average fleet age of just over 2.5 years and each with 189 seats.

Kulula [6] have a fleet of three 165 seat 737-400's, one 145 seat 737-300, one 737-800 and one 737-200. They have ordered at least two more new 186 seat 737-800's which suggests they are in the process of modernizing their fleet. Mango was launched in 2006 with three new 186 seat 737-800s which they currently operate [13]. The 1Time fleet consists of four 157 seat MD82s, five 157 seat MD83s (one stored) and four 132 seat MD87s (two stored) [14].
3. RESEARCH METHODOLOGY

3.1 Information sources

A variety of data gathering techniques were employed. The LCA industry is competitive and therefore airline companies are reluctant to provide operational data beyond that required by legislation. Some data can be obtained from public sources such as company websites, airport and aviation authority websites and airline association websites. North American airline companies are required by law to publish certain information in the form of “20-F” statements.

The Johannesburg Stock Exchange (JSE) requires listed South African airline companies to publish financial data regularly. Both 1Time and Comair are listed on the JSE, but Comair only provides information for Kulula combined with the other BA and Comair businesses. Other data on 1Time and Kulula was gathered by interviews with staff and field observations.

Apart from the website and some news items, there is little published information about Mango. The staff did not respond to requests for interviews or information.

3.2 Quantitative data

3.2.1 1Time

The annual financial reports were obtained [14]. These contained the annual passenger numbers, available seats, number of aircraft, operating revenue and operating profit. The company provided the available seat miles (ASM) and employee numbers via interview [15].

3.2.2 Kulula

On-time performance is published on the Kulula website [6]. Additional data was obtained using field observations. This involved the recording of flight data for a number of actual flights, including ground travel times, aircraft turnaround times, aircraft load factors and departure and arrival times. When delays occurred, reasons and delay lengths were recorded. In total, data from 54 Kulula flights over a two month period was recorded.

3.2.3 Southwest

Information was obtained from the American Bureau of Transportation Statistics [16], MIT’s airline data project [9], AirlineFinancials.com [17] and the securities information “10-K” filing [18].

3.2.4 Ryanair

Information was obtained from the European Low Fares Airline Association [19], published financial reports and the “20-F” statement [4].

3.2.5 Currency fluctuations and conversions

Comparisons between airlines’ performance in different countries is complicated by exchange rate volatility. Currencies were converted to US dollars for the purpose of comparisons. The largest currency exchange rate fluctuations which affected this study occurred in 2008 for the exchange rate of local SA currency and the US dollar. Twelve percent fluctuations around the mean were observed [20]. The yearly mean value was used in calculations. Upper and lower limits are also presented relating to the upper and lower conversion rates experienced during the year.
4. RESULTS

4.1 Financial and operational performance indicators

Performance indicators for the two international LCAs and 1Time were calculated. Historical information from 2005 to 2008 was included where possible and is included in graphical form in Figures 1 to 7. The employee to aircraft ratios were found to be 35, 65 and 55 for Ryanair, Southwest and 1Time respectively.
4.2 Recorded performance measures for Kulula

4.2.1 Turnaround Times

At the two largest SA airports, Cape Town and Oliver Tambo (Johannesburg), the minimum turnaround time measured was 40 minutes at Cape Town and the average of 13 recorded measurements was 55 minutes. Of the smaller airports, including Port Elizabeth, Durban and George, the minimum measured turnaround time was 25 minutes at Port Elizabeth, followed by 26 minutes at Durban. The average for the turnaround times at the smaller airports was 41 minutes (of 12 recorded measurements).

Eleven turnaround time measurements were recorded at Lanseria Airport. The minimum was 28 minutes and the average was 42 minutes. This sample mean differed from the largest airports turnaround time mean of 55 minutes at the 85% T-test confidence level.

4.2.2 Load factors and timing measures

Other operational performance data recorded for the Kulula flights is shown in Table 1 and compared with available data from the other airlines.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Kulula</th>
<th>Southwest</th>
<th>Ryanair</th>
</tr>
</thead>
<tbody>
<tr>
<td>Passenger load factor (%)</td>
<td>76</td>
<td>71</td>
<td>81</td>
</tr>
<tr>
<td>Average on-time performance (%)</td>
<td>79.7</td>
<td>81.7</td>
<td>84.8</td>
</tr>
<tr>
<td>Average departure delay (minutes)</td>
<td>13.0</td>
<td>10.3</td>
<td>-</td>
</tr>
<tr>
<td>Average arrival delay (minutes)</td>
<td>8.0</td>
<td>5.3</td>
<td>-</td>
</tr>
<tr>
<td>Average taxi out time (minutes)</td>
<td>8.7</td>
<td>10.4</td>
<td>-</td>
</tr>
<tr>
<td>Average taxi in time (minutes)</td>
<td>4.0</td>
<td>4.4</td>
<td>-</td>
</tr>
</tbody>
</table>

5. DISCUSSION

5.1 Performance indicators

Figures 1 and 2 show that Ryanair has the highest operating margins and Revex Ratio. Both Ryanair and Southwest have experienced decreasing margins since 2006 while 1Time has been relatively stable. 1Time’s most recent Revex Ratio was similar to that of Southwest. BELF and CASM shows that Ryanair has better cost efficiency compared to the other two airlines. Operating costs of both 1Time and Southwest are similar.

Table 2 shows fuel costs experienced by Southwest and Ryanair in 2007. Although Ryanair uses less fuel per ASM, Southwest have a lower fuel cost because of the lower fuel price in North America. Table 3 shows costs experienced by the two airlines in 2007. Southwest also have lower depreciation and amortization costs per ASM. Ryanair’s lower total cost per ASM is largely due to staff costs, and, to a lesser extent, maintenance costs.
Table 2 - Fuel Costs [6][18]

<table>
<thead>
<tr>
<th></th>
<th>fuel price (US$/gallon)</th>
<th>fuel usage/ASM (gallons)</th>
<th>fuel cost/ASM (US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southwest</td>
<td>1.7</td>
<td>0.015</td>
<td>0.025</td>
</tr>
<tr>
<td>Ryanair</td>
<td>2.5</td>
<td>0.012</td>
<td>0.030</td>
</tr>
</tbody>
</table>

Table 3 - Cost Breakdown per ASM for 2007 (US cents) [6][18]

<table>
<thead>
<tr>
<th></th>
<th>staff</th>
<th>depreciation</th>
<th>fuel &amp; oil</th>
<th>maintenance</th>
<th>other</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southwest</td>
<td>3.23</td>
<td>0.56</td>
<td>2.55</td>
<td>0.62</td>
<td>2.16</td>
<td>9.12</td>
</tr>
<tr>
<td>Ryanair</td>
<td>0.98</td>
<td>0.62</td>
<td>3.00</td>
<td>0.18</td>
<td>2.86</td>
<td>7.64</td>
</tr>
</tbody>
</table>

Ryanair have a substantially lower employee to aircraft ratio than 1Time and Southwest. Staff costs account for approximately 23% of total airline costs [21]. This is the main reason for Ryanair’s lower costs and accounts for approximately 5% of the 13% comparative cost advantage Ryanair has over Southwest. Both 1Time and Southwest have more inclusive services in their product offerings which may partially explain the higher number of staff. Both Southwest and 1Time operate smaller capacity and older aircraft than Ryanair, hence the similarity in overall costs (CASM).

Achieving high labour productivity is considered an integral part of the LCA business strategy [21]. Most of the LCA employee cost advantage over traditional airlines is due to flexible work rules and cross utilization of staff, as well as a culture of cooperation among labour groups. Belobaba [21] mentions Southwest as an example of this. This research suggests Ryanair have outperformed Southwest in this respect.

The three airlines show mildly increasing trends in PY from 2006 to 2008. This is more likely a reflection of cost increases rather than demand increasing more than supply as suggested by Holloway [8]. 1Time has a PY comparable to Ryanair’s PY and has three stored aircraft. This indicates an oversupply of seats in the local market. The introduction of Mango in 2006 put pressure on PYs, as indicated by the demise in 2008 in SA of the LCA Nationwide Airlines.

Southwest has higher passenger yields than Ryanair and 1Time and therefore better profitability per passenger. The differences in RASM between the three airlines is small. Southwest’s overall profitability has fallen in recent years due to increasing costs and low ALFs. Southwest per passenger costs per mile (PPC) can be calculated by dividing CASM by ALF. In 2008, PPC for Southwest in USD was 14.2 US cents, higher than Ryanair (9.8 US cents) and 1Time (11.6 US cents) due to their smaller capacity airplanes compared to Ryanair and lower load factors compared to Ryanair and 1Time. Further investigation reveals that Southwest have a differentiated price structure, with distinct products on offer (Southwest 2010). As a result of this and lower ALFs, Southwest’s RASM is not correspondingly better than that of the other two airlines. Note that product differentiation is likely to limit the aircraft seating density achievable compared to a strictly low-cost offering such as that of Ryanair. Less seats per aircraft combined with higher prices for one of two offered products results in Southwest’s lower ALFs.
As opposed to Southwest and 1Time, Ryanair charge separately for luggage and other cost items. This strategy encourages lower cost passenger behavior and ensures fewer employees are required, but at the expense of lower PYs.

The ALF minus BELF (figure 7) is a good indication of operational performance and risk. Although both Southwest and Ryanair show declining margins over the years, Ryanair still had a 14% margin in 2008. Both Southwest and 1Time’s comparatively small 3% margin in 2008 indicates that they should consider cost reduction. In practice, cost reduction will involve staff cost reduction which could adversely affect passenger yields.

5.2 Turnaround times

The recorded turnaround times show that the smaller airports and Lanseria have lower average turnaround times than the two larger airports. Kulula have a turnaround time advantage of 13 minutes on average at Lanseria compared to Oliver Tambo. Few of the measured turnaround times are below or equal to the average reported values for Southwest and Ryanair of 25 minutes. 1Time and Mango will have longer average turnaround times than Kulula since they do not use Lanseria in Johannesburg. Since most flights are approximately 1 hour long, the time lost per day will add up to a loss in ASM, driving up CASM. In practice, in a typical 14 hour day the turnaround time savings may allow Kulula to add an additional daily flight for shorter routes such as Durban to Johannesburg where they could currently operate nine daily flights.

The higher turnaround times SA LCAs experience compared to international LCAs is partly due to use of secondary airports in North America and Europe. High volumes of LCA traffic at secondary airports compel airports to focus efforts on lowering turnaround times, since the LCAs are their primary customers. Secondary and smaller airports in SA such as Lanseria cater for a wide variety of customers and aircraft sizes. Therefore they do not specifically focus efforts on LCA requirements.

1Time has low value planes and higher variable aircraft operation costs such as fuel and maintenance compared to Kulula and Mango, therefore high turnaround times will have a lesser effect on costs. Kulula and Mango’s higher fixed costs due to newer, more expensive aircraft means more flight time and lower block speeds are important for reducing costs per passenger. Therefore the adverse effects of high turnaround times on them will be exacerbated.

Figures 8 and 9 show turnaround procedures for Easyjet (20 mins, a large UK based LCA) and Kulula (45 mins) respectively. Easyjet have reportedly achieved turnaround times as low as 15 minutes. The diagrams show where Kulula experiences comparative time penalties in the turnaround process. The biggest time penalty of 13 minutes is for passenger embarking. Other areas which need attention in terms of speed are passenger disembarking, cleaning and bag search for unboarded passengers, which have time penalties of 5, 4 and 4 minutes respectively.
Performance of South African Low Cost Airlines measured against International market Leaders - Campbell, Pratley

**Figure 8** - Easyjet Turnaround Processes [11]

**Figure 9** - Kulula Turnaround Processes [11]
International LCAs commonly do not pre-assign seat numbers whereas all SA LCAs do. SA airlines have a booking-in process whereby every passenger receives a boarding pass and is pre-assigned a seat. Embarkation time is wasted by passengers having their boarding passes checked, looking for their seats and then blocking aisles for other passengers while they stow their hand luggage. During boarding, passengers in the departure area are slow to move to the aircraft. International LCAs such as Ryanair and Southwest do not pre-assign seat numbers to passengers. This incentivises passengers to speed up the boarding process to obtain better plane seats [3]. It may also reduce the number of unboarded passengers. It is also possible to direct passengers to fill the aircraft from the back to the front and thereby reduce boarding times further.

International LCAs reduce carry-on (cabin) luggage to speed boarding times, since passengers don’t block aisles while stowing large amounts of luggage in already full overhead lockers. They achieve this by measuring the dimensions of all carry-on luggage at check-in and only allowing each passenger to carry on one item of a predefined maximum size. Local LCAs often have similar rules but do not strictly enforce them.

Average departure delay and arrival delay times are marginally worse for Kulula than Southwest Airline, with average measured 2.7 min and 1.7 min penalties respectively. Kulula has an advantage over Southwest in taxi in and out times. This is because of the reduced amount of traffic in SA compared to North America.

50% of recorded sample take-offs were delayed, mostly due to rotation. This relates to fueling activities, cleaning activities, or baggage handling activities. Some previously documented reasons for delays are given in Table 4 for Kulula circa 2005.

<table>
<thead>
<tr>
<th>Description</th>
<th>Number of Delays (rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unboarded passengers</td>
<td>1 (most)</td>
</tr>
<tr>
<td>Missed slots</td>
<td>2</td>
</tr>
<tr>
<td>Operations (other)</td>
<td>3</td>
</tr>
<tr>
<td>Aircraft rotation due to customer services - late arrival of incoming flight</td>
<td>4</td>
</tr>
<tr>
<td>Aircraft defects - on departure</td>
<td>5</td>
</tr>
<tr>
<td>Computer failure</td>
<td>6</td>
</tr>
<tr>
<td>Late delivery of loading or loading of catering</td>
<td>7</td>
</tr>
<tr>
<td>Aircraft rotation due to information management</td>
<td>8 (least)</td>
</tr>
</tbody>
</table>

Most delays are due to unboarded passengers. Encouraging less luggage and not pre-assigning seats would result in less delays due to fewer incidents of unboarded passengers and the related baggage search.

There are also many delays due to connections. Most international LCAs only offer a point to point service to prevent connection delays. Since missed slots are often caused by schedule disruption because of delays, total delays could be reduced substantially by not pre-assigning seat numbers, improved control of baggage and not offering connecting flights.

5.3 Operating strategies
Ryanair offers less services included in the fare but allows passengers the option of paying extra for them. This has caused ancillary revenue in 2009 to reach 20% of total revenue [4]. In comparison, Southwest and 1Time have more product differentiation and more included services, hence more staff costs.

Kulula’s measured ALF of 76% is between Southwest’s 71.2%, and 1Time’s 82% in 2008. They have a targeted marketing strategy and strong market position, being the first SA LCA and largest SA LCA. 1Time would be unprofitable at this level of load factor. Therefore Kulula may have a lower BELF and hence lower cost/revenue ratio than 1Time.

Assuming an MD-82 19% fuel premium over the 737-800, similar maintenance costs, a 10 block-hr operating day, 1% per month interest cost of capital, a value of R21m for an MD-82 [22][23], R720m for a 737-800 [24], and R7/litre jet A1 price [25], a comparison can be made between operating costs of a 737-800 with an MD-82 (table 5). The 737-800 costs R3500 more to fly per block hour, but is 6% cheaper per passenger seat because of its larger capacity. The difference will be exacerbated during time periods of high fuel costs. This calculation suggests that at current fuel prices 1Time and Southwest may be operating at a small cost disadvantage to their competitors with newer aircraft. Both airlines will also be more sensitive to fuel price increases than their competitors. Table 3 shows that maintenance costs for Ryanair are less than those of Southwest. Similarly, maintenance costs for 1Time are likely to be higher than for Kulula due to its older fleet.

Table 5 - Fuel and ownership cost comparison

<table>
<thead>
<tr>
<th></th>
<th>737-800</th>
<th>MD-82</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuel cost (R/hr)</td>
<td>26083</td>
<td>30778</td>
</tr>
<tr>
<td>Cost of ownership (R/hr)</td>
<td>8500</td>
<td>300</td>
</tr>
<tr>
<td>Plane capacity (seats)</td>
<td>186</td>
<td>157</td>
</tr>
<tr>
<td>Total cost per hour (R/hr)</td>
<td>34583</td>
<td>31078</td>
</tr>
<tr>
<td>Cost/passenger seat-hour (R/hr)</td>
<td>186</td>
<td>198</td>
</tr>
</tbody>
</table>

1Time and Southwest have a marketing disadvantage compared to competitors with newer fleets. Many customers consider older aircraft more accident prone than newer aircraft. However, 1Time and Southwest have a strategic advantage in that the cost of not using such older aircraft is relatively low. This allows 1Time valuable flexibility and means they have lower risk in the volatile SA market.

Because of 1Time’s older aircraft, they are limited to operational improvements involving product differentiation and employee productivity improvements. This research suggests that staff costs are the major cost driver that should be addressed by SA LCAs. Improving passenger yields with older aircraft such as Southwest do is also a possible strategy to increase profits.
6. CONCLUSIONS

This research has revealed that higher margins may be achievable by SA LCAs through cost reduction strategies, particularly with regard to staff costs. SA LCA’s could position themselves as a “true LCA” such as Ryanair have done in Europe. Alternatively, higher passenger yields can be achieved by product differentiation, as the largest low-cost airline (Southwest) has shown.

Possible operational changes for SA LCAs could result in decreased turnaround times and higher aircraft utilisation. This would improve specifically Mango’s and Kulula’s performance. However, due to low value aircraft it will have a lesser effect on 1Time’s performance.

None of the SA LCAs are using all of the cost reduction strategies international competitors such as Ryanair use. This suggests there may be opportunity to introduce some or all of these strategies to reduce costs per passenger and hence increase profitability. Many of these strategies have been shown to work synergistically together to create better cost-conscious employee and customer behavior.

7 REFERENCES

Performance of South African Low Cost Airlines measured against International market Leaders - Campbell, Pratley


FINANCIAL MAINTENANCE MODELLING IN THE FRAMEWORK OF THE UNIFIED FINANCIAL ANALYSIS METHODOLOGY

Dr J.L. Coetzee\(^1\) and Prof Dr W. Breymann\(^2\)

\(^1\)M-Tech Consulting Engineers (Pty) Ltd
Vanderbijlpark, South Africa
jasper.coetzee@m-tech.co.za

\(^2\)Institute of Data Analysis and Process Design
Zürich University of Applied Sciences, Switzerland
wolfgang.breymann@zhaw.ch

ABSTRACT

Maintenance budgets are a perpetual problem because the understanding of maintenance people regarding the necessity of maintenance actions are not easily communicated to the general management of the business. This often leads to necessary maintenance being postponed. Although some effort has been put into explaining the relationship between maintenance expenditures and production outcomes by various authors, this remains one of the areas of misunderstanding, unavoidably leading to at best a sub-optimization of company output and thus profit.

A similar problem of a waste of resources because of a misunderstanding between different departments in large organizations also exists in the financial industry. During the last years this problem has been successfully approached by a methodology called Unified Financial Analysis (UFA). Given the stochastic nature of changes in markets and other economic realities, UFA methodology makes it possible to manage the risk and profitability of loans and investments well.

The principles embedded in the UFA approach are well suited for making the realities of maintenance expenditure versus the benefit of achieving high operational readiness of plant visible to higher management. It can also serve to assist maintenance management in managing maintenance risk.

Because of the high level of complexity involved, extensive research is envisaged to achieve this objective. The present research is driven conceptually from the maintenance engineering side and tactically from the financial side.

The group working on the project is mainly situated at the Zurich University of Applied Sciences (ZHAW) in Winterthur, Switzerland, but with maintenance engineering inputs from the University of Johannesburg and the Terotechnica Maintenance College. Some funding has been obtained in Switzerland for the necessary UFA research. Being a joint research project, funding for the conceptual research has to be obtained from South African sources.

OPSOMMING

Instandhoudingsbegrotings is ‘n voortdurende probleem omdat die insig van instandhoudingsbestuur rakende die noodsaaklikheid van instandhoudingsaksies meesal nie maklik kommunikeerbaar is aan die besigheid se topbestuur nie. Dit lei diikwels daartoe dat noodsaaklike instandhouding uitgestel word. Alhoewel verskeie outeurs werk gedoen het om die verband tussen instandhoudingsuitgawes en produksie-uitkomstes toe te lig, bly dit

\(^1\) Managing Director, M-Tech Consulting Engineers (Pty) Ltd.
\(^2\) Professor of Solid Physics, School of Engineering, Zurich University of Applied Sciences.
'n Soortgelyke probleem van `n vermorsing van hulpbronne as gevolg van `n misverstand tussen die verskillende afdelings in groot organisasies bestaan ook in die finansiële sektor. Gedurende die afgelope paar jaar is hierdie probleem suksesvol aangespreek deur `n metode genaamd `Unified Financial Analysis` (UFA). Gegewe die stogastiese aard van die veranderinge in die markte en ander ekonomiese werklikhede, maak die UFA metode dit moontlik om die risiko's en die winsgewendheid van lenings en beleggings goed te bestuur.

Die beginsels opgeneem in die UFA benadering is uitses geskik om die realiteit van die balans tussen instandhoudingsuitgawes en die voordelte van die berekening van hoë operasionele gereedheid van produksietoerusting sigbaar te maak vir hoër bestuur. Dit kan verder ook dien om instandhoudingsbestuur te ondersteun in die bestuur van instandhoudingsrisiko.

As gevolg van die hoë vlak van kompleksiteit wat betrokke is, word uitgebreide navorsing in die vooruitsig gestel om hierdie doel te bereik. Die huidige navorsing word konseptueel vanuit `n instandhoudingsingenieurswese hoek aangedryf, terwyl die finansiële navorsing hoofsaaklik takties van aard is.

Die groep wat aan die projek werk is hoofsaaklik gesentreer by die Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) in Winterthur, Switserland, maar met instandhoudingsingenieurswese insette vanaf die Universiteit van Johannesburg en die Terotechnica Instandhoudingskollege. Gedeeltelike befondsing vir die nodige UFA navorsing is reeds vanuit bronne in Switserland bekom. Aangesien die navorsingsprojek `n samewerkingskarakter het, moet befondsing vir die konseptuele navorsing vanuit Suid Afrikaanse bronne bekom word.
1. INTRODUCTION

It is well known that the right level of maintenance increases the overall production capability of technical assets, i.e. that there is an optimum level of maintenance yielding a maximum contribution to overall profit. Even though the eventual increase of profit by means of a given increase of the level of maintenance is somewhat indirect, on a conceptual level this relationship is in principle understood [3] and agreed upon. Unfortunately, to this day this theoretical insight is practically useless for operational decisions because in general it is impossible to quantify how much maintenance is needed (in financial terms) to generate a given amount of additional income. In particular, it is not possible to determine the trade-off between preventive and reactive maintenance and, a fortiori, the level of maintenance needed for maximum profit. In this situation it is comprehensible that, according to their stake, different parties involved have different judgments of the situation. While the maintenance managers focus on operational readiness, upper management is primarily concerned with the costs involved. These differences in perception may explain the gap - well known and largely deplored - between maintenance managers and upper management.

Interestingly, even though with the advent of ERP systems it is possible to capture all expenditure and income, this didn’t really change the situation. The reason is that ERP systems do not provide a well-founded methodology for experimenting with different levels of prevention and its outcomes. In fact there are no clear concepts of how the important structures of the productive process should be modelled in financial terms, because of a lack of a set of adequate guiding principles or organising principles.

Evidently, new, yet undiscovered concepts are needed. A very few attempts into this direction have been undertaken up to now [4]. However, there is a sector where such concepts do already exist - the financial sector. Due to the risk management requirements of the regulators various approaches for the risk and profitability management at the level of the whole institutions have been developed.

The problem in finance is similar: A proper assessment of the expected profit and the riskiness of a bank require the valuation of all expected future cash flows, and the evaluation of their risk involved. To do this properly for a whole institution, boundaries between different departments (the silo structure) must be overcome. Second, to get a good estimation of future cash flow streams it is not enough to just register all past cash flows. To be able to derive an expected future situation from the past and the present, the cash flow patterns must be organised according to the correct principles.

One of the most systematic and all-encompassing approach for organising cash flow patterns is a methodology recently published under the name of “Unified Financial Analysis” [1], which relies on the concept of standard contract types. A contract in the sense of this methodology can be imagined as the algorithmic counterpart of a legal contract that determines the cash-flow generating rules. In the simulation the actual cash flows are generated in interaction with a model of the environment (i.e., the financial market, counter parties, costs, and some additional statistical rules). Any imaginable kind of analysis can then be applied to the generated cash flow patterns. A considerable efficiency gain is obtained through standardisation by means of standard contract types since about 95% of the financial volume of a typical financial institution can be modelled with about two dozens of standard contract types.

The idea we propose in this paper is to use an analogous approach for investigating the influence of different strategies and levels of maintenance on a producing company’s profit. Our approach provides a step in the direction of quantifying the relationship between the level of maintenance of a given technical asset, and the productive income generated through it, in financial terms. The central idea is to use the principles of Unified Financial Analysis and extend them in such a way that it encompasses the cash flow
patterns created in relation to physical assets. In this respect, we will present financial analogies of the cash flow stream generated by operating a physical asset and the different maintenance strategies that may be used. The different maintenance strategies can be viewed as an extension of financial contracts. In this way the future cash flow pattern associated can be derived, and described in terms familiar to financial management. In this process, it will also become clear which parameters must be estimated to get quantitative results, and how the data must be organised such that parameter estimation becomes possible.

This paper is organized as follows. Section 2 summarizes the maintenance context. This is then followed by a short description of the Unified Financial Analysis (UFA) principles (section 3). The analogy of cash flow patterns associated with physical assets and maintenance, on the one hand, and financial cash flow patterns on the other hand are the subject of section 4. Section 5 concludes the paper.

2. THE MAINTENANCE CONTEXT

2.1 The problematic of Maintenance Budgeting

Maintenance decision-makers have a perpetual problem of obtaining satisfactory budget funds. They need to optimise the impact of maintenance on the company’s profit through practicing optimal levels of prevention\(^3\).

Business decision-makers, on the other hand, mostly see maintenance only as a necessary expense. They perceive maintenance to have no relevance towards the business outcomes, apart from keeping production machinery running when it fails or are busy failing. In particular, the benefit of preventive maintenance is not adequately perceived and valued.

The business can thus never obtain optimal results due to this dichotomy in understanding the role of maintenance towards the business outcomes. On the one hand, maintenance decision-makers understand that better business outcomes are possible through practicing relatively high levels of prevention, without however knowing the probable amount of additional income that is due to such a maintenance policy. Conversely, business decision-makers believe they can achieve better results through the limitation of maintenance expenditure - without being aware of the probable loss of income that this is likely to cause.

It is clear to both parties that there is a direct relationship between maintenance expenditure and business profit. Maintenance expenditure is both an expense, decreasing the business’ profit, and a profit-earning measure through it achieving additional income through higher levels of Operational Readiness of manufacturing equipment.

2.2 The objective of maintenance expenditure

There are at least four reasons for maintenance expenditure in its quest towards keeping operating equipment operational:

\( a) \) The most obvious reason is that one aimed at repairing the equipment following failure. Sometimes such expenditure might also follow on indisputable signs that such equipment is about to fail. These funds are aimed at restoring operational capability, following a loss of such capability, and constitute reactive maintenance expenditure.

\(^3\) Prevention is a combination of basic care (adjustment settings, lubrication, and awareness of adverse signals), Condition Measurement, and Use Based interventions. Its purpose is to prevent the negative effects of failure on the business outcomes as far as is possible.
b) The second use of maintenance funds is to provide basic equipment care. Equipment that are not regularly serviced, lubricated, adjusted, etc. will fail more often. Such maintenance expenditures are the most basic form of preventive maintenance expenditure. It is consequently the most basic part of the level of prevention practiced by maintenance.

c) Thirdly, money spent to predict failure before it occurs. One of the best ways of improving the level of prevention is by taking action when some physical parameter indicates an impending risk of failure. Such expenditure takes on the form of the purchase of Condition Monitoring Equipment, or purchasing an outside service for achieving such pre-emptive warning.

d) Another means of prevention is practiced through use-based replacements or reconditioning of components based on their historical incidence of failure. The use of such historical evidence, combined with the effect of the ratio of the cost of failure to the cost of prevention can produce huge profit gains.

2.3 The effect of the level of prevention on the profit of the business

The level of prevention has at least three effects on the profit of the business. These include the cost of the maintenance, the effect on Operational Readiness, and the direct effect on the profit of the business.

2.3.1 Achieving lower maintenance cost

When practising a low level of prevention, most maintenance interventions will be breakdowns, resulting in a high cost of maintenance, cf. Fig. 1. Because breakdowns are unplanned events, with a high incidence of secondary damage, the maintenance costs at low levels of prevention are particularly high. On the other hand, a high level of prevention will result in the cost of breakdowns to be low.

Naturally prevention also comes at a cost. Thus, the cost of prevention will be low when practising a low level of prevention, cf. Fig. 1, green line. On the other hand, the cost of prevention will be high when practising a high level of prevention. The particular shape of the graph of the cost of prevention will depend on the specific mix of prevention practiced. For simplicity sake we assume a straight line relationship.

The total cost of maintenance is given by the addition of the cost of breakdowns and the cost of prevention. This results in a graph with the very distinct minimum, cf. Fig. 1, blue
It should be clear that, if cost was the only consideration, that level of prevention resulting in the minimum total cost of maintenance will be practiced. However, there are other factors, which in most cases play a more significant role, as follows.

2.3.2 Achieving a high level of Operational Readiness

The second effect of the Level of Prevention is that it increases the level of Operational Readiness\(^4\) of the equipment of the business [2]. When no prevention is practised, a relatively low level of Operational Readiness exists, cf. Fig. 1b. As the level of prevention is increased, the level of Operational Readiness increases, first quite dramatically before flattening out. This increase in Operational Readiness sets the scene for increases in production output and thus business income.

2.3.3 Achieving higher Operational Profit

The following picture shows the classical profit model of an organisation that manufactures and sells goods for profit. The total cost of manufacturing consists of a combination of fixed cost and variable cost. A part of this cost is maintenance cost (both fixed cost and variable cost). If the quantity Q is manufactured, the profit as shown will be the result.

![Figure 2 - Profit impact of the Level of Prevention](image)

If the maintenance department achieves a higher level of operational readiness through practising a higher level of prevention, then it will be possible for the production function to increase the productive output of the business. Such increase is shown by \(\Delta Q\) in figure 2. This will result in an increase in profit (shown by \(\Delta P\) in the figure). One can then say that the extra profit is the result of the increased prevention practised by the maintenance department\(^5\).

---

\(^4\) Operational Readiness includes Availability, Reliability, and Operability of equipment, as well as Quality of manufactured product.

\(^5\) The increase in the level of prevention will of course come at the price. This will result in the gradient of the total cost line fractionally increasing, causing a small decrease in the value of \(\Delta P\) as shown in the figure. This increase in cost is so small when compared to the increase in income, that it can be neglected.
It was shown above that maintenance has two ways of impacting on the profit of the business. These were shown to be the cost of maintenance itself, and its positive effect on operational readiness (leading to the resultant increase in profit resulting from the use of such increased operational readiness). The level of prevention practised is the source of both these effects. So, while a level of prevention exists that will result in the lowest cost of maintenance, the benefit of practising a high level of prevention will in general be the better option. The increased profit resulting from such higher maintenance input will have a substantial benefit for the organisation, when compared to the option of driving maintenance costs down. However, computing the additional benefit is by no means easy because the important effect - the increase of income - is indirect: It consists in computing the value of a kind of real option, which is a difficult task that up to now has not been undertaken. Before undertaking this task we will recast the maintenance situation in the language of modern financial risk and profitability as exposed in any textbook on modern finance (e.g. [5]).

3. MAINTENANCE FROM THE POINT OF VIEW OF MODERN FINANCE

From the point of view of finance, a plant is an option that gives the owner the right but not the obligation to produce some products through operating it. This requires the plant to be operationally ready, which in turn requires maintenance. Without maintenance the plant would continuously degrade: its operational readiness would decrease to zero, the option represented by the plant (and called “plant option” in the following) would lose its value. This would happen even if the plant is not operated at all. One can thus say that maintenance is the part of the option premium that has to be paid continuously (in addition to the initial investment) for the plant option to keep its value. This is one important difference compared to standard financial options. Another important difference is that the value of the plant option depends on the level of maintenance, i.e., on the amount of the continuous part of the option premium the plant’s owner is willing to pay. From an accounting point of view this is part of the (fixed and variable) costs.

The expected pay-off of the plant option during the time period $[0,T]$ is a cash flow stream that can be written as follows:

$$E \left[ \int_0^T \max[I(t) - C_v(t), 0] \, dt \right]$$

Here, $I(t)$ signifies the revenues and $C_v(t)$ the variable costs. Both quantities depend on the volume produced, and on additional conditions, such as the price of raw materials and the selling price of sold products. The option is exercised (the plant is operated), whenever the difference between income and variable cost is positive. Since the pay-off flow is stochastic, we have to compute its expected value.

To value this option, the pay-off flow must still be discounted properly. In the present context this is best done by using a deflator, $D(t)$, which is also stochastic$^6$. The value of the option is then given by:

$$R = E \left[ \int_0^T D(t) \max[I(t) - C_v(t), 0] \, dt \right]$$

$^6$ Valuing cash flow streams using a universal deflator is exposed on a technical level in detail in [6]. For a summary for non-specialists the reader is referred to [1], section 4.5.
where the assumption $D(0) = 1$ has been used.

In Eq. (2) risk is not taken into account because for pure financial options this is not needed. However, here it should be done, and can be done by adding an explicit risk premium $C_{\text{Risk,F.O.}}$ as costs. This quantity measures the financial risk of the pay-off stream caused by plant failure or breakdown.

The variable costs correspond to the option’s strike price. Running the plant is reasonable whenever $l(t) > C_0(t)$, i.e., whenever revenues exceeds variable costs. However, this is valid only for the short term. To be profitable in the long term, revenues should exceed variable costs plus fixed costs plus maintenance costs, which means that the option value $R$ should be larger than fixed costs plus maintenance costs.

To maximize profit one has to maximize the difference:

$$\max \{R(S_M, V) - C_F - C_M(S_M)\}$$

where $S_M$ indicates the maintenance strategy. Also on the cost side, a risk premium $C_{\text{Risk,Maint.}}$ can be added, which is a measure of the cost-related uncertainty.

The important point here is that while the maintenance costs $C_M$ are directly visible (and reported as such), the value of maintenance enters the calculation through an increase in value of the plant option. Valuation of this option is thus the central problem. In this respect the most important aspect here in the present context is the dependence of the option value on the maintenance strategy, $S_M$. We emphasize that it's not simply dependence on maintenance expenditure but on the maintenance strategy. Different maintenance strategies are associated with different cash flow patterns, consisting of the cost flow pattern part and the plant option pay-off part.

This means that it is essential to find a convenient way to discriminate between different maintenance strategies, and a convenient parameterization of such strategies. This will be done in the next paragraph according to the principles of Unified Financial Analysis [1].

### 4. UNIFIED FINANCIAL ANALYSIS AND MAINTENANCE MANAGEMENT

#### 4.1 Unified Financial Analysis: A contract centred approach towards financial analysis

Due to the increase in complexity of the financial world, it is no longer sufficient to manage finances through a double-entry bookkeeping system. The risks inherent to modern finance makes it necessary to take a much more sophisticated view towards financial management, driven among others by the requirement of national regulars, and on the international level by the so-called Basel agreements. Accordingly, a number of commercial risk and profitability management systems are available, which together are able to simulate, analyse and value the future cash flow pattern of a financial institution. Of importance here are the principles or foundations of such an approach, which are exposed in a rigorous way in [1] under the name of “Unified Financial Analysis”.

The central idea is to start with the financial contract and to introduce standard contract types for efficient and transparent encoding of future cash flows. This is quite natural and obvious for a financial institution because the (legal) financial contract (as bonds, stocks, futures, options, etc.) is at the centre of its business. The cash flow patterns are derived from the contracts in interaction with the environment, which consists of the market, counter parties, and additional statistical rules named behavioural elements. They are all
linked to the individual contracts. In addition, costs which occur at the level of the whole company have been added as a fifth input element, cf. chapter 7 in [1]. Thus, altogether there are the following input elements:

a) Financial Contracts - represent contractually binding agreements between two counter parties that govern the exchange of cash flows (time and amount).

b) Risk Factors - Many financial contracts contain clauses with reference to market conditions. There are two subgroups of risk, namely Market Risk (interest rate, exchange rates, stock and commodity prices) and Insurance Risk (frequency and severity of claims, mortality rates).

c) Counter parties - Financial Contacts are promises to exchange cash flows according to some patterns. Whether the promises can be kept depends on the standing or rating of the counter parties. If a counter party fails to keep its promise a so-called default occurs resulting in a financial loss. The associated risk is called Counter Party Risk or Credit Risk.

d) Behavioural Elements - Contracts contain rules concerning the exchange of cash flows. However, there are particular rules that affect cash flows which can be observed only statistically. Examples are contacts with an undefined cash flow profile (i.e. savings accounts), mortgages where the debtor has the legal right to prepay, and bonuses in life insurance contracts. Insurance risk factors (frequency and severity of claims, mortality rates) are also modelled in this section.

e) Costs - Costs and revenues occurring at the level of the whole companies through central services, etc. They can be divided into fixed costs and variable costs. In view of a faithful profitability analysis, costs should be attributed to individual contracts whenever possible.

From these input elements, financial events are derived at an intermediate level. The evaluation of financial events means that the information encoded in the input elements is unfolded onto the time line. From the financial events the cash flows are computed; they then provide the basis for the Elements of Financial Analysis, namely liquidity, value and income, risk and sensitivity analysis:

a) Liquidity - money and financial instruments form the counter flow to the flow of goods. Money, or cash flow, is the only “visible”, “touchable”, or “tangible” artefact of finance.

b) Value and Income - double-entry bookkeeping added two new concepts to the simple cash statement, namely value and income (or profit), which is the change in value over time.

c) Risk and Sensitivity Analysis - in the early 1970s, the savings and loans crisis and the invention of financial options created the need for more forward-looking and more efficient analytical methods. Whereas risk analysis introduces the notion of distributions, sensitivity analysis attempts to quantitatively answer questions concerning future outcomes.

The analysis method described so shortly above results in financial institutions being able to manage the inherent risk of financial business well. What is envisaged in this paper is to use the same inherent principles, and to extend them if necessary, to manage the risk inherent to the maintenance business in an organisation that produces goods for selling, or produces a service of some kind. In particular, the risk involved in spending money (or better: in not spending enough money) to maintain equipment needs to be made transparent to upper management.

4.2 Maintenance in the framework of Unified Financial Analysis

The aim of this section is to outline how the financial description of maintenance activities can be integrated into the UFA framework sketched above. Although actual maintenance
relationships have been the basis of what follows, much of the detailed analysis still has to be worked out in sufficient detail to obtain a model that could serve as the major vehicle for communication between maintenance management and upper management. The envisaged structure will directly follow the UFA methodology.

According to section 3, the central part of the financial modelling of the effects of maintenance is the plant option. In the UFA framework it should be represented at the contract level as a plant option contract. Such a contract is an extension of the existing UFA framework.

Apart from the initial investment the cash flow stream produced by this option consists of two parts:

- the option premium, i.e., maintenance expenditures
- the pay-off consisting of
  - revenues generated by operating the plant
  - the variable operating costs (raw material plus operating staff)
  - fixed operating costs that occur only when the plant is operated.

Additional fixed costs that also occur when the plant is not operated (but are not part of the maintenance costs), must be linked to the plant option contract but not considered when evaluating the option pay-off.

### 4.2.1 The option pay-off

A key quantity for evaluating the pay-off is the volume produced, because the revenues and the variable costs depend on this quantity. The revenues and the cost of the raw materials are computed by means of the market prices of the end product and the raw materials, respectively. Market prices are modelled in the risk factor section. The salaries of the operating staff are modelled in the behaviour section. Fixed operating costs can be modelled directly as part of the contract as occurring whenever the plant is operated.

The volume produced is thus a key quantity. It depends on:

a) Total Possible Production Time
b) Percentage of Time that the Plant is Not Utilised
c) Production made per hour of production time

Another way to represent the volume produced is by means of the maximum possible volume and Operational Readiness. The reason is that Operational Readiness is the major output of the maintenance function. It is this quantity that is to be increased through an increase in the option premium.

The important element is to connect Operation Readiness to the level of maintenance provided through the option premium. Even though this requires knowledge of the system under investigation a necessary prerequisite is the discrimination between different modes of maintenance. Only this makes it possible to clearly work out the beneficial effect of preventive maintenance, and to determine the optimal maintenance level.

If necessary, Operational Readiness can be subdivided further. It is based on the Availability of equipment, as well as the equipment’s operating capability (Operability), and its ability to produce quality output (Qualibility). The Availability of equipment in turn consists of the elements Reliability (MTTF), and Maintenance Reaction Time (MTTR). Operational Readiness determines the ability of the plant (or business) to produce a high output of the manufactured product (or service).
4.2.2 The option premium

The option premium consists of maintenance expenditures. In a global way they can be characterized by the value of the maintenance budget (annual % of capital value). However, without giving details of how much of the maintenance budget is spent for the different modes of maintenance, no further analysis of the optimal maintenance strategy is possible.

A mixture between reactive and preventive maintenance can be described, e.g., by the level of prevention, which influences both maintenance costs and operational readiness.

The Level of Prevention is measured through the quantities Scheduling Intensity (the intensity with which the organisation practices prevention in preference to reactive maintenance), and Breakdown Severity (the impact of breakdown cost on the total cost of maintenance).

From the specific maintenance strategy (mixture of preventive and reactive maintenance, etc.) the Operational Readiness can be derived by means of the following intermediate elements through a plant-specific model:

a) % Uptime
b) Number of Downtime Occasions

4.2.3 Maintenance Risk

While Operational Readiness provides a measure of the direct influence of maintenance on the productive process, the inherent risk of downtime related losses should always be kept in mind as well. Maintenance Risk has at least two components, i.e. Repair Risk, and Production Loss Risk. These risk elements need to be evaluated with the objective of minimising the possibility and the impact of unforeseen events on the productive process, and thus the gross profit. Since this is not a market risk but specific to a given production plant, these risk factors should be modelled by suitable statistical models in the behaviour section. Their integration into value and premium of the plant option has been described in section 3.

5. CONCLUSION

By modelling and analysing the plant option along the lines presented above it will be possible to evaluate the extra profit indicated in Fig. 2 that can be achieved through good maintenance. This extra profit is a function of the external environment consisting of markets prices, salaries, business cycle, etc. In addition, plant-specific effects must be taken into account. The analysis of maintenance in the context of the plant option explicitly considers the two parts. The first regards the direct cost effect (Maintenance Cost Intensity), and the second the benefit of the money spent on maintenance in terms of a decrease in the level of lost production (Maintenance Improvement Justification). They are reframed as the premium and the pay-off of a real option, namely a plant option, which consists in the possibility to produce some products by operating a plant.

The plant option pay-off and the option premium described above provide all the necessary information to give the business’ upper management and macro tool for maintenance budget management. However, the specific mechanisms of this “macro tool” have to be carefully researched and defined. The usefulness of this instrument will be very much dependent on the quality of this research. Also, both option premium and option pay-off have to be refined further to increase their usefulness, apart from the fact that practical ways have to be found for their evaluation. This should be done in the light of modern
Financial maintenance modelling in the framework of the unified financial analysis methodology - Coetzee, Breymann

maintenance approaches, such as RCM, in which the critical failure modes are determined using an FMECA analysis.

A large part of the challenge in building a maintenance budget evaluation tool that will suffice for a proper level of communication between maintenance management and upper management resides exactly here. A lot of thought and research have to be expended in this area to enable one to devise a method that can be encapsulated in a standard software interface. In any case we believe that the UFA methodology provides the right organizing principles on the macro level.

4. REFERENCES


REFINING THE OPERATING MODEL CONCEPT TO ENABLE SYSTEMATIC GROWTH IN OPERATING MATURITY

Marné de Vries†, Alta vd Merwe‡, Aurona Gerber§, and Paula Kotze¶

†Department of Industrial Engineering
University of Pretoria, South Africa
Marne.devries@up.ac.za

‡Meraka Institute
CSIR, South Africa
Avdmerwe3@csir.co.za

§Meraka Institute
CSIR, South Africa
AGerber@csir.co.za

¶Meraka Institute
CSIR, South Africa
PKotze1@csir.co.za

ABSTRACT
To stay competitive, enterprises of today need to rely on a sound foundation for execution that incorporates the infrastructure and digitised processes for automating a company’s core capabilities. Once this foundation has been established, management could move their attention away from focusing on lower-value activities to innovative ways to increase profits and growth. The Business-IT Alignment Framework (BIAF) defines business-IT alignment in terms of a paradigm of alignment, three dimensions for alignment, and mechanisms and practices. The BIAF could provide a business-IT alignment perspective on the foundation for execution approach. Using the BIAF perspective, this paper comments on some of the deficiencies related to the foundation for execution approach regarding the systematic identification of opportunities for enterprise-wide process standardisation. The goal is to define a list of requirements that should direct the design of appropriate mechanisms and practices to address the identification of process re-use opportunities for multiple levels of operating maturity.

OPSOMMING
Om kompeterend te bly, benodig organisasies vandag ’n goeie fondasie van uitvoering. Hierdie fondasie van uitvoering behels die infrastruktuur en gedigitaliseerde prosesse wat die onderneming se kern bekwaamhede automatiseer. Eers wanneer die fondasie gevestig is, kan bestuur hul aandag verskuif van laer-vaarde aktiwiteite en fokus op innoverende wyes om winste en groei te verhoog.

Die Besigheids-IT Belyningsraamwerk (BIBR) definieer die besigheids-IT belying in terme van ’n paradigma van belyning, drie dimensies van belyning, en mekanismes en praktyke. Die BIBR sou ’n besigheids-IT belyningsperspektief kon gee vir die fondasie van uitvoering benadering. Hierdie artikel maak gebruik van die BIBR perspektief om kommentaar te lever op die tekortkominge van die fondasie van uitvoering benadering rakende die sistematiese identifisering van onderneming-wye prosesstandaardisasie geleenthede. Die doel is om ’n lys van behoeftes te identifiseer wat sal lei tot die ontwerp van geskikte mekanismes en praktyke vir die identifisering van proses her-gebruikgeleenthede vir veelvuldige operasionele volwassenheidsvlakke.
1. INTRODUCTION

According to the 2007 survey by Luftman and Kempaiah [1], both IT and business alignment and attracting, developing, and retaining IT professionals have consistently been the major IT management concerns since 1994. Since the early 1980’s, Business-IT alignment has been an important challenge in both private and public/non-profit sectors (Knoll and Jarnvenpaa, [2]). According to Luftman’s definition of business-IT alignment [3], there is strong evidence of a link between business-IT strategy alignment and organisational performance (Luftman and Kempaiah, [4]). However, Luftman and Kampaia [4] point out that organisations still face significant challenges in aligning business with IT. Still an art form rather than a science (Wegman, [5]), there is still no silver bullet for aligning business with IT.

The Business-IT Alignment Framework (BIAF) (De Vries, [6]) defines business-IT alignment in terms of a paradigm of alignment, three dimensions for alignment, and mechanisms and practices. The BIAF could provide a business-IT alignment perspective on the foundation for execution approach of Ross, Weill and Robertson [7]. Using the BIAF perspective, Ross et al. [7] have a specific approach towards business-IT alignment. This approach has two key artefacts: the operating model (OM) and core diagram (CD), which are used in combination with four stages of operating maturity. The OM and CD communicate the enterprise-wide vision for process standardisation and data centralisation, with the aim to define long-term process and data rationalisation decisions that may be implemented enterprise-wide. These decisions direct the evolution of the IT landscape only up to the third (optimised core) of the four levels of operating maturity. Once the third level of maturity has been achieved, companies typically implement a core set of systems to support standard enterprise processes, whilst the unique needs of business units are not accommodated. The fourth level of operating maturity should enable an enterprise to preserve global standards, while enabling local differences. The fourth level of operating maturity thus requires a business modularity architecture, which accommodates the unique needs of business units. In addition, previous action research revealed deficiencies in deriving the key artefacts (OM and CD) of the foundation for execution approach.

At present, there are no requirements for the set of practices and mechanisms that are needed to address the identification of process re-use opportunities for both the third and fourth levels of operating maturity. The goal of this paper is therefore to extract such a list of requirements for the identification of process standardisation opportunities in an organisation. These requirements should direct the design of appropriate mechanisms and practices that enables systematic identification of opportunities for enterprise-wide process standardisation.

The paper is structured as follows: In Section 2 we provide some background on the Business-IT Alignment Framework (BIAF) and the foundation for execution approach. Section 3 relates the foundation of execution approach with the BIAF to highlight alignment deficiencies. Additional deficiencies are also identified through an action research approach executed. In Section 4 the operating model concept is extended and the set of requirements are identified that should address the systematic identification of opportunities for enterprise-wide process standardisation/replication. The paper is concluded in Section 5.

2. BACKGROUND

There are various alignment theories, each with its own alignment paradigm, combination of alignment dimensions, mechanisms, and illustrative case studies. Yet there are very few case studies to demonstrate the integrated use of these theories. De Vries [6] introduced a BIAF to unpack business-IT alignment in terms of generic approaches, dimensions,
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

mechanisms and practices. This section discusses the use of the BIAF to provide a business-IT alignment perspective on the foundation for execution approach of Ross et al. [7]. From this perspective, various deficiencies of the foundation for execution approach will be discussed. The deficiencies provide the rationale for the identification of a list of requirements that could address the deficiencies.

2.1 The Business-IT Alignment Framework

Alignment endeavours/programmes/projects are usually founded on defendable value propositions. These value propositions are based on certain belief systems about value-creation in an organisation and the capability of marketing the propositions to the owners/funding parties of the organisation. With reference to Figure 1, the BIAF starts with the foundation part (alignment belief/paradigm of creating value), which determines the business-IT alignment approach with reference to the other BIAF parts (De Vries, [6]).

A selection of alignment mechanisms and practices (e.g. methodologies, processes, methods, tools, governance structures) are required to facilitate change/evolution to re-align business with IT. The selected alignment mechanisms and practices are related to one or more BIAF dimensions. The framework contains three key alignment dimensions, represented by three panes in Figure 1:

- Architecture abstraction layers;
- Perspectives/stakeholder viewpoints; and
- Organising scope.

![Figure 1 - BIAF (De Vries, [6], adapted)](image)

The dimensions part provides the context of alignment, answering the questions, ‘What needs to be aligned?’ , ‘To what extent?’ and ‘From which perspective(s)?’. The alignment mechanisms and practices part provides the means for alignment.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

In support of the value-creation paradigm and scope of alignment, the mechanisms and practices that are selected could be further classified according additional classifiers. The additional classifiers relate to (1) the version or versions of alignment (current state / future state), (2) the starting point for doing architecture work (top-down, bottom-up or middle in), (3) the alignment frequency (periodic vs continuous) and, (4) different ways of addressing the changing/dynamic nature of the alignment components.

De Vries [6] used BIAF to demonstrate the interpretation of the framework for four prominent Enterprise Architecture (EA) models: the Zachman Enterprise Framework (ZEF), The Open Group Architecture Framework (TOGAF), the Federal Enterprise Architecture (FEA), and the Gartner Enterprise Architecture Framework (GEAF). Each of the analysed models has its own paradigm of value-creation and different demarcations of the three BIAF dimensions, if demarcated at all. The alignment mechanisms and practices are supposed to support the value-creation paradigm and the extent of alignment, as defined and demarcated by the specific EA model.

It is possible to compare the different EA models in terms of various factors, such as the efficiency of demarcating the three dimensions, or the ability of the EA models to address the scope defined by the demarcated dimensions via the alignment mechanisms and practices. Zachman for example, believes that contrary to most other EA models; his ontology provides a scientific approach in defining architecture abstraction layers (inventory sets, process transformations, network nodes, organisation groups, timing periods and motivation reasons) and perspectives (scope contexts, business concepts, system logic, technology physics, component assemblies and operations instance classes) (Zachman [8]:20). Scrutiny of other models reveals considerable overlap between the different abstraction layers, which emphasises the need for an ontology that provides one fact in one place. Although the ZEF provides an ontology for defining suitable mechanisms and practices for doing EA work, Zachman [8] is not prescriptive about a required set of mechanisms and practices. These depend on the project team using the ZEF (O’Rourke, Fishman and Selkow [9]). On the contrary, the Architecture Development Method (ADM) of the Open Group Architecture Framework (TOGAF) defines a rich set of mechanisms and practices that support enterprise-wide business-IT alignment (The Open Group [10]). Although the Open Group defined their own Content Framework to demarcate the three BIAF dimensions, the ADM could also apply the dimensions defined by the Zachman Enterprise Framework (The Open Group [10]:56).

2.2 The foundation for execution approach

The foundation for execution approach [7] aims to rationalise and digitise both the routine, everyday processes and competitively distinctive capabilities of an organisation. Ross et al. [7] recommend eight steps in creating a foundation for execution. During the first three steps, key artefacts are defined, which should be used to establish technology, data, and process rationalisation objectives for an organisation. The authors used BIAF to contextualise and evaluate the foundation for execution approach in terms of business-IT alignment and identify deficiencies inherent in the approach and its artefacts.

3. THE FOUNDATION FOR EXECUTION APPROACH RELATED TO BIAF

The foundation for execution is “the infrastructure and digitised processes automating a company’s core capabilities” (Ross et al., [7]:4). In terms of the BIAF, the value-creation paradigm of this approach is that an enterprise requires a foundation for execution to digitise operational processes. This will free up management time to focus on strategic issues.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

Relating to the additional classifiers of an approach, the foundation for execution approach focuses mainly on the future state architecture. Ross et al. ([7]:44) believe that a company needs to articulate a vision (future view) of how the company will operate, called the operating model (OM). This vision needs to provide direction for building a foundation for execution. Once a vision is created, business and IT leaders define key architectural requirements of the foundation for execution, which may be communicated via a core diagram (CD). The OM and CD are used in combination with four stages of operating maturity to elevate an organisation to a higher level of operating maturity. An IT engagement model is used to ensure that new business initiatives contribute to the foundation for execution. Continuous alignment takes place to evolve the architecture. A top-down approach (starting at the contextual perspective, translating through subsequent perspectives) is followed in terms of architecture development, emphasising the contextual perspective (see Figure 2). Although the foundation for execution approach aims at reducing architectural complexity by rationalising data and processes according to the OM requirements, the mechanism and practices do not explicitly address the problems associated with the changing/dynamic nature of architecture components.

3.1 Dimensions of alignment

Ross et al. [7] do not stipulate different architecture abstraction layers, perspectives or organisation entities to demarcate the three BIAF dimensions. Hence using the ZEF as an ontology for defining the architecture abstraction layers and perspectives/stakeholder viewpoints, the foundation for execution approach emphasises two main architecture abstraction layers: data (WHAT: inventory sets) and process (HOW: process transformations), and one perspective: scope contexts (see Figure 2).

The main contribution of Ross et al. [7] is to define on a contextual level the data (WHAT: inventory sets) that could be shared and the processes (HOW: process transformations) that could be replicated across different business units (shaded in Figure 2). Ross et al. purposefully omit alignment with the motivational aspects (WHY: motivation reasons) of the business (see Figure 2). The rationale is that strategic initiatives, derived from the strategic direction, often lead to IT-enablers for each strategic initiative. This creates the delivery of piece-meal IT solutions that are not integrated (Weill & Ross [11]). The IT department constantly reacts to the latest strategic initiative and is always a bottleneck, operating in a reactive mode.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

Figure 2 - The foundation for execution (Ross et al.[7]) alignment focus

3.2 Alignment mechanisms and practices

Building a foundation for execution requires mastery of three key disciplines, the operating model, Enterprise Architecture and operating maturity as well as the IT engagement model. In the following section, the three disciplines are discussed in context of the foundation of execution approach.

(1) The Operating Model

The OM is the “necessary level of business process integration and standardisation for delivering goods and services to customers” (Ross et al. [7]:8). The level of process integration relates to the extent to which business units share data. Process standardisation is used to create efficiencies across business units, but limits opportunities to customise services.

Based on the different levels of process standardisation and process integration, Ross et al. [7] define four types of OMs. These OMs are not only dependent on the levels of process standardisation and integration, but are defined based on a group of characteristics. Figure 3 gives an overview of these characteristics.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

Figure 3 - Characteristics of four operating models (Ross et al. [7]:29)

(2) Enterprise Architecture and Operating Maturity

According to Ross et al. ([7]:9) EA is the “organising logic for business processes and IT infrastructure, reflecting the integration and standardisation requirements of the company’s operating model”. A core diagram (CD) is used to provide a graphical representation of the required enterprise architecture (EA). The CD should be used to:

- Facilitate discussions between business and IT managers to clarify requirements for the company’s foundation for execution, and
- Communicate the vision (high-level business process and IT requirements of a company’s operating model).

An example of the unification OM is given in Figure 4. As a unification OM requires high levels of process standardisation and process integration (data sharing), the standard (core) and linked processes, and shared data are depicted on the diagram. The diagram also depicts key customer types and automating technologies.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

Key customers
Linked and standard (core) processes
Shared data
Linking and automating technologies

Figure 4 - Core diagram template for a unification OM (Ross et al., [7]:54)

Ross et al. ([7]:71) believe that organisations need to follow a systematic transformation process in changing towards the future EA, as required by the OM. Companies should build out their enterprise architectures through four stages of operating maturity:
1. **Business Silos architecture**: where companies maximise individual business unit needs or functional needs.
2. **Standardised Technology architecture**: gaining IT efficiencies through technology standardisation and increased centralisation of technology management.
3. **Optimised Core architecture**: providing companywide data and process standardisation, appropriate for the OM.
4. **Business Modularity architecture**: where companies manage and reuse loosely coupled IT-enabled business process components to preserve global standards while enabling local differences.

(3) **IT Engagement Model**

The IT Engagement model portrays the set of governance mechanisms that are required to ensure that business and IT projects achieve both local and companywide objectives. Ross et al. [7] state that top performing companies provide the following mechanisms and practices as part of their IT Engagement Model:
1. **Company-wide IT governance**, defined as the “decision rights and accountability framework to encourage desirable behaviour in using IT” (Ross et al. [7]:119). Company-wide IT governance should be based on five types of interrelated IT governance decisions (Weill & Ross [12]). The OM has profound influence on one of these decision types: the IT principle decisions. IT principles decisions need to delineate how IT will support the desired OM.
2. **Project management**, which requires a formalised project methodology with clear deliverables and checkpoints.
3. **Linking mechanisms**, which incorporates processes and decision-making bodies that need to align incentives and connect the project-level activities to the companywide IT governance.
3.3 Partial validation of the foundation for execution approach

In order to receive feedback on the perceived practicality of the two key artefacts, the OM and CD, we were engaged in an action research activity (also reported on in [13]). The purpose was to receive qualitative feedback on the difficulties experienced in defining the current OM and CD for an organisation / sub-division. The following interpretations of difficulties with the OM and CD were identified (De Vries, [13]):

- Difficulty in selecting a single operating model is linked to the identification of the degree of process standardisation / integration for the analysed organisation/business unit. Extensive implicit/explicit knowledge is implied during the evaluation of the OM characteristics that define the degree of process standardisation / integration.
- Difficulty in finding the correct information to perform an OM classification or select core diagram components. Identification of OM characteristics and core diagram components require knowledge about the strategic choices (markets, products/services), operating/organising logic, business processes and main databases & technologies of the organisation. Some baseline architectures are thus required and this knowledge is not necessarily available or in an explicit format.
- Respondents experienced difficulty in selecting the main components of the core diagram & understanding the core diagram templates. This may be related to the limited set of examples provided in the textbook. Case studies would be required to demonstrate inputs that would be required (e.g. baseline architectures) to define the core diagram components.

3.4 Identifying additional deficiencies

Based on the qualitative feedback received from the action research effort, the authors revisited the purpose/use of the operating model, summarised as follows:

- Senior managers need to debate and select an appropriate OM to establish a vision for how the company will operate (Ross et al. [7]).
- The OM is also a “choice about what strategies are going to be supported”. The choice of an OM drive the implementation of a whole set of strategic initiatives (Ross et al. [7]:26).
- The OM should be used to direct IT principles decisions (Weill & Ross [11], Weill & Ross [12]).
- Each OM enables different strategies for achieving growth and profits (Weill & Ross [11]).
- The enterprise OM should be used to establish interoperability requirements. According to The Open Group ([10]:331), the corporate OM “will normally indicate what type of interoperability approach will be appropriate” and should be determined during the Architecture Vision or Business Architecture phases of the Architecture Development Method (ADM). The OM provides a vision of what needs to be shared.

If senior managers are to use the OM to guide them during the strategic decision-making processes, the artefacts should be based on a more rigorous approach in attaining the artefact outputs. The author consequently revisited the OM and made some critical evaluations. In our evaluations there were significant deficiencies related to method and elevating to a fourth level of operating maturity. These deficiencies are discussed further in 3.4.1 and 3.4.2 respectively.

3.4.1 Method deficiency

The characteristics of the OM (see Figure 3) could be classified according to different categories, which imply different timings. The characteristics relate to:
Current business architecture configurations that pose opportunities for sharing data and replicating similar processes/functions (e.g. shared customers/products/suppliers; operationally unique business units or functions).

Shared data and standardised processes (e.g. shared customer/supplier/product data; standardised processes).

Suggestions in terms of business and IT governance arrangements that go hand-in-hand with the other characteristics (e.g. autonomous business management; IT decisions made centrally).

An implicit process is thus suggested to derive a required OM (see Figure 5).

The organisation needs to analyse certain business architecture parameters to establish rationalisation opportunities.

Rationalisation opportunities could be identified within two main areas: (1) Data (sharing data across organisational entities), and (2) Process (replicating/re-using processes across organisational entities). The levels of data sharing and process replication will provide opportunities for sharing certain technologies. A pure coordination OM could use common portals and middleware technology; a replication OM could use common system components; while a unification OM could use common application systems (Weill & Ross [11]).

Once rationalisation opportunities have been established an organisation needs to derive a future OM that would exploit these opportunities.

The future OM then needs to direct the design of appropriate governance mechanisms.
Refining the operating model concept to enable systematic growth in operating maturity

de Vries, van der Merwe, Gerber, Kotze

Analyse business architecture parameters.

Identify opportunities to:

- Means for analysing organising entities, their suppliers, products, customers. Which are shared across organising entities?
- Means for identifying dependencies between organising entities.

Share DATA

Re-use PROCESSES

Decide on the future OM by exploiting sharing/re-use opportunities.

Design IT Governance Mechanisms to support IT decision-making.

Means for identifying similar data entities that may be shared across organisational entities.

Means for identifying and representing similar processes across organising entities.

Deficiencies in elevating to a third level of operating maturity

- Means for analysing organising entities, their suppliers, products, customers. Which are shared across organising entities?
- Means for identifying operational similar organising entities.

Deficiencies in elevating to a fourth level of operating maturity

- Means for identifying and representing similar process components across organising entities.

3.4.2 Deficiency in elevating to a fourth level of operating maturity

Ross et al. ([7]:26) believe that the choice of an OM is a critical decision for a company and that "it’s the first step in building a foundation for execution". Re-visiting the role of the OM in transforming an organisation through different levels of operating maturity however revealed insightful results. Section 3.2 indicates that the OM is only required to elevate a company from a second level of operating maturity to a third level of operating maturity, which is also supported by a more recent publication of Weill & Ross [11], where standardisation objectives are defined for each type of OM as differentiators. The four OMs all require ‘shared services’ and common ‘infrastructure technology’ objectives (objectives for level two operating maturity). Data sharing and process replication objectives differentiate the four OMs from one another and are objectives for reaching the third level of operating maturity. While the third level operating maturity objectives are derived from the OM and exploit rationalisation opportunities across the enterprise, the fourth level of operating maturity acknowledges the unique needs of business units and need to be

Figure 5 - Current deficiencies in defining and using the OM
supported via IT-enabled process components. The OM however does not facilitate the identification of process components that may be IT-enabled and re-used across the organisation (see Figure 5).

4. EXTENDING THE OPERATING MODEL CONCEPT

The left-hand side of Figure 5 portrays the process that could be followed in defining and using the OM. The deficiencies discussed previously are indicated on the right-hand side. Smith & Fingar [14] believe that many companies already seized the opportunity of implementing centralised data management systems (sharing data). Based on the deficiencies highlighted in Figure 5, a set of requirements were derived (see Table 1) to define the scope of supplementing mechanisms and practices. The following requirement categories were addressed:

1. User(s) of the practices and related mechanisms
2. Generality
3. Process categories included
4. Current architecture capabilities
5. Process representation
6. Replication constraints
7. Feasibility analyses

The requirements that have been identified depict the scope for developing applicable mechanisms and practices.
<table>
<thead>
<tr>
<th>No</th>
<th>Category</th>
<th>Requirement Detail</th>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>User(s) of the practices and related mechanisms</td>
<td>Any EA practitioner who wants to use the OM specified by Ross et al. [7] and needs to collaborate with other stakeholders in defining the required level of process standardisation/replication.</td>
<td>The practices and mechanisms are created for the purpose of enhancing the OM concept as defined by Ross et al. [7].</td>
</tr>
<tr>
<td>2</td>
<td>Generality</td>
<td>The practices and mechanisms should be generic in their application to different types of industries. An EA practitioner should be able to apply the practices and mechanisms to either a profit-driven, not-for-profit/government organisation within any industry, in combination with the foundation for execution approach.</td>
<td>The foundation for execution approach is generic in its application.</td>
</tr>
<tr>
<td>3</td>
<td>Process categories included</td>
<td>The practices and mechanisms may be applied to all processes in the organisation.</td>
<td>The foundation for execution approach is based on the paradigm of creating a foundation for execution, which not only focuses on competitive distinctive capabilities, but also rationalising and digitising everyday processes that a company requires to stay in business (Ross et al., [7]:4).</td>
</tr>
<tr>
<td>4</td>
<td>Current architecture capabilities</td>
<td>The practices and mechanisms need to take current work in terms of Enterprise Architecture, Business Architecture and Process Architecture into account, but also need to provide sufficient detail if none of these architectures have been explicated.</td>
<td>According to Ross et al. ([7]:26), the first step in building a foundation for execution is to define the OM for the organisation. No pre-conditions are defined for defining this model. The ability to define this model however is dependent on current architecture capabilities and documented/explicated architectures. Immature architecture capabilities may require additional architecture work, such as defining enterprise-wide process management standards and a centralised process repository (Smith &amp; Fingar ([14]:177).</td>
</tr>
<tr>
<td>No</td>
<td>Category</td>
<td>Requirement Detail</td>
<td>Motivation</td>
</tr>
<tr>
<td>----</td>
<td>----------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
</tbody>
</table>
| 5  | Process representation | The practices and mechanisms should encourage consistent process representation to ensure re-use. The extent of re-use includes the following:  
- It should be possible to add process measures if required for the purpose of performance measurement and/or process improvement.  
- The process representations should support end-to-end views of processes.  
- Process representations should not hamper the transition from the third to fourth levels of operating maturity.  
- The representations that are used to communicate process replication opportunities should be understandable to business users (from the contextual and conceptual viewpoints). | A consistent representation may enhance communication about how the business operates, enable efficient handoffs across organisational boundaries and allow for consistent performance measurement across organisation entities or similar competitors (Davenport [15]). In addition, transitioning from a third to fourth level of operating maturity (as defined by Ross et al., [7]) requires the identification of business services that may be shared among different organisational entities. Heinrich et al. [16] believe that the identification of business services requires a consistent representation of the enterprise’s processes. |
| 6  | Replication constraints | The mechanisms and practices should identify process/functional similarities across different type of entities. The practitioner should receive guidance in deciding on the type of demarcation used. | Well et al. ([11] mention that replication opportunities may be defined across various types of entities (business units, regions, functions and market segments). |
| 7  | Feasibility analyses | The mechanisms and practices should not suggest the means for assessing or measuring the feasibility of process replication/rationalisation. Feasibility analysis, e.g. operational, cultural, technical, schedule, economic and legal feasibility (Whitten [17]) that may be associated with process rationalisation solutions are therefore excluded. | Although a feasibility analysis may direct the required level of process standardisation, this set of mechanisms and practices will merely propose a way of identifying replication opportunities, based on process similarity. |

Table 1: Requirements for addressing the deficiencies highlighted in Figure 5
4. CONCLUSION AND THE WAY FORWARD

The BIAF was used to provide a business-IT alignment perspective on the foundation for execution approach of Ross et al. [7]. From this perspective, the main contribution of Ross et al. [7] is to define on a contextual level the data that could be shared and the processes that could be replicated across different business units. Within this context, current deficiencies have been highlighted. This paper focused on process replication, rather than data sharing, and defined a set of requirements for the systematic identification of opportunities for enterprise-wide process standardisation/replication.

Based on the requirements identified in this paper, a method-building approach will be used to define a set of mechanisms and practices that may be used in combination with the foundation for execution approach. The proposed set of mechanisms and practices will be based on research within the disciplines of Enterprise Architecture, Business Architecture, Process Architecture and Service Oriented Architecture. The added mechanisms and practices should ultimately enhance the practical use of the OM concept.

5. REFERENCES

AN EMPIRICAL STUDY OF FACTORS INFLUENCING PROJECT QUALITY IN THE DESIGN PHASE: A CASE OF BUILDING PROJECTS IN THE DEPARTMENT OF BUILDING AND ENGINEERING SERVICES OF BOTSWANA

E.N.D. Dodoo, C.C. van Waveren and K.-Y. Chan
Graduate School of Technology Management
University of Pretoria, South Africa
corro@up.ac.za

ABSTRACT

There were numerous complaints regarding the quality of building projects ran by the Department of Building and Engineering Services in Botswana. The aim of this empirical study is to explore the inter-relationships between 16 identified factors in the project design phase and the five measures which the design phase takes cognizance of. Involvement of building contractors plays a significant and positive role in four out of the five cognizant measures, namely cognizance of time, value management, project life cycle view and communication management. Moreover, the 16 factors are classified into categories and their impacts on the measures are also explored.
1. INTRODUCTION: MOTIVATION OF THIS STUDY AND ITS AIMS

The Department of Building and Engineering Services (DBES) in Botswana is a merger of the previous Department of Architecture and Building Services (DABS) and Department of Electrical and Mechanical Services (DEMS). This newly merged department, DBES, is responsible for the implementation of government building infrastructure development projects. DBES has a pure project management structure where people working on a project are grouped together under a project manager that report to an executive officer. It appears that the merging of the two departments has not lead to better project quality. There were numerous complaints regarding to DBES building projects. This paper aims at evaluating participants involved in DBES projects against project success factors in order to identify key focus areas. Project success factors often fall into two groups: project planning and subsequent tactical operationalization [1]. In previous research, the project design phase has a greater impact on overall project quality [2]. This paper identified 16 factors during the design phase that may influence the quality of projects. During the design phase, certain aspects are considered to ensure project success, for example, constructability which may reduce the need of revision and changes later on during project execution [3, 4]. This study takes into account five aspects or measures which the design phase of DBES projects should takes cognizance of. Moreover, the underlying relationships (using multivariate analysis) between the 16 factors and the five measures during project design phase will be uncovered. The remainder of this paper is structured as follows: A short description of the DBES in Botswana will be given in Section 2. Section 3 gives a brief background of the factors (during project design phase) that may influence project success and the aspects that project design phase take cognizance of (denoted as cognizant measures). Section 4 describes the research methodology used. The statistical results will be reported in Section 5. The final section will conclude this paper and provide recommendations for managers and future studies.

2. DBES IN BOTSWANA

The Republic of Botswana is situated in Southern Africa, nestled between South Africa, Namibia, Zimbabwe and Zambia. The country is democratically ruled, boasts of a growing economy and a stable political environment. Botswana’s economic progress since independence has been one of the few success stories of the African continent. According to World Bank (in Economic Indicators 2004), twenty years ago, the country was one of the 20 poorest countries in the world. Anderson [5] stated that today, Botswana is considered as the richest non-oil producing country in Africa. Botswana has seen a substantial and consistent growth in its (GDP) with a corresponding increase in development activities, processes and projects in the country. This has led to a consistent growth in the construction industry, stimulated by many government projects initiated by ministries. The increase in the number and complexity of the project being implemented has led to severe strain on the resources of the two departments, namely, Department of Architecture and Buildings Services (DABS) and Department of Electrical and Mechanical Services (DEMS). Projects ran by these two separate departments had huge cost overruns, took too long to complete and were at times perceived to be of poorer quality. The Botswana Government constituted two consortia namely the Organization Capacity Assessment (OCA) Study and Operational Efficiency Improvement (OEI) Study to review the project management procedures and organizational structure with a view to improving project delivery and efficiency.

Following the study, government took decision to merge the DEMS and the DABS. Government considered that merging the two departments would lead to improve project delivery, in particular implementing project within budget, time and meeting the clients’ expectations with high performance. In 2004 the government of Botswana decided to merge the then existing two departments and created a new department now called the
Department of Building and Engineering Services (DBES). DBES organizational structure is a pure project management structure. According to Steyn et al [6], in the pure project or "projectized" structure, people working on a project are grouped together under a project manager that report to an executive officer. They further indicated that the pure project structure is preferred where projects are large and have a long duration. Moreover when a project is of great strategic importance and where it is important to complete the project within a short time, a pure project organization structure should be considered for such a project. DBES employs about 168 professional staff and operates through three-core division, namely Project Implementation Division (PID), Technical Services Division (TSD) and Maintenance Division (MD) and has other division or sections providing support services. According to the mission statement, the DBES defines its vision and mission as shown hereafter.

"The Department will be a model of excellence in the delivery and maintenance of public buildings, engineering infrastructure and related services to the full satisfaction of our stakeholders." (Vision)

"The Department will provide building and related services in partnership with stakeholders to client ministries/departments consistent with National Development Plans and other Government Policy pronouncements. We commit ourselves to local capacity building within time with available resources and to acceptable standards." (Mission)

The merging of the two departments has not lead to any significant reduction in delays and cost overruns of projects and therefore its poor quality. This study will look at the factors that influence project quality in the design phase of building projects with specific focus on projects in the Department of Building and Engineering Services (DBES). A justification for the study of the factors that influence the quality of project in the design phase of projects is that incorrect and defective designs do have an enormous impact on the successful implementation of building projects. The quality of designs has influence on the implementation of the project and therefore the success or failure of the project. To a large extent the satisfaction that the end user derive from a completed building depends on the quality input at the initial stages of the designs. To allow the contractors to transform the design concepts and ideals into physical product, the designers must provide well illustrated drawing and detailed specifications. Any attempt to improve on the quality of designs at the design phase of the project will consequently improve the construction process and finally a completed building that will meet the end-user expectation. In addition the study of the factors that influence the quality of design will create the awareness of quality of design to the project sponsors to allow sufficient project resources at the early stages of the project.

3. RESEARCH FRAMEWORK

There are two parts in this section: firstly, the factors that influence project quality in the design phase of building projects; and secondly the aspects that the design process take cognizance of (denoted as cognizant measures).

3.1 Factors that influence project quality in design phase

Building projects start with the project design phase where it ‘translates the primary concept into an expression of a spatial form which will satisfy the owner’s requirements into an optimum and economic manner’ [2]. During the project design phase the objectives, constraints and specifications of the project are defined by a team of individuals [7]. In building projects many uncertainties could be eliminated during the early stage by addressing the scope of the project, the deliverables, the cost involved, as well as the schedule for delivery [6]. Effective management for project quality starts when the deliverables (that satisfy the need) are determined at this phase [8]. Customer satisfaction
An empirical study of factors influencing project quality in the design phase
Dodoo, van Waveren, Chan

is considered as one of the four project quality pillars that emerge from the conceptual foundation [9] and thus customer / user requirements at the initial stages of projects (i.e. design phase) should be specified or defined clearly to avoid design deficiencies which may cause contract modifications during the implementation of the project [10].

In this study the authors identified three categories of factors in the project design phase that may influence project quality. These categories are: project management principles, involvement of role players, and project scope.

3.1.1 Project management principles

These principles are most often learnt from experience and may be applied to all projects to ensure success. In this study the following principles to be included during the design phase of DBES's projects have been identified:

- DBES's selection process of the design team on meeting the client’s requirements and expectations
- Adherence to design brief from clients
- DBES's commitment to continuous quality improvement
- DBES leadership commitment to the adherence of procedures and specifications at the design phase
- Comprehensiveness of drawings and specifications
- Visibility of the total life cycle cost during the design phase
- Use of lessons learned to improve the quality of designs

3.1.2 Involvement of role players

Inputs from project participants (role players) during the design phase will clarify project scope, planning and scheduling and eliminate uncertainties in the future [11]. Involvements of the following role players during design phase are important to DBES project quality:

- Building Contractors
- End-users
- Management of DBES
- Other Stakeholders

3.1.3 Project scope

Inadequate or poor scope definition may negatively collate to the project performance and result in failure of projects [12]. Moreover, it may lead to poor design basis and adversely affect projects [13]. Project scope related factors identified in this study are:

- Sufficient project information provided by DBES
- Completeness and clarity of the scope definition provided by DBES
- Pre-project planning effort on government projects by DBES
- Stability (less changes in) of design brief or project requirement during the design phase
- Management of changes in design brief at the design phase

3.2 Cognizant measures

There are certain aspects that the design process should consider in order to achieve a positive outcome at a later stage. For example, value management is important “when developing the design from the project brief at the conceptual design stage, as well as in limiting any variations to those that are absolutely essential [2]. Another important aspect that the design process of construction projects should take into account is the project life
An empirical study of factors influencing project quality in the design phase
Dodoo, van Waveren, Chan

cycle view so that projects are designed with constructability, maintainability, supportability and the control of total cost in mind. The aspects that the design process of construction projects should take into considerations are identified in this study as (denoted as cognizant measures):

- Cognizance of time management
- Cognizance of value management
- Cognizance of project life cycle view
- Cognizance of requirements and expectations of the end users
- Cognizance of communication management

4. RESEARCH METHODOLOGY

4.1 Data collection

This research applies a quantitative research methodology and uses a questionnaire survey to collect data. A total of 115 questionnaires were distributed to participants (units of analysis) involved in DBES building projects. The survey sample comprised of 5 Project Managers, 15 Architects, 15 Quantity Surveyors, 15 Civil Engineers, 15 Structural Engineers, 15 Mechanical Engineers, 15 Electrical Engineers, 10 client ministries/departments and 10 PPADB (Public Procurement and Asset Disposal Board) registered contractors in grade E. A total of 70 completed questionnaires were returned resulting in a response rate of 61%. From the returned questionnaires, it was found that respondents who have more than 5 years of experience constitute to 70% of the population and all of the respondents have participated in DBES building projects; this indicates the validity of the sample and data collected.

4.2 Questionnaire design and statistical methods

There are three parts in the questionnaire design:

i. Attributes of the respondents: years of experience, professions or trainings, sectors of employment.
ii. Factors during design phase that influence building project quality.
iii. Aspects that design phase takes cognizance of (cognizant measures).

Respondents were asked to use a Likert scale of 5 to rate each factor from 1 (very low) to 5 (very high). These ratings were used to evaluate participants involved in DBES projects against project success factors in order to identify key focus areas.

In the first part of the data analysis, the SPSS statistical software was used to determine the means and standard deviation of the attributes of the respondents, the factors and the measures in the design phase. Moreover, a reliability test was done on the factors to determine how well the items measured a single, unidimensional latent construct; in this study, how well the factors could be represented as a category. The second part is a multivariate analysis which explored the interrelationships between the factors and the cognizant measures in the design phase.

5. ANALYSIS OF DATA AND DISCUSSION

5.1 Descriptive statistics: attributes of respondents

The frequency counts of the respondents’ attributes are illustrated in Table 1.
The years of experience of the project participants are spread more or less equally across 1-20 years. If one takes the cut-off point of 10 years, there are more or less two groups of participants with regards to their years of experience: 58.6% of participants with less and equal to 10 years of experience and 41.4% of more than 10 years. This suggests that the DBES are not dominated by participants who are inexperienced. In the profession or training attribute, those professions which were constructional related (i.e. architectural & quantity surveyor services, civil & structural engineering) constituted to 48.6% while the other professions were non-constructional (i.e. electrical and mechanical engineering) and constituted to 51.4%. These two groups may be associated with participants in the previous two departments mentioned in the introduction section (i.e. DABS & DEMS). This result suggests that there is a balanced mixture of participants with building construction related background and those with engineering (electrical & mechanical) background. Participants working on DBES building projects who are employed in the public sector and the private sector have similar group sizes. This indicates that the DBES projects are neither public oriented nor private oriented. Therefore the data gathered in this study is not biased.

5.2 Descriptive analysis: Factors in design phase

As mentioned in the research methodology section, questionnaire survey was used and the respondents were asked to rate the factors in the questionnaire to explore how DBES in behave with regards to these factors in the design phase. In Table 2, descriptive statistics of all the factors are presented.

From Table 2 one can notice that none of the factors has a mean score above 3 (which is the mid-point on a scale of 5). This observation implies that respondents working on DBES building projects do not perform above or equal to satisfactory level under these factors assessed.
An empirical study of factors influencing project quality in the design phase  
Dodoo, van Waveren, Chan

There is internal consistency (or reliability) in the factors measured under each category; i.e. all the Cronbach’s α’s are greater than 0.6, which indicates reliable scales. In other words, the average scores of all the means of factors under the corresponding category can represent the score of that particular category. It was found that the average score for the category Adherence to Project Management Principles (mean=2.7082) was the lowest amongst all three categories. Moreover, within this category, DBES’s commitment to continuous quality improvement, and use of lessons learned to improve the quality of design had the two lowest means (below 2.6) and visibility of the total life cycle cost during design phase (the third lowest) amongst the seven factors. This finding implies that DBES’s poor project quality may be the result of participants involving in DBES building projects not satisfactorily adhere to project management principles, especially with regards to the three particular factors mentioned above, during its project design phase.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category A: Adherence to project management principles in project design phase</strong>&lt;br&gt;α = 0.655; mean = 2.7082; S.D. = 0.85833</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1: Selection process of the design team</td>
<td>2.83</td>
<td>1.129</td>
</tr>
<tr>
<td>A2: Design brief from clients</td>
<td>2.83</td>
<td>1.204</td>
</tr>
<tr>
<td>A3: DBES’s commitment to continuous quality improvement</td>
<td>2.57</td>
<td>1.149</td>
</tr>
<tr>
<td>A4: DBES leadership commitment to adherence of procedures and specifications at the design phase</td>
<td>2.80</td>
<td>0.957</td>
</tr>
<tr>
<td>A5: Comprehensiveness of drawings and specifications</td>
<td>2.73</td>
<td>0.977</td>
</tr>
<tr>
<td>A6: Visibility of the total life cycle cost during the design phase</td>
<td>2.63</td>
<td>1.066</td>
</tr>
<tr>
<td>A7: Use of lessons learned to improve the quality of designs</td>
<td>2.57</td>
<td>1.124</td>
</tr>
<tr>
<td><strong>Category B: Involvements of role players in project design phase</strong>&lt;br&gt;α = 0.898; mean = 2.7357; S.D. = 0.75767</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1: Involvement of building contractors</td>
<td>2.63</td>
<td>1.132</td>
</tr>
<tr>
<td>B2: Involvement of the end-users</td>
<td>2.90</td>
<td>1.276</td>
</tr>
<tr>
<td>B3: Involvement of the management of DBES</td>
<td>2.80</td>
<td>1.001</td>
</tr>
<tr>
<td>B4: Involvement of other stakeholders</td>
<td>2.61</td>
<td>0.906</td>
</tr>
<tr>
<td><strong>Category C: Project scope</strong>&lt;br&gt;α = 0.647; mean = 2.7600; S.D. = 0.65106</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1: Clarity and completeness of the project information provided by DBES</td>
<td>2.84</td>
<td>0.973</td>
</tr>
<tr>
<td>C2: Completeness and clarity of the scope definition provided by DBES</td>
<td>2.81</td>
<td>0.952</td>
</tr>
<tr>
<td>C3: Pre-project planning effort on government projects by DBES</td>
<td>2.64</td>
<td>1.077</td>
</tr>
<tr>
<td>C4: Stability (less changes in) of design brief or project requirement during the design phase</td>
<td>2.70</td>
<td>0.998</td>
</tr>
<tr>
<td>C5: Process of management of changes in design brief</td>
<td>2.80</td>
<td>1.016</td>
</tr>
</tbody>
</table>

There are internal consistency (or reliability) in the factors measured under each category; i.e. all the Cronbach’s α’s are greater than 0.6, which indicates reliable scales. In other words, the average scores of all the means of factors under the corresponding category can represent the score of that particular category. It was found that the average score for the category Adherence to Project Management Principles (mean=2.7082) was the lowest amongst all three categories. Moreover, within this category, DBES’s commitment to continuous quality improvement, and use of lessons learned to improve the quality of design had the two lowest means (below 2.6) and visibility of the total life cycle cost during design phase (the third lowest) amongst the seven factors. This finding implies that DBES’s poor project quality may be the result of participants involving in DBES building projects not satisfactorily adhere to project management principles, especially with regards to the three particular factors mentioned above, during its project design phase.
5.3 Descriptive analysis: Cognizant measures in design phase

These variables were measured using a 4-points Likert scale. All the means are above 3. This indicates that the project participants take all five aspects (cognizant measures) into considerations during the project design phase at more than a satisfactory level (above the mid-point of 2.5). During the design phase, project participants take the most considerations on *end users’ requirements and expectation* (with the mean of 3.71) and the least on *project life cycle view* (with the mean of 3.19).

5.4 Multivariate analysis: factors and measures

The purpose of this analysis to to explore the relationships between the factors and the cognizant aspects. These relationships will indicate how the project participants’ performances (with respect to a specific success factor) and their cognizance of a specific cognizant measure relate to each other. The framework of this analysis is shown in Fig 1 and the statistical results are shown in Table 4.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Means</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1: Cognizance of time management</td>
<td>3.33</td>
<td>0.675</td>
</tr>
<tr>
<td>M2: Cognizance of value management</td>
<td>3.2</td>
<td>0.844</td>
</tr>
<tr>
<td>M3: Cognizance of project life cycle view</td>
<td>3.19</td>
<td>0.873</td>
</tr>
<tr>
<td>M4: Cognizance of end users’ requirements and expectation</td>
<td>3.71</td>
<td>0.486</td>
</tr>
<tr>
<td>M5: Cognizance of communication management</td>
<td>3.49</td>
<td>0.583</td>
</tr>
</tbody>
</table>

Table 3 - Means and standard deviations of cognizant measures

Fig 1 - Research framework for multivariate analysis
An empirical study of factors influencing project quality in the design phase  
Dodoo, van Waveren, Chan

Involvement of building contractors (B1) has positive and significant (at significant level of \( p<0.01 \)) impact on all cognizant measures, except the cognizance of end users’ requirements and expectations. This means that the more the involvement of building contractors, the more the project participants consider time management, value management, project life cycle view and communication management.

Process of management of changes in design brief (C5), has negative impact (significant at \( p<0.05 \)) on cognizance of time management and positive impact on cognizance of value management. This result suggests that the better the project participants handle the process in managing the changes in design brief, they considered time management less important.

There are five factors that have slightly significant impacts (\( p<0.1 \)). Adhere to design brief from clients (A2) has negative impact and the use of lessons learned to improve the quality of designs (A7) has positive impact on cognizance of time management. Process of management of changes in design brief (C5) has positive impact on cognizance of time management. Clarity and completeness of the project information provided by DBES (C1) has negative impact with cognizance of project life cycle view. Pre-project planning effort on government by DBES (C3) has positive impact on cognizance of end users’ requirements and expectations.

5.5 Multivariate analysis: Categories and measures

The purpose of this analysis is to reduce the list of these factors into the three main categories and explore how these categories relate to the cognizant measures. With the purpose to identify which category has the most impact on a specific cognizance measure, the three categories are entered into the regression model one at a time as shown in three steps:

<table>
<thead>
<tr>
<th>Variables</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.964***</td>
<td>1.343**</td>
<td>2.016***</td>
<td>3.365***</td>
<td>2.684***</td>
</tr>
<tr>
<td>A1</td>
<td>-0.052</td>
<td>-0.267</td>
<td>-0.174</td>
<td>-0.043</td>
<td>-0.066</td>
</tr>
<tr>
<td>A2</td>
<td>-0.339*</td>
<td>0.154</td>
<td>0.046</td>
<td>0.233</td>
<td>0.163</td>
</tr>
<tr>
<td>A3</td>
<td>-0.066</td>
<td>0.096</td>
<td>-0.119</td>
<td>-0.061</td>
<td>0.325</td>
</tr>
<tr>
<td>A4</td>
<td>0.291</td>
<td>-0.200</td>
<td>0.305</td>
<td>0.357</td>
<td>-0.153</td>
</tr>
<tr>
<td>A5</td>
<td>-0.013</td>
<td>0.062</td>
<td>0.078</td>
<td>-0.295</td>
<td>0.017</td>
</tr>
<tr>
<td>A6</td>
<td>-0.143</td>
<td>0.102</td>
<td>0.119</td>
<td>-0.102</td>
<td>0.090</td>
</tr>
<tr>
<td>A7</td>
<td>0.368*</td>
<td>0.302</td>
<td>0.293</td>
<td>-0.072</td>
<td>-0.186</td>
</tr>
<tr>
<td>B1</td>
<td>0.392***</td>
<td>0.478**</td>
<td>0.454**</td>
<td>0.056</td>
<td>0.345***</td>
</tr>
<tr>
<td>B2</td>
<td>-0.244</td>
<td>-0.104</td>
<td>-0.176</td>
<td>0.033</td>
<td>-0.016</td>
</tr>
<tr>
<td>B3</td>
<td>0.127</td>
<td>-0.078</td>
<td>-0.041</td>
<td>0.057</td>
<td>0.046</td>
</tr>
<tr>
<td>B4</td>
<td>0.205</td>
<td>0.073</td>
<td>0.089</td>
<td>-0.255</td>
<td>-0.243</td>
</tr>
<tr>
<td>C1</td>
<td>-0.019</td>
<td>-0.003</td>
<td>-0.329*</td>
<td>0.192</td>
<td>0.072</td>
</tr>
<tr>
<td>C2</td>
<td>-0.046</td>
<td>-0.129</td>
<td>0.124</td>
<td>-0.123</td>
<td>0.224</td>
</tr>
<tr>
<td>C3</td>
<td>0.204</td>
<td>0.012</td>
<td>0.014</td>
<td>0.394</td>
<td>-0.087</td>
</tr>
<tr>
<td>C4</td>
<td>-0.112</td>
<td>0.169</td>
<td>-0.046</td>
<td>-0.091</td>
<td>-0.104</td>
</tr>
<tr>
<td>C5</td>
<td>-0.364*</td>
<td>0.290*</td>
<td>-0.090</td>
<td>0.024</td>
<td>0.168</td>
</tr>
</tbody>
</table>

\* \( p < 0.10 \); \* \( p < 0.05 \); *** \( p < 0.01 \)

Table 4 - Results from regression analysis
An empirical study of factors influencing project quality in the design phase  
Dodoo, van Waveren, Chan

Model 1: Model with only Category A (Adherence to project management principles)  
Model 2: Model 1 + Category B (Involvement of role players in project design phase)  
Model 3: Model 2 + Category C (Project scope)

Each time when a variable is added into the model, its associated $\Delta R^2$ will indicate if the added variable is significantly accountable to the variance in the model. The statistical results are shown in Table 5.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Dependent variable: cognizant measures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M1</td>
</tr>
<tr>
<td>Model 1</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>Category A</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R^2$</td>
</tr>
<tr>
<td></td>
<td>F-value</td>
</tr>
<tr>
<td></td>
<td>$\Delta F$-value</td>
</tr>
<tr>
<td>Model 2</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>Category A</td>
</tr>
<tr>
<td></td>
<td>Category B</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R^2$</td>
</tr>
<tr>
<td></td>
<td>F-value</td>
</tr>
<tr>
<td></td>
<td>$\Delta F$-value</td>
</tr>
<tr>
<td>Model 3</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>Category A</td>
</tr>
<tr>
<td></td>
<td>Category B</td>
</tr>
<tr>
<td></td>
<td>Category C</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
</tr>
<tr>
<td></td>
<td>$\Delta R^2$</td>
</tr>
<tr>
<td></td>
<td>F-value</td>
</tr>
<tr>
<td></td>
<td>$\Delta F$-value</td>
</tr>
</tbody>
</table>

* p < 0.10;  ** p < 0.05;  *** p < 0.01

Table 5 - Results from multivariate regression analysis

The Variable Inflation Factor (VIF) values associated with variables in the regression models were less than 10, indicating that serious multicollinearity problems do not exist in these models.

Observing the three models with the dependent variable as cognizance of time management (M1), model 2 has the best fit with the addition of Category B (regression model as a whole has $\Delta F$-value of 3.153 with p<0.05). This result suggests that as compared to the other two categories, Category B has the positive and significant (at p<0.05) impact on the cognizance of time management.
An empirical study of factors influencing project quality in the design phase
Dodoo, van Waveren, Chan

Amongst all models with dependent variable M2 (cognizance of value management), model 3 has the best fit (ΔF-value of 3.670 with p<0.1). The effect of Category C (project scope) in model 3 accounted for approximately 4.9% of the variance in M2. Moreover, Category C has a positive impact (significant at p<0.1) on M2.

Model 1 has the best fit with dependent variable M3 (cognizance of project life cycle view). Addition of the other two categories did not result in any significant ΔF-value. Category A (adherence to project management principles) has positive impact on M3 with significance at p<0.05. Moreover in the analysis with last two dependent variables (M4 & M5), Category A has positive impact on M4 (cognizance of end users’ requirements and expectations) and M5 (cognizance of communication management).

6. CONCLUSION AND RECOMMENDATIONS

Due to numerous complaints about DBES building projects, the purpose of this empirical paper is to assess participants in DBES building projects under 16 factors during the project design phase which influences project quality. Moreover, this paper also explores the interrelationships between these factors and the cognizant measures identified. In this section the most important findings are summarized and discussed. Further recommendations are suggested.

This paper has shown statistical evidence of less than satisfactory performance of project participants during DBES building projects design phase. That is, all factors have mean values less than 3 (mid-point on a scale of 5). The key focus area that needs the most attention is Adherence to Project Principles during design phase (which has the lowest score out of the three categories). DBES may improve upon three factors in order to address this problem, namely: commitment to continuous quality improvement, use lessons learned to improve quality of designs and visibility of the total life cycle cost during design phase.

Moreover, this paper has revealed certain relationships between the success factors and the cognizant measures. The most important factor is the involvement of contractors which has a positive influence on how well the participants take cognizance of time management, value management and communication management. Perhaps the contractors have provided the project participants with the relevant knowledge. Involvement of contractors has no effect on the cognizance of end users’ requirements and expectations. This may be attributed to the fact that contractors and end users do not communicate directly. One may assume that the involvement of end users (B2) should enhance the cognizance of end user’s requirements and expectations (M4). However, it is not the case in DBES projects. Participants of DBES projects from the public sector have direct communication with the end-users and pass on the information regarding design requirements to the participants in the private sector (e.g. design consultants). Information may be inaccurately interpreted and may be lost when transferring it from the public sector to the private sector. Inaccurate information may contribute to the reasons for poor design quality (which does not meet specifications). Thus for all the project participants to take cognizance of end users, this study recommends that all three parties (i.e. end users as well as project participants from both the public and the private sectors) should be involved during the project design phase (see Fig 2).
An empirical study of factors influencing project quality in the design phase
Dodoo, van Waveren, Chan

Fig 2 - Direct communications amongst three parties to improve cognizance of end users’ requirements and expectations

There are three negative relationships identified in the first stage of the multivariate analysis. Although these results are slightly significant, one can still identify the problem areas of DBES projects.

- The more project participants adhere to the design brief from clients, the less they take cognizance of time management: This finding suggests that time is subordinated to project quality. In other words, project quality is of prime importance in the case of DBES building projects; and the less consideration of time may be attributed to the time delays of the DBES projects.

- When project participants perform better in the process of management of changes in the design brief, they tend to take less cognizance of time management: To effectively manage changes in the project design, a configuration management (CM) approach can be used where controlling the changes to the item and its documentation is part of such approach. When clients request changes in the design brief, one has to consider how this change may affect the time and cost of the project. Project participants know that CM is time consuming and time delays are inevitable. Therefore they may not put time as prime priority and consider project quality and cost to be more important.

- The more clearer and complete the project information is, as provided by DBES, the less the project participants take cognizance of project life cycle view: The clearer and more complete information provided by DBES may include the necessary tasks allocated to participants in more detail (e.g. participant A has to perform task A at Phase A of the project life cycle). Thus participants at each phase feel that it’s not necessary for them to consider or take care of other phases in the life cycle because they know in advance what their tasks are and when to perform these tasks. All they need to do is ‘follow the plan’ in their allocated phase. On the other hand, if the information provided by DBES is vague or incomplete, the project participants may be more uncertain about their tasks and feel that they need to see the ‘whole picture’ (i.e. project life cycle view) in order to figure out where their tasks fall within the life cycle.

When comparing the three categories, Category A (adherence to project management principles) has more influence on how project participants take cognizance of project life cycle view (M3), end users’ requirements and expectations (M4) and communication management (M5). Since project management principles are organization specific (i.e. they differ from one organization to another) and in DBES it seems these principles take the above mentioned aspects into account. It is recommended that these principles also take into account the other two aspects (communication management & time management) to
An empirical study of factors influencing project quality in the design phase

Dodoo, van Waveren, Chan

further improve DBES projects. For examples, time management enhances projects to complete within time schedule and communication management allow project participants ‘to understand profoundly the status and level of progress in the project’ [14].

Several future studies are suggested. Firstly, DBES may also be evaluated under factors during other project phases, e.g. execution phase. This may further point out other key focus areas to ensure better project quality. Secondly, this study only considers three key focus areas (i.e. the three categories) of the design phase; the framework may be extended further by exploring other focus areas as well. For example, environmental impact factors can also be taken into consideration during the project design phase. Thirdly, similar studies may be performed in other African countries for benchmarking purpose. There may be a common trend with the problem areas in the African environment, which the investors should take note of.

REFERENCES

DEFINITION AND ROLE OF AN INNOVATION STRATEGY

B.R. Katz1*, N.D. du Preez2, C.S.L. Schutte3
1Department of Industrial Engineering
Stellenbosch University, RSA
bernard@indutech.co.za

2 Department of Industrial Engineering
Stellenbosch University, RSA
niek@indutech.co.za

3 Department of Industrial Engineering
Stellenbosch University, RSA
corne@sun.ac.za

ABSTRACT

The concept of an innovation process to formalize Innovation Management in a company has been suggested [1]. As part of the formalization of innovation as a core business process the role and importance of an innovation strategy has come to the fore.

The questions are; what is the definition of an innovation strategy? What is the role of an innovations strategy? What are the components of an innovation strategy?

This paper presents a definition for an innovation strategy, the components of an innovation strategy and describes how the innovation strategy is positioned in relation to traditional business strategies.

OPSOMMING

Die konsep van ‘n innovasieproses om Innovasiebestuur in ‘n maatskappy te formaliseer word voorgestel, en as deel van innovasie se formaliseringsproses word die rol en belangrikheid van ‘n innovasiestrategie benadruk.

Die volgende vrae word gestel: Wat is die definisie van ‘n innovasiestrategie, wat is die rol van ‘n innovasiestrategie en wat is die komponente van die innovasiestrategie?

Hierdie artikel stel ‘n definisie voor vir ‘n innovasiestrategie en die komponente van so ‘n strategie, en beskryf hoe die innovasiestrategie geposisioneer moet word ten opsigte van tradisionele besigheidstrategieë.

1 The author is enrolled for a PhD (Industrial Engineering) degree in the Department of Industrial Engineering, Stellenbosch University
*Corresponding author
INTRODUCTION

In recent times the role and importance of innovation in companies has grown significantly [2]. The amount of time, money and effort focused on innovation and innovation management in both academic and commercial environments has too increased significantly. The concept of an innovation process to formalize innovation management in a company has been suggested [1]. Innovation processes are receiving more attention in companies and several attempts to describe these processes have been published.

As part of the formalization of innovation as a core business process the role and importance of an innovation strategy has come to the fore. As with all business processes the decisions made and the way in which the process is implemented are all governed by a strategy. The strategy governing innovation in a company is termed an innovation strategy.

This paper aims to define what an innovation strategy is and to describe the role of an innovation strategy in a company and how an innovation strategy can be positioned in context with other business strategies. This paper concludes by identifying the different components required in a comprehensive innovation strategy.

The questions addressed in this paper are:

1. What is the full scope of a definition for an innovation strategy?
2. What is the role of an innovation strategy in a company?
3. What are the main components of an innovation strategy?

In order to address the first question and develop a comprehensive definition of an innovation strategy, the various definitions of innovation and of strategy as separate domains need to be understood. The paper therefore starts with an investigation into the definition of both innovation and strategy with each of these sections concluding with a comprehensive definition.

Before the first research question can be finalized and a comprehensive definition for an innovation strategy presented, the second research question has to be answered. The role of an innovation strategy is described based on the role of innovation and the structure and role of general business strategies in a company.

Once the definition of an innovation strategy is developed and the role of an innovation strategy is described the final research question is answered by suggesting the components required in a comprehensive innovation strategy in order to fulfill the previously defined definition and role. The process followed is illustrated in Figure 1.
2 DEFINITIONS FOR INNOVATION AND STRATEGY

The difficulty in defining what exactly is an innovation strategy is due to the many different definitions within the two separate domains of innovation and strategy. Therefore, in order to develop an appropriate definition for innovation strategy it is important to first understand the range of definitions for both innovation and strategy as separate domains.

2.1 Defining Innovation

It is important to clarify a complete definition for innovation, in order to develop a complete definition for an innovation strategy.

The definition of innovation is highly varied, possibly due to the fact that many different disciplines have focused on innovation from their specific perspective. Also, the concept of innovation has become more complex over time [3].

An early definition of innovation, from an economics perspective, was presented by Schumpeter. An innovation - by definition -

“had a substantial economic impact. An innovation was something that changed the market place in a profound way. The innovating organization was, thus, likely to become the new market leader and to gain an immense advantage over its competitors” [4].

The 21st Century Working Group defines innovation as follows:

“Innovation transforms insight and technology into novel products, processes and services that create new value for stakeholders drive economic growth and improve standards of living” [5].

Several other definitions also include the concepts of new and novel [6].

2.2 Categorizing Innovation

Several different ways of categorizing innovation have been suggested [2]. These include:

- Types of innovation,
- Newness of the innovation,
- Impact of the innovation.

2.2.1 Types of Innovation

Types of innovation can also be described as the end result of the innovation. This can be clearly seen from the following six different innovation types [6]:

- New products
- New services
- New methods of production
- Opening new markets
- New sources of supply new ways of organizing

As reported by Katz [7]:

---

*Definition and Role of an Innovation Strategy*

*Katz, du Preez, Schutte*

*2 DEFINITIONS FOR INNOVATION AND STRATEGY*

The difficulty in defining what exactly is an innovation strategy is due to the many different definitions within the two separate domains of innovation and strategy. Therefore, in order to develop an appropriate definition for innovation strategy it is important to first understand the range of definitions for both innovation and strategy as separate domains.

2.1 Defining Innovation

It is important to clarify a complete definition for innovation, in order to develop a complete definition for an innovation strategy.

The definition of innovation is highly varied, possibly due to the fact that many different disciplines have focused on innovation from their specific perspective. Also, the concept of innovation has become more complex over time [3].

An early definition of innovation, from an economics perspective, was presented by Schumpeter. An innovation - by definition -

“had a substantial economic impact. An innovation was something that changed the market place in a profound way. The innovating organization was, thus, likely to become the new market leader and to gain an immense advantage over its competitors” [4].

The 21st Century Working Group defines innovation as follows:

“Innovation transforms insight and technology into novel products, processes and services that create new value for stakeholders drive economic growth and improve standards of living” [5].

Several other definitions also include the concepts of new and novel [6].

2.2 Categorizing Innovation

Several different ways of categorizing innovation have been suggested [2]. These include:

- Types of innovation,
- Newness of the innovation,
- Impact of the innovation.

2.2.1 Types of Innovation

Types of innovation can also be described as the end result of the innovation. This can be clearly seen from the following six different innovation types [6]:

- New products
- New services
- New methods of production
- Opening new markets
- New sources of supply new ways of organizing

As reported by Katz [7]:
“When looking at types of innovation, Baker [2] states that a company’s ability to support product and process innovation is no longer adequate and that a third type of innovation, strategy innovation needs to be introduced in order to provide further support. This type of innovation specifically emphasizes the importance of a longer term view of the contribution of innovation towards competitiveness and success as a company.”

Hence three main innovation types can be defined. These are product innovation, process innovation and strategic or business model innovation.

2.2.2 Newness of Innovation

The newness of an innovation can be described on a continuum from radical innovation to incremental innovation.

“Radical innovations are those that produce fundamental changes in the activities of an organization and large departures from existing practices, and incremental innovations are those that result in a lesser degree of departure from existing practices” [8].

2.2.3 Impact of Innovation

The impact of innovation can be described on a continuum from sustaining to disruptive.

“Sustaining innovations improve the performance of established products or services. Discontinuous innovations bring to market very different products or services that typically undermine established products and services in the particular market sector” [2].

2.3 Comprehensive Definition of Innovation

Taking into consideration the key concepts of innovation from the literature as well as the different ways of categorizing innovation a comprehensive definition of innovation can be developed.

“The successful generation, development and implementation of new and novel ideas, which, introduce new products, processes and/or strategies to a company or enhance current products, processes and/or strategies leading to commercial success and possible market leadership and create value for stakeholders, driving economic growth and improving standards of living” [7].

This definition is used to develop a comprehensive definition for an innovation strategy.

2.4 Defining Strategy

In order to develop a comprehensive definition for an innovation strategy an understanding of the types and definition of strategy is required. In this section an explanation of the evolution of strategy is provided along with several definitions for strategy.

There are almost as many different definitions for strategy as there are writers on the topic [9]. As cited by Heath [9], Donald C. Hambrick suggested two main reasons for the lack of consensus: first, strategy is a multi-disciplinary concept and second, strategy is situational and will consequently tend to vary by industry.
2.4.1 A Brief History of Strategy in a Business Context

The origins of strategy can be traced back through history, initially as a military concept. General Ulysses Grant in the 1860’s viewed strategy as:

“the deployment of one’s resources in a manner which is most likely to defeat the enemy” [10].

The application of a structured strategy in a business context can be traced back to the early 1970’s [11].

The history of strategy in the context of business started with the framework of Richard D. Irwin in 1971. This framework defined strategy as a match between what a company is capable of doing within the given environmental circumstances. The company’s capabilities are considered its strengths and weaknesses, while the environmental circumstances are considered as external opportunities or threats. At the time the required techniques to analyze the internal and external environments were not available, which made the practical use of this framework challenging [12].

In 1980, Michael E. Porter’s book, Competitive Strategies: Techniques for Analyzing Industries and Competitors, introduced a range of breakthrough approaches for developing company strategies based on the forces which exist within a given industry [12]. These forces are commonly known as the five competitive forces that shape strategy [13]. Porter’s approach does take the internal capability of the company into account; however it has a strong focus on the competitive positions at industry level.

At a later stage the focus of strategy shifted to a more internal perspective. Emphasis was placed on the internal skills and capability of the company. The strategic quest for a learning organization was deemed essential to achieve a competitive strategic position. The analysis of the industry and external environment received little attention [12].

In the mid-nineties a resources based view of strategy development was suggested. This approach includes both the external analysis and internal perspective promised by Irwin’s original framework [12].

2.4.2 Defining Strategy

As with the definition of innovation there are several different concepts that are suggested in the many different documented definitions of strategy.

There are definitions which consider strategy to be management’s action plan for running the business [14]. Other definitions include the concept of competition, where the aim of a strategy is to gain a competitive advantage in the market. In several definitions the concept of determining long-term goals and objectives is included. The allocation of resources at a high-level is another concept used in the definition of strategy. The definition by A.D. Chandler incorporates the majority of these concepts.

“Strategy can be defined as the determination of the basic long-term goals and objectives of an enterprise, and the adoption of courses of action and the allocation of resources necessary for the carrying out of these goals” [15].

64
The problem with a single definition for strategy is that it limits the use of strategy and therefore limits the potential benefits companies may gain from a wider application of Strategic Management. Mintzberg addresses this issue by providing five definitions for strategy and discusses the interrelationships between these definitions [16]. Table 1 summarizes these five definitions.

<table>
<thead>
<tr>
<th>Strategy as a ...</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan</td>
<td>Consciously intended course of action, a guideline to deal with a situation.</td>
</tr>
<tr>
<td>Ploy</td>
<td>A specific “maneuver” intended to outwit an opponent or competitor.</td>
</tr>
<tr>
<td>Pattern</td>
<td>Stream of actions revealing a consistency in behavior.</td>
</tr>
<tr>
<td>Position</td>
<td>Locating a company in the greater business environment. A niche.</td>
</tr>
<tr>
<td>Perspective</td>
<td>Ingrained way of perceiving and interacting with the world. A company’s personality.</td>
</tr>
</tbody>
</table>

Table 1 - Five definitions for strategy [16]

2.4.3 Strategy Hierarchy

An important part of understanding the role, structure and importance of an innovation strategy is to understand other business strategies. More specifically, the way in which traditional business strategies create various layers of a company’s strategy.

In the same way that a company’s organizational structure is seldom completely flat, a company’s strategy also requires hierarchical levels. These levels start with the overall corporate strategy and then filter down into more detailed functional strategies. Three generic hierarchical strategy levels can be defined. These levels are presented in Table 2 [17].

<table>
<thead>
<tr>
<th>Strategic Level</th>
<th>Primary Focus</th>
<th>Questions Answered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate</td>
<td>Scope of business.</td>
<td>What set of businesses should we be in?</td>
</tr>
<tr>
<td>Business</td>
<td>Distinctive competencies and competitive advantages.</td>
<td>How do we compete in a particular business?</td>
</tr>
<tr>
<td>Functional area</td>
<td>Coordination and integration of activities within a single function.</td>
<td>How do we maximize resource productivity?</td>
</tr>
</tbody>
</table>

Table 2 - Hierarchy of strategies [17]

Traditionally the following functional strategies are common in companies:

- Finance and Accounting
- Human Resources
- Information Systems
- Marketing
- Production/Operations
- Research & Development
2.4.4 Rational and Incremental Approaches to Strategy

There has been a well documented debate among business strategists about the pros and cons of rationalist and incrementalist approaches to strategy.

“Rationalist strategy has been heavily influenced by military experience, where strategy consists of the following steps: (1) describe, understand and analyze the environment; (2) determine the course of action in the light of the analysis; (3) carry out the decided course of action” [18].

Tidd [18] goes on to explain that,

“incrementalists argue that the complete understanding of complexity and change is impossible: Our ability both to comprehend the present and to predict the future is therefore inevitably limited.”

Therefore incrementalist strategies evolve over time and are adjusted far more frequently than rationalist strategies. These adjustments are made as uncertainty is reduced through a better understanding of the complexities involved.

Due to the high-levels of uncertainty and complexity associated with innovation, an incrementalist approach would best fit the management of an innovation strategy. Despite our best efforts companies are not able to predict the future with any amount of certainty. This means that even an explicit and documented innovation strategy should be adjusted on an ongoing basis in incremental steps as new certainties reveal themselves.

2.5 Comprehensive Definition of Strategy

A single comprehensive definition for strategy is not possible due to the different roles strategy plays, the different hierarchical strategic levels and even the different strategy development processes.

The best one can do is to highlight concepts, which are prominent in the strategy literature and ensure these are represented in the definition of an innovation strategy. Therefore a strategy is a:

- Determination of long-term goals.
- Conscious and pre-determined action plan to achieve goals.
- Allocation of resources required by an action plan.
- External analysis and an internal perspective.
- Ploy, pattern, position or perspective.
- Corporate, business or functional level strategy
- Rationalist or incrementalist strategy

These concepts are used, along with the comprehensive definition of innovation to develop a definition for an innovation strategy.

3 DUAL ROLES OF AN INNOVATION STRATEGY

In order to develop a comprehensive definition for an innovation strategy the role or purpose of an innovation strategy needs to first be described. In this section the dual roles of an innovation strategy are described.
3.1 The Dual Focus of Innovation

The roles of an innovation strategy are closely linked to the role of innovation in a company. Innovation plays two major roles in the success of a company.

Firstly innovation can play a role in achieving a company’s current corporate objectives by enabling a company to launch innovative products, find innovative ways to enter new markets or improve internal efficiencies. However, achieving current corporate objectives is not the only role of innovation in a company.

Another important role of innovation in a company is to change the company direction, when required. Rather than innovation being used to achieve current corporate objectives, it is the mechanism for changing corporate direction and objectives.

A company which can successfully manage its current business optimally, using continuous improvement and innovation to do things better, while at the same time creating the business of the future, doing things differently, (Figure 2) is known as an ambidextrous company [19].

Based on these dual roles of innovation, two roles for an innovation strategy can be described. The first role is an improvement role. The innovation strategy playing this role can be called an improvement innovation strategy. The second role is a future business role. The innovation strategy playing this role can be called a future business innovation strategy.

3.2 Role of an Improvement Innovation Strategy (Doing Things Better)

In this role an improvement innovation strategy:

- Aligns innovation objectives with corporate objectives.
- Is a guideline on the type, level and impact of innovation required to achieve current corporate objectives.
- Allocates resources between current operations and innovation initiatives.
- Is a plan to best utilize resources for innovation.

The goal of an improvement innovation strategy is to ensure the optimal plan and resource allocation in order to achieve the company’s corporate objectives. Anthony [20] writes:
Definition and Role of an Innovation Strategy
Katz, du Preez, Schutte

“by allocating resources more efficiently and accelerating the highest potential innovations, companies can enjoy a winning streak of innovation successes that will throw competitors off balance.”

3.3 Role of a Future Business Innovation Strategy (Doing Things Differently)

In this case the role of an innovation strategy is to help decide when and how to selectively abandon the past in order to focus on the future business (Figure 2). It is about managing the transition between the company’s current S-Curve and its future S-Curve (Figure 3).

The S-Curve reveals how the growth of yesterday’s innovation, which over time has been improved and optimized, eventually starts to decline. Today’s innovation is the key to future growth. Deciding which S-Curve to follow next and when is the correct time for stepping from the old to the new is the role of a future business innovation strategy.

![Figure 3 - Transition between S-Curves](image)

Another way of understanding the role of a future business innovation strategy is to look at the different types of strategies presented in Table 1. A future business innovation strategy would be required to consciously change either the pattern, position or perspective strategies of a company. Changing a company’s historic pattern of actions or its position in the business environment or its inherent personality would require a significant departure from the current business and the creation of the future business.

Traditional corporate, business and functional area strategies are required to provide consistency and reduce uncertainty [21]. Therefore these strategies are a stabilizing force in a company allowing the company to concentrate on the detailed operations without having to keep worrying about the long-term direction. This means that once these strategies are set they become a barrier to change and therefore a barrier to innovation. The role of the future business innovation strategy is therefore to counteract this barrier and provide a framework in which these other strategies can be changed if and when required.
Thus, a future business innovation strategy is required to:

- Selectively abandon the past and start a new S-Curve.
- To change pattern, position and perspective strategies.
- Provide a framework to change traditional corporate, business and functional area strategies.

4 DEFINING AN INNOVATION STRATEGY

In both the International and European Journal of Innovation Management not a single result is found when searching for the term “innovation strategy”. This lack of literature about the definition, structure and use of an innovation strategy is reflected in the fact that very few companies have an explicit, documented innovation strategy. Companies, which do include innovation in their overall business strategy either, include it as a company value, a pillar of their overall strategy or as an important attribute of their business.

There are a few definitions of an innovation strategy in the literature, but each of these cover only a section of the overall role of an innovation strategy.

If, in its simplest form a company’s strategy is defined as a plan designed to achieve a particular long-term aim, then an innovation strategy can be defined as a plan, which will enable a company to achieve its long-term goals through the use of innovation.

“An innovation strategy helps firms decide in a, cumulative and sustainable manner, about the type of innovation that best match corporate objectives” [22].

If strategy is defined as a guide for the allocation of resources in order to achieve the company’s objectives then:

“An innovation strategy guides decisions on how resources are to be used to meet a firm’s objectives for innovation and thereby deliver value and build competitive advantage. [22]”

These definitions are in line with the improvement role of an innovation strategy. If the future business role of an innovation strategy is also taken into consideration, along with the identified concepts in the definitions of innovation and strategy then the following, comprehensive definition of an innovation strategy can be defined. An innovation strategy is:

- an incrementalist, functional, predetermined plan governing the allocation of resource to different types of innovations in order to achieve a company’s overall corporate strategic objectives and,
- a decision framework guiding a company about when and how it should selectively abandon the past and/or change its corporate strategy and objectives in order to focus on the business of the future.

In the following section these definitions as well as the previously defined roles of an innovation strategy are used to identify the components of an innovation strategy.
5 COMPONENTS OF AN INNOVATION STRATEGY

Based on the different roles and the comprehensive definition of an innovation strategy a set of components for an innovation strategy can be identified. These components provide the structure and content of an innovation strategy.

For the first part of the definition of an innovation strategy components are required to guide innovation selection and to allocate resources. For the second part of the definition components are required to guide the decision making process and to prepare the company for the possibility of significant change.

5.1 Components for Guiding Innovation Selection

These components are required to guide the selection of innovation initiatives. The type, level, impact and risk profile are four components which can be used to define the required portfolio of innovations and guide the selection of new innovation initiatives.

The innovation strategy describes the required blend of innovation types, levels, impacts and risks required to support the company’s strategic objectives. This then allows the gates and filters in the innovation process to be tuned to achieve this blend. Therefore, once the innovation strategy is aligned with the corporate strategic objectives the innovation process will ensure the correct innovation initiatives are selected in order to achieve those strategic objectives.

5.2 Components for Allocating Innovation Resources

These components are required to guide the allocation of resources for innovation. The distribution of resources for innovation is influenced by three main factors these are the ratio between innovation resources and operational resources, who will be involved in the innovation activities and where the innovation activities will take place. Based on these factors three innovation strategy components can be identified. These components are; collaboration, place and resources.

The resource component includes the split of resources between operating the current business, executing innovation initiatives and improving the company’s capability to innovate.

The place component helps balance the types of resources which are required. In-house innovation requires more internal personnel while outsourced innovation requires less internal personnel but often more financial resources. Greenfields innovation involves spinning out a new company or business unit. This often requires both high-levels of internal personnel and financial resources.

The collaboration component also has an influence on the level of human and financial resources. If the innovation is closed, more internal resources and capabilities are required. If however open innovation is applied less internal personnel are required but more financial resources are needed.
5.3 Components for Guiding the Decision Process and Preparing for Change

These components are required to guide the decision process regarding the future business of the company as well as to determine the required capability of the company to adapt to the future business.

The driver component describes the various external drivers, which at some point in the future will require the company to abandon the past and create its future business. The innovation strategy provides guidance on which drivers the company should focus.

The maturity component describes the current capability of the company to successfully innovate and the maturity improvement path the company should take in order to be prepared for its future business.

5.4 Nine Components of an Innovation Strategy and Options for Each

Therefore, nine components have been identified as crucial for an innovation strategy in order for the strategy to fulfill the roles described and align with the comprehensive definition defined. The components of an innovation strategy are presented in Figure 4.

![Figure 4 - Components of an innovation strategy](image)

Each component has several options, which either lie along a continuum or are discrete. These options need to be understood and selected by a company when developing an appropriate innovation strategy (Table 3).

For example, the four components for guiding innovation selection each have a blend of options. A specific innovation strategy could therefore require:

- Innovation type: 30% product, 20% process, 50% strategic
- Innovation level: 20% incremental, 50% semi-incremental, 30% radical
- Innovation impact: 10% disruptive, 90% sustaining
- Risk profile: 30% low risk, 40% medium risk, 30% high risk
Table 3 - Components and options of an innovation strategy

The combination of options for each strategy component influences the way in which a company develops its innovation process and capabilities.

The appropriate combination of options for each of the innovation strategy components depends on the industry in which the company operates, the current and medium-term economic conditions, the company’s other business strategies and the company’s internal capabilities and resources.

6 CONCLUSION

In order to develop a comprehensive definition for an innovation strategy the definitions of innovation and of strategy first had to be separately understood. The different roles of an innovation strategy also had to be described.

6.1 Definitions for Innovation and Strategy

Defining what is an innovation strategy is a difficult task due to the many and varied definitions for both innovation and strategy. In its simplest form the definition of innovation is a new or novel idea which adds value. There seems to be no simple definition for strategy. A strategy plays several significant roles in managing a company. These range from charting a direction, describing the way in which a company tries to gain a competitive advantage to allocating resources.

6.2 Different Roles of an Innovation Strategy

Two distinct roles for an innovation strategy are described. They are an improvement innovation strategy role and a future business innovation strategy role. The improvement innovation strategy plays a role in doing current things better in order to achieve current strategic objectives. The future business innovation strategy plays a role in selectively abandoning the past and selecting the business of the future.
6.3 Definition of an Innovation Strategy

It is clear that an innovation strategy needs to function across the different hierarchical strategy levels in a company. At one level an innovation strategy is a functional strategy used to plan the allocation of resources for innovation and to guide the selection of the type, level, impact and risk profile of innovation initiatives.

At another level it is a decision framework assisting corporate management to decide when to abandon their past and focus on the business of the future. At this level it no longer supports the traditional corporate strategy, but aims to change the very nature of the company.

6.4 Components of an Innovation Strategy

Based on the definition and the described roles of an innovation strategy nine components of an innovation strategy are identified.

The identified components of an innovation strategy play a role in supporting the different functions of an innovation strategy. The type, level, impact and risk profile components are mainly used to guide innovation selection. The collaboration, place and resource components influence the distribution of resources for innovation. The maturity and driver components prepare the company to innovate their future business.

The balance between the different options in each of the innovation strategy components has many influencing factors; including the industry, the economic conditions and the internal capabilities of the company. The development of an appropriate innovation strategy for a specific company, based on the identified components and options, is significantly influenced by these factors.

The significance of each factor on each of the innovation strategy components is yet to be determined. The next phase of this research into innovation strategy will concentrate on the various factors influencing the innovation strategy and will try developing an approach for a company to analyze these factors and then decide on an appropriate innovation strategy for its specific circumstances.

7 REFERENCES

Definition and Role of an Innovation Strategy
Katz, du Preez, Schutte


IMPROVING LOADING TIMES AND TRUCK ROLLOVER TENDENCY USING A PYRAMID STACKING METHOD

F. Kienhöfer¹, R. Crichton², J. Grobler³, and T. Dessein⁴

¹,⁴Department of Mechanical, Industrial and Aeronautical Engineering
University of the Witwatersrand, South Africa
frank.kienhofer@wits.ac.za, thomas.dessein@wits.ac.za

²,³Hall Longmore
Murray & Roberts, South Africa
robert.crichton@wits.ac.za, johannes.grobler@dhl.com

ABSTRACT
The new multi-product pipeline (NMPP) is a multibillion Rand project. The transport of the pipes to site presents an interesting multidisciplinary transport engineering case study. The loading configuration of the pipes was found to not only influence the loading and unloading times and load utilisation, but also the truck rollover tendency. Trucks transporting pipes in South Africa previously only used a rectangular stacking configuration. An alternative pyramid stacking configuration was successfully trialled. The loading and unloading times were reduced by 38% and 42% respectively. Vehicle loads were increased from 9 to 10 pipes even though the vehicle rollover safety was not compromised.

OPSOMMING
Die nuwe multi-produkte pyplyn (NMPP) is ’n multibiljoenrand projek. Die vervoer van die pype dien as ’n interessante multi-dissiplinêre vervoeringenieurswese gevallestudie. Die gewigsverhouding van die pype beïnvloed nie alleenlik die op en aflaai tye nie, maar ook die rolveiligheid van die voertuig. Voorheen was pype in Suid-Afrika slegs in ’n vierkantige konfigurasie vervoer. ’n Alternatiewe piramide-konfigurasie was suksesvol geëvalueer. Die op- en aflaai tye is met 38% en 42% onderskeidelik verminder. Met die nuwe konfigurasie kan 10 pype vervoer word (in vergelyking met 9 wanneer die ou konfigurasie gebruik word), sonder dat die rolveiligheid van die voertuig verander.

*Corresponding author
1. INTRODUCTION

1.1 Background

Achieving excellence and reducing costs in logistics systems are vital for competitiveness and profitability. In 2009, the total logistics cost relative to GDP (Gross Domestic Product) in South Africa was 14.7% [1] and has remained steady near this mark since the first State of Logistics survey was conducted for 2004 by the CSIR [2]. These figures compare unfavourably to the cost of logistics in the US. The US reduced the logistics cost relative to GDP from 16.2% in 1981 to 8.7% in 2002 [3]. In 2009, the total logistics cost relative to GDP in the US was 7.7% [4]: almost half the South African cost of logistics.

The above statistics not only indicate that South Africa needs to improve and find ways to reduce transport costs, but also highlight that the costs of logistics is significant in comparison to GDP. Global competitiveness is directly linked to the effectiveness of a nation’s logistics systems to transport goods [5]. In order to improve South Africa’s logistics systems and competitiveness, government needs to ensure appropriate funding of transport infrastructure and legislation [6]. However, on a microeconomic level, the opportunity exists for organisations and firms to improve their transport operations and logistics [7]: in so doing this will contribute to global competitiveness.

The new multi-product pipeline (NMPP) is a multibillion Rand project to transport refined petroleum products (petrol, diesel and jet fuel) from Durban to Gauteng. The project was initially valued at R11.5-billion [8] but has escalated by R2.75-billion to R15.42-billion [9]. A project of this size means that improvements in delivering steel pipes to site can save millions of Rands.

Hall Longmore (the pipe manufacturing company awarded the contract to supply the steel pipes for the NMPP) initiated an operations improvement plan to improve the transport configuration of the pipes, as well as the loading and unloading methods of the pipes. This forms an interesting multidisciplinary case study in that changing the transport configuration of the pipes not only affected the loading and unloading times, but also the rollover tendency of the fully loaded trucks. The success story is described using metrics developed in operations research, as well as vehicle dynamics.

1.2 Literature Review

In order to improve the logistics system, the logistics system must first be measured using the correct metric. Chow et al. [10] provide a summary of the research conducted on defining and measuring logistics performance from five leading logistics journals between 1982 and 1992. A comprehensive review of logistics metrics is given by Caplice and Sheffi [11] who further developed evaluation criteria to evaluate the appropriateness of the metrics. A more recent review of logistics metrics was conducted by Gunasekaran and Kobu [12]. The argument is made that traditional metrics which tend to be financial, focused on function and inward looking, have their place but more modern measures which are integrative and include intangibles like flexibility and innovation are required to shape strategy.

The transported load configuration on a truck not only affects logistics operational performance but also vehicle safety and, in particular, the rollover tendency of the fully loaded truck. The static rollover threshold (SRT) is the maximum steady state lateral acceleration that can be sustained by a vehicle in a constant radius high speed turn. The metric directly measures vehicle rollover stability. Using US accident statistics, Winkler [13] showed that a strong relationship exists between SRT and rollover accident risk. As roll stability decreases, so the probability of rollover in an accident increases rapidly until the
Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method
Kienhöfer, Crichton, Grobler, Dessein

vehicle becomes very likely to roll over in nearly any accident. A similar study by Mueller et al. [14] conducted in New Zealand showed an equally strong correlation between SRT and rollover accident risk. The study recommended a minimum SRT value of 0.35 g.

The SRT of a vehicle can be measured directly using a tilt table test and the procedure developed by the Society of Automotive Engineers [15]. The test method is, however, expensive and not many experimental facilities have a tilt table. Elischer and Prem [16] have developed an empirical prediction method for SRT which is accurate to ±7%.

No literature exists detailing the influence of loading configuration on logistics performance in conjunction with the rollover tendency of a fully loaded truck using SRT. This study emphasizes that the optimal truck loading method should consider both operations research and vehicle dynamics.

1.3 Aim of This Study

A case study is presented illustrating the benefits that Hall Longmore achieved in revising the loading and unloading method of pipes unto the required trucks used for transportation. The case study highlights the need for a multidisciplinary approach to analysing problems involving truck transport loading, as this embraces the fields of operations research and vehicle dynamics.

2. CASE STUDY DETAILS AND METHODOLOGY

Section 2.1 details the freight task i.e. the weights and dimensions of the pipes to be transported and the constraints imposed by vehicle legislation. Section 2.2 describes the pipe loading and unloading equipment as this influenced the pipe stacking configurations which are discussed in Section 2.3. The choice of logistics metrics to analyse the operational research performance of the revised pipe stacking method is presented and evaluated in section 2.4. Section 2.5 explains the method used to estimate the rollover tendency of the fully loaded truck.

2.1 The Freight Task and Constraints Imposed by Vehicle Legislation

The new multiproduct pipeline (NMPP) is a multibillion Rand project to pump fuel from Durban to Gauteng. The main trunk line from Durban to Jameson Park will comprise 544 km of 24-inch diameter pipes with a wall thickness of 11.6 mm and length of 18.288 m. The freight task of transporting these pipes via truck from the manufacturing site, Hall Longmore in Wadeville Johannesburg, to one of the unloading sites in Vrede, is used as an in-the-field testing ground to compare the pyramid stacking method of pipes to the previously used rectangular stacking method.

Tractor semi-trailer combinations are used to transport the 3200-kg pipes. The number of pipes that can be transported per vehicle is constrained by vehicle legislation [17]. These constraints include maximum axle loadings, a permissible vehicle combination mass limit of 49 700 kg, a maximum vehicle height of 4.3 m and a maximum vehicle width of 2.6 m. Because an 18.288 m pipe is an indivisible load the transport contractors can apply for an abnormal load permit to permit them to use a vehicle longer than the legal 18.5 m limit (actual vehicle length is 21.9m).

2.2 The Pipe Loading and Unloading Equipment

Hall Longmore previously used a crane, spreader beam and slings to load pipes onto the truck trailer (See Figure 1a). The limitations of this system are:

1. Centring of the pipe in the slings takes time and effort,
2. The pipes must be stacked at least 100 mm apart on the trailer so that the slings can be removed and consequently the loaded pipes on the trailer take up more width, and

3. The pipes must be loaded onto saddles which lift the pipe up from the loading bed to allow the slings to be removed after loading. The saddles add extra weight.

To improve on the above limitations, new pipe loading equipment was designed: a trellis spreader beam was used which could allow for three pipes to be lifted at once (See Figure 1b). A further advantage of the new equipment is that the pipes are fastened onto the spreader beam with end attachments which allow for the pipes to be loaded next to each other. The loading spaced width is reduced and the pipes can be loaded into a pyramid stacking configuration as opposed to the rectangular stacking configuration which must be used with saddles.

![Figure 1 - Pipe spreader beams (a) for a single pipe with slings (b) for 3 pipes with end attachments](image)

At the offloading site, a multi-terrain loader with a vacuum lifter is used to offload the pipes. The vacuum lifter obviates the need for hooks, chains and slings using pads to suck onto and grip the pipe using a vacuum.

2.2 Rectangular Stacking versus Pyramid Stacking

The rectangular stacking of pipes using saddles is illustrated in Figure 2a. Because of the required 100 mm spacing between pipes to allow the slings to be removed, only three 610 mm diameter pipes can be loaded next to each other. Placing a fourth pipe on the first level would exceed the maximum allowable width of 2.6 m as prescribed by the South African road authorities [17].

The revised pyramid stacking configuration, made possible using the newly trialled pipe loading equipment, is illustrated in Figure 2b. Each consecutive row of pipes is placed in the gaps of the previous row which holds the pipes in place (separated with two rubber sheets to protect the pipe coatings). The rubber sheets are of negligible mass and obviate the need for saddles. The new proposed pyramid stacking method influences the loading
and unloading times and load utilisation (discussed in section 2.3), as well as the rollover safety of the vehicle (discussed in section 2.4).

2.3 Measuring the Loading and Unloading Times and Load Utilisation

The logistics metrics chosen to compare the rectangular stacking versus the revised pyramid stacking method of pipes were:

1. Loading time per item (pipe),
2. Unloading time per item (pipe),
3. Loading operations required to load truck per item (pipe),
4. Unloading operations required to unload truck per item (pipe), and
5. Number of items (pipes) transported per truck.

At the loading site the loading times were measured specifically for this loading study and recorded with a resolution of one minute. At the unloading site, the unloading inspection sheets were used to determine the unloading times. The unloading times were recorded with a resolution of five minutes.

The logistics metrics are used for an operational level comparison rather than to shape strategy, thus traditional metrics as argued by Gunasekaran and Kobu [12] are valid. A more detailed evaluation of the logistics metrics using the criteria recommended by Caplice and Sheffi [11] follows: the metrics are valid and appropriate for the comparison because the same item (a pipe of the same dimensions) is being loaded in both cases. The metrics are robust (not prone to be misinterpreted across organisations) and useful (easily understood by the decision maker). Measuring the loading times was not compatible with the existing logistics system. However, taking a small sample of loading time readings proved to be economical. Using the unloading inspection sheets ensured that measuring the unloading time was both compatible with the existing logistics system and economical.

Measuring the loading times specifically for the purposes of this investigation raises behavioural soundness uncertainties. A worker who is being measured with a stop watch for a short period would likely alter his productivity level over this short period. Additionally, management may be tempted to influence the result even at an unconscious level to
Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method
Kienhöfer, Crichton, Grobler, Dessein

manipulate a given outcome. The behavioural soundness of the unloading times metric is better in this regard as the measurements were taken from unloading inspection sheets.

A weakness of the chosen logistics metrics is the poor level of integration. The metrics are focused on the loading and unloading operation: a part of the entire logistics process and system. The effect of the improvement on the logistics system as a whole needs to be evaluated. For example: the unloading and loading times would need to be related to the required days a truck is hired for the given freight task. This weakness and future work to improve this weakness will be discussed in the conclusions.

2.4 Estimating the Rollover Tendency of the Fully Loaded Truck

Using a pyramid stacking configuration of the pipes compared to the previously used rectangular stacking configuration influences the payload centre of gravity height, (See Figure 1). The payload centre of gravity height affects the rollover tendency of the truck, which was analysed by calculating the SRT. The SRT was predicted using the estimation method developed by Elischer and Prem [16]:

$$SRT = \frac{T}{2HF}$$  

where:

- $T$ = track width [m]
- $H$ = height of centre of gravity for the vehicle (tare and payload) [m]
- $F = 1 + \frac{W_P (H_P - H_E)}{H (W_E + W_P)}$
- $W_P$ = payload mass [kg]
- $W_E$ = empty vehicle mass [kg]
- $H_P$ = height of centre of gravity of payload [m]
- $H_E$ = height of centre of gravity of empty vehicle [m]

3. RESULTS AND DISCUSSION

3.1 Comparison of Rectangular versus Pyramid Pipe Stacking using Logistics Metrics

The raw data for the loading and unloading times of the pipes are presented in the Appendix. As a summary, Table 1 compares rectangular versus pyramid pipe stacking in terms of loading and unloading times per pipe.

<table>
<thead>
<tr>
<th></th>
<th>Rectangular</th>
<th>Pyramid</th>
<th>Minimum Improvement [95% confidence]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading time per pipe</td>
<td>6.6 min</td>
<td>2.8 min</td>
<td>38% reduction</td>
</tr>
<tr>
<td>Unloading time per pipe</td>
<td>4.9 min</td>
<td>2.2 min</td>
<td>42% reduction</td>
</tr>
</tbody>
</table>

Table 1 - Rectangular versus pyramid pipe stacking in terms of loading and unloading time

The pyramidal pipe stacking has effected an approximate reduction of at least 40% in loading and unloading time per pipe transported as compared to the rectangular pipe stacking method.
The calculated loading times per pipe for the rectangular and pyramid stacking methods are based on five and two sample times respectively (See Appendix). In spite of the relatively few samples, the appropriate one-sided Welch-Satterthwaite t-test can be used to show that the pyramid pipe stacking method has reduced the loading times by 38% with a 95% confidence level. A total of sixty sample times have been used for calculating the equivalent unloading time comparison. Similarly, the appropriate one sided pooled variance t-test can be used to show that the pyramid pipe stacking method has reduced the unloading times by 42% with a 95% confidence level.

The improvement in the loading and unloading times when using a pyramid stacking method as compared to the rectangular method is due to the fewer loading and unloading operations required. A detailed breakdown of the loading and unloading operations for the rectangular versus the pyramid pipe stacking methods is given in the Appendix. A summary in Table 2 compares the rectangular versus pyramid pipe stacking method in terms of loading and unloading operations per pipe.

<table>
<thead>
<tr>
<th>Operations</th>
<th>Rectangular</th>
<th>Pyramid</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loading operations per pipe</td>
<td>2.0</td>
<td>0.8</td>
<td>60% reduction</td>
</tr>
<tr>
<td>Unloading operations per pipe</td>
<td>2.0</td>
<td>1.0</td>
<td>50% reduction</td>
</tr>
</tbody>
</table>

Table 2 - Rectangular versus pyramid pipe stacking using loading and unloading operations

The pyramid pipe stacking method has reduced the loading and unloading operations required per pipe by 60% and 50% respectively. Reducing the operations required accounts for the reduced time required for loading and unloading. The elimination of the saddles (placement and removal in the loading and unloading process) accounts for the saving in operations and time.

The introduction of the loading spreader beam (See Figure 1) which is able to load three pipes at a time has resulted in a greater reduction in loading operations as compared to unloading operations (60% to 50%) but the loading and unloading time improvements are very similar (38% compared to 42%).

The pipes were transported by two transport logistics companies. Company one used older truck trailer combinations which were heavier and could thus only transport 8 pipes using the rectangular stacking method before reaching the maximum legal mass limit of 49 700 kg [17]. Company two used newer truck trailer combinations and could thus transport 9 pipes using the rectangular stacking method. Using the pyramid pipe stacking method, both companies freed up approximately 900 kg (6 x 150 kg per saddle) of payload, allowing for an extra pipe to be transported. The load utilisation (defined as the number of pipes transported per truck) increased by 12.5% and 11.1% for company one using older trucks and company two using newer trucks respectively, using the pyramid pipe stacking method (See Table 3).
Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method

Kienhöfer, Crichton, Grobler, Dessein

<table>
<thead>
<tr>
<th></th>
<th>Rectangular</th>
<th>Pyramid</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load utilisation older trucks</td>
<td>8</td>
<td>9</td>
<td>12.5% increase</td>
</tr>
<tr>
<td>Load utilisation newer trucks</td>
<td>9</td>
<td>10</td>
<td>11.1% increase</td>
</tr>
</tbody>
</table>

Table 3 - Rectangular versus pyramid pipe stacking using loading utilisation

3.2 Comparison of Rectangular versus Pyramid Pipe Stacking using SRT

The effect of the pyramid pipe stacking method as compared to the previously used rectangular pipe stacking on the rollover tendency as measured by the SRT is presented in Table 4. In the case of the older trucks the SRT (rollover stability) is increased by 5% and for the newer trucks the SRT has remained unchanged.

<table>
<thead>
<tr>
<th></th>
<th>Rectangular</th>
<th>Pyramid</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRT [g] older trucks</td>
<td>0.41 (8 pipes)</td>
<td>0.43 (9 pipes)</td>
<td>5% increase</td>
</tr>
<tr>
<td>SRT [g] newer trucks</td>
<td>0.39 (9 pipes)</td>
<td>0.39 (10 pipes)</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 4 - Comparison of rectangular versus pyramid pipe stacking using SRT

The results in Table 4 indicate that vehicle safety in terms of rollover stability has not been compromised by increasing the load utilisation using the pyramid pipe stacking method as compared to the rectangular stacking method.

4. CONCLUSIONS

A case study has been presented which compares the stacking configurations of pipes transported by truck trailer combinations. Trucks transporting pipes in South Africa previously only used a rectangular stacking configuration. An alternative pyramid stacking configuration was trialled. The new pyramid stacking method was shown to effect an approximate 40% improvement in loading and unloading times and allows for an extra pipe to be transported per vehicle combination. Using the SRT to estimate the rollover tendency, it was shown that these improvements did not compromise the vehicle safety. The research paper is an interesting multidisciplinary transport engineering comparison incorporating traditionally considered divergent fields of operations research and vehicle dynamics.

5. LIMITATIONS OF THIS RESEARCH AND FUTURE WORK

This case study involves analysing only one company or one industry. It may be argued that the results are thus not necessarily representative of other companies and industries. This general criticism of the case study is countered by the advantage that analysing a single company or industry allows focus on detail, highlighting interesting nuances such as the multidisciplinary nature of this problem.

The pyramid stacking method of pipes as compared to the rectangular stacking method was shown to reduce the loading and unloading times but the significance of these
Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method

Kienhöfer, Crichton, Grobler, Dessein

Improvements to the logistics system as a whole has not been evaluated. Future work may use discrete simulation to analyse the improved stacking effect on the overall time of transporting the pipes and relate this to the number of trucks required for the given freight task. Additionally, the impact of the pipe stacking method on fuel consumption and emissions is being considered.

6. ACKNOWLEDGEMENTS

The presented research has been made possible by funding from Murray & Roberts, Hall Longmore, and the National Research Foundation’s Technology and Human Resources for Industry Programme. The authors would like to thank Bahle Ngcobo for his assistance in recording loading times.

7. REFERENCES

APPENDIX

<table>
<thead>
<tr>
<th>Stacking Method</th>
<th>Pipes</th>
<th>Time [min]</th>
<th>Time per pipe [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>8</td>
<td>42</td>
<td>5.3</td>
</tr>
<tr>
<td>Rectangular</td>
<td>8</td>
<td>41</td>
<td>5.1</td>
</tr>
<tr>
<td>Rectangular</td>
<td>8</td>
<td>41</td>
<td>5.1</td>
</tr>
<tr>
<td>Rectangular</td>
<td>8</td>
<td>70</td>
<td>8.8</td>
</tr>
<tr>
<td>Rectangular</td>
<td>8</td>
<td>62</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Table 5 - Loading times using the rectangular pipe stacking method

<table>
<thead>
<tr>
<th>Stacking Method</th>
<th>Pipes</th>
<th>Time [min]</th>
<th>Time per pipe [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyramid</td>
<td>10</td>
<td>27</td>
<td>2.7</td>
</tr>
<tr>
<td>Pyramid</td>
<td>10</td>
<td>28</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Table 6 - Loading times using the pyramid pipe stacking method
## Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method

*Kienhöfer, Crichton, Grobler, Dessein*

<table>
<thead>
<tr>
<th>Load Number</th>
<th>Pipes</th>
<th>Time [min]</th>
<th>Time per pipe [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2524</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>2263</td>
<td>8</td>
<td>60</td>
<td>7.5</td>
</tr>
<tr>
<td>2262</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>2241</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1691</td>
<td>8</td>
<td>30</td>
<td>3.8</td>
</tr>
<tr>
<td>1676</td>
<td>8</td>
<td>30</td>
<td>3.8</td>
</tr>
<tr>
<td>1686</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>1673</td>
<td>8</td>
<td>50</td>
<td>6.3</td>
</tr>
<tr>
<td>1641</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1636</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1658</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1604</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1599</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1605</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1595</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1602</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1578</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1580</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>1579</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1567</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>1565</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1576</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1532</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1446</td>
<td>8</td>
<td>45</td>
<td>5.6</td>
</tr>
<tr>
<td>1478</td>
<td>8</td>
<td>30</td>
<td>3.8</td>
</tr>
<tr>
<td>1447</td>
<td>8</td>
<td>40</td>
<td>5.0</td>
</tr>
<tr>
<td>1432</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1438</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1445</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
<tr>
<td>1416</td>
<td>8</td>
<td>35</td>
<td>4.4</td>
</tr>
</tbody>
</table>

**Table 7 - Unloading times using the rectangular pipe stacking method**
Table 8 - Unloading times using the pyramid pipe stacking method

<table>
<thead>
<tr>
<th>Load Number</th>
<th>Pipes</th>
<th>Time [min]</th>
<th>Time per pipe [min]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2519</td>
<td>9</td>
<td>30</td>
<td>3.3</td>
</tr>
<tr>
<td>2255</td>
<td>9</td>
<td>30</td>
<td>3.3</td>
</tr>
<tr>
<td>2256</td>
<td>9</td>
<td>30</td>
<td>3.3</td>
</tr>
<tr>
<td>1401</td>
<td>9</td>
<td>25</td>
<td>2.8</td>
</tr>
<tr>
<td>1378</td>
<td>10</td>
<td>25</td>
<td>2.5</td>
</tr>
<tr>
<td>1374</td>
<td>9</td>
<td>25</td>
<td>2.8</td>
</tr>
<tr>
<td>1679</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1678</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1677</td>
<td>10</td>
<td>20</td>
<td>2.0</td>
</tr>
<tr>
<td>1675</td>
<td>10</td>
<td>25</td>
<td>2.5</td>
</tr>
<tr>
<td>1601</td>
<td>9</td>
<td>30</td>
<td>3.3</td>
</tr>
<tr>
<td>1583</td>
<td>9</td>
<td>25</td>
<td>2.8</td>
</tr>
<tr>
<td>1575</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1582</td>
<td>10</td>
<td>25</td>
<td>2.5</td>
</tr>
<tr>
<td>1577</td>
<td>10</td>
<td>30</td>
<td>3.0</td>
</tr>
<tr>
<td>1566</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1529</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1531</td>
<td>10</td>
<td>25</td>
<td>2.5</td>
</tr>
<tr>
<td>1530</td>
<td>10</td>
<td>35</td>
<td>3.5</td>
</tr>
<tr>
<td>1534</td>
<td>10</td>
<td>35</td>
<td>3.5</td>
</tr>
<tr>
<td>1484</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1482</td>
<td>9</td>
<td>25</td>
<td>2.8</td>
</tr>
<tr>
<td>1483</td>
<td>10</td>
<td>15</td>
<td>1.5</td>
</tr>
<tr>
<td>1480</td>
<td>10</td>
<td>20</td>
<td>2.0</td>
</tr>
<tr>
<td>1486</td>
<td>9</td>
<td>30</td>
<td>3.3</td>
</tr>
<tr>
<td>1485</td>
<td>9</td>
<td>20</td>
<td>2.2</td>
</tr>
<tr>
<td>1439</td>
<td>10</td>
<td>20</td>
<td>2.0</td>
</tr>
<tr>
<td>1443</td>
<td>10</td>
<td>15</td>
<td>1.5</td>
</tr>
<tr>
<td>1405</td>
<td>9</td>
<td>15</td>
<td>1.7</td>
</tr>
<tr>
<td>1413</td>
<td>10</td>
<td>15</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Improving Loading Times and Truck Rollover Tendency Using a Pyramid Stacking Method

Kienhöfer, Crichton, Grobler, Dessein

<table>
<thead>
<tr>
<th>#</th>
<th>Rectangular Loading*</th>
<th>Pyramid</th>
<th>Rectangular Unloading</th>
<th>Pyramid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Move saddle with crane</td>
<td>Lift pipes (3)</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>2</td>
<td>Centre and lift pipe</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>3</td>
<td>Centre and lift pipe</td>
<td>Strap layer</td>
<td>Move saddle with vacuum</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>4</td>
<td>Centre and lift pipe</td>
<td>Lift pipes (3)</td>
<td>Move saddle with vacuum</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>5</td>
<td>Strap layer</td>
<td>Strap layer</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>6</td>
<td>Move saddle with crane</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>7</td>
<td>Move saddle with crane</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>8</td>
<td>Centre and lift pipe</td>
<td>Strap layer</td>
<td>Move saddle with vacuum</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>9</td>
<td>Centre and lift pipe</td>
<td></td>
<td>Move saddle with vacuum</td>
<td>Lift pipe (1)</td>
</tr>
<tr>
<td>10</td>
<td>Centre and lift pipe</td>
<td></td>
<td>Lift pipe (1)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Strap layer</td>
<td></td>
<td>Lift pipe (1)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Move saddle with crane</td>
<td></td>
<td>Lift pipe (1)</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Move saddle with crane</td>
<td></td>
<td>Move saddle with vacuum</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Centre and lift pipe</td>
<td></td>
<td>Move saddle with vacuum</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Centre and lift pipe</td>
<td></td>
<td>Move saddle with vacuum</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Strap layer</td>
<td></td>
<td>Move saddle with vacuum</td>
<td></td>
</tr>
</tbody>
</table>

Table 9 - Loading and unloading operations for the rectangular versus the pyramid pipe stacking methods

*The numbered operations for loading are illustrated in Figure 2.
LEON HEALTHCARE: A CASUALTY OF INEFFICIENCY

Ronald A. Mandavha\(^1\) and Dieter Hartmann\(^2\)

\(^1\)Generation Business Engineering, Systems Integration
Eskom Enterprises, South Africa
Ronald.Mandavha@eskom.co.za

\(^2\)School of Mechanical, Industrial and Aeronautical Engineering
University of the Witwatersrand, South Africa
Dieter.Hartmann@wits.ac.za

ABSTRACT

This paper investigates the perception that public healthcare is less efficient or less effective than private healthcare. This is done by outlining the causes of inefficiency inherent to patient flow processes in healthcare. These causes were identified by conducting a benchmarking process between Public and Private Healthcare Institutions, using qualitative and quantitative research methods. Doctors and staff were interviewed and the operations of the medical casualty were observed. A time, work and motion study was performed. On the basis of gathered data, a benchmarking exercise was done and improvements were simulated.

This study found that the principal disparities between the two sectors fell into one of five categories, namely the institutions’: customer orientation, morale, facility layout, doctor scheduling or other resource issues.

In particular it was found that doctor scheduling does not target peak patient flow and that together with inefficient layouts, causes significantly longer transportation and overall stay times. Further it was found that soft issues reduce the quality and consistency of service.

The positive aspects that were identified from the Private Healthcare Institutions during the benchmarking process were incorporated into the patient flow process of Public Healthcare Institutions. Several operational improvements were tested using a simulated model of the medical casualty unit in the public institution.

Inefficiencies were found in both sectors and it can be concluded that they can benefit greatly from further involvement from industrial engineers.

OPSOMMING

Hierdie artikel ondersoek die persepsie dat openbare gesondheidsdienste minder effektief en doeltreffend as die privaatsektor is. Ons identifiseer die oorsake van swak doeltreffendheid wat die vloei van pasiënte in ‘n hospitaal beïnvloed. Hierdie oorsake is geïdentifiseer deur kwalitatiewe en kwantitatiewe navorsingsmetodes en die twee sektore is met mekaar vergelyk. Onderhoude is met doktors en ander personeel gevoer. ‘n Tyd - bewegings- en werkstudie van die mediese ongevalle eenheid is gedaan en verbeteringe is gesimuleer.

Ons het gevind dat die hoof verskille tussen die sektore in een van vyf kategorieë val, naamlik die hospitaal se: filosofie teenoor kliëntediens, moraal van die personeel, fasiliteitsuitleg, dokter skedulering en ander probleme met hulpbronne.

Ons het gevind dat dokters se werkskedules nie die tydperke van maksimale pasiëntevloei teken nie, en dat dit, saam met ondoeltreffende fasiliteitsuitleg die algehele verblyf tyd ingrypend verleng. Verder het ons sagte kwessies gevind wat die kwaliteit en konsekwentheid van dié diens benadeel.

Die positiewe aspekte van beide sektore is in ons simulasiemodel getoets, en verbeteringe is so in die mediese ongevalle eenheid ingevoer.
1. INTRODUCTION

The South African healthcare system is divided into two sectors, the tax funded public sector and the corporate private sector[1]. The public sector provides a service to the general population while the private sector provides a service to individuals that are able to afford the cost of private healthcare. Consequently, the public sector is large compared to the private sector. As the public sector provides healthcare services to most of the South African population, there has been a drive, since 1994, to increase access to healthcare for the poor, to improve the quality of healthcare and to make the public sector systems more cost effective[2]. Public institutions face serious challenges. This can be seen in high infant mortality rates and recent industrial action by staff member for better working conditions [3][4]. All these challenges have created a perception that public healthcare is less effective and less efficient. Whilst it is also held that private healthcare is more efficient and effective, the cost of attaining this service is high. This paper aims to investigate the perception that public healthcare is less efficient or less effective than private healthcare. A benchmarking process is conducted between Public and Private Healthcare Institutions. A Value Stream Map for both public and private healthcare institution was also identified.

The investigation was carried out in the medical casualty department of two healthcare institutions. The Medical casualty department is where patients who are “feeling sick” go to, i.e. a patient with a heart attack, stroke or malaria. This department was chosen as it has long queues and high length of stay for patients. This compromised the safety of patients, the quality of the healthcare service, patient throughput and staff morale.

The challenges that face South Africa’s healthcare system are not unique. Attempts to improve the healthcare industry are being made worldwide. Innovations and other quality improvement techniques continue to improve healthcare. However, the practice that seems to have offered the most significant improvement in patient care, to date, is the time tested Lean manufacturing philosophy [5]. Consequently, leading healthcare organisations across the world are beginning to implement Lean [6][7]. The Lean principle helps organisations to eliminate waste, streamline processes and cut costs [5]. It is an integrated set of industrial principles that emerged in the post World War II Japanese automotive industry and gained traction in the USA in the 1970s [8]. Lean eliminates waste by taking out unnecessary processes and redirecting human effort towards value adding business operations [5].
2. METHOD

The steps followed when conducting the project are shown in Figure 1. A seven step methodology was developed. The investigation involved using qualitative and quantitative research methods (step 1); this included interviewing personnel, conducting general observations as well as conducting time, work and motion studies. The data from step 1 and step 2 were analysed and a benchmarking process was conducted. Credible research articles were also sourced, this, together with best practice obtained in private healthcare institution were used to build simulation models for improving public healthcare. The most feasible option was selected; conclusions and recommendation were then made.

3. RESULTS

3.1 Observations

When investigating both institutions, several observations were made. Some of the observations were later confirmed by the time study. The investigation found that:

3.1.1 The principal disparities between the two institutions fell into one of five categories, namely the institutions’ customer orientation, staff morale, facility layout, doctor scheduling or other resource issues.

3.1.2 One obvious difference between the two institutions is that a private hospital is more customer-orientated than a public hospital. This is self explanatory, as such an institution has a commercial driver, and customer satisfaction is thus fundamental. By contrast, the non commercial nature of public healthcare doesn’t prioritise patient satisfaction [9]. This creates challenges in these institutions as patients and their relatives become aggressive towards seemingly apathetic staff. This increases stress levels for staff whilst decreasing their morale.

3.1.3 The facility layout in private institutions is simple and centralises all the different sections, whereas in the public healthcare layout, the different sections are widely distributed. Appendix A shows the facility layouts for public healthcare institutions. Notably in the public institution, an ill patient needs to walk just under 700 metres to receive full treatment.
3.1.4 The flow of patients, from patient arrival to patient departure, is different and creates challenges too. Patients in public healthcare institution are screened by a doctor before they can be admitted for assessment [10]. The screening process ensures that patients will receive care where they reside and that patients are not turned away while they are in life threatening medical conditions. All patients in private healthcare are admitted without screening. Figure 2 and Figure 3 show the flow of patient in a public and a private healthcare institution, respectively. It is important to note, that in these figures, a shaded station represents a delay of more than an hour.

Figure 2 - Flow of patients in public healthcare institutions
Staff levels and shift structures also differ between the two sectors. Private healthcare employs more nurses than public healthcare institutions (Table 1 - Available staff members) - this is relevant, as nurses can perform tasks such as bandaging, administering of medication, but also by taking samples and moving patients, samples and lab results, thus taking pressure off doctors [11]. Furthermore, Doctor scheduling is different and below optimal compared to [12]; this is shown in institutions for both types of institution.

Table 1 shows the available resources during day and night shifts for both types of institution.
Lean Healthcare: A Casualty Of Inefficiency
Mandavha, Hartmann

Table 1 - Available staff members

<table>
<thead>
<tr>
<th>Staff</th>
<th>Public Healthcare</th>
<th>Private Healthcare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Night shift</td>
<td>Day shift</td>
</tr>
<tr>
<td>Assessment Doctors</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Screening Doctor</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nurses</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Receptionist</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

3.1.2 Results from the time study

Figure 5 shows the average normal time that patients will spend in the different sections in a hospital for both healthcare institutions. This time includes patient’s queuing time waiting for service and the actual services times in the different sections. These figures show the average total time that patients spend in the system. An average patient will spend more than three times longer in a public hospital, than in a comparable private hospital. It is also significant that many of the delays that exist in the public sector don’t exist in the private sector, such as the delay at the cashier.

![Figure 5 - Normal time patients spend in different area for both healthcare institutions](image)

The average actual (Value Added (VA)) service time is shown in Figure 6 for both public healthcare institution and private healthcare institution.

![Figure 6 - Normal service time for both healthcare institutions](image)
Figure 7 and Figure 8 show the total average normal time that patients are spending in a public healthcare institution and private healthcare institution over a 24 hour period, respectively.

**Figure 7** - Average total time patients spend in a public healthcare institution

![Figure 7](image)

**Figure 8** - Average total time patients spend in a private healthcare institution

![Figure 8](image)

Figure 9 and Figure 10 show that doctor shift patterns do not address patient arrivals patterns, thus resulting in accumulation of patients. This confirms the observation made in section 3.2.1.

**Figure 9** - Queue length vs. Number of doctors on duty for public healthcare

![Figure 9](image)
Figure 10 - Queue length vs. Number of doctors on duty for private healthcare

The Value stream map for both public and private healthcare institution were identified, these are shown in Figure 11 and Figure 12 respectively.

Figure 11 - Value stream map for Public healthcare institutions

Figure 12 - Value stream map for Private healthcare institutions
3.2 Simulation results
Best practice benchmarks from the private sector were incorporated into a simulation model of the public institution. Different simulation models were designed tested and validated. On the basis of best resource utilisation and shortest queue length, the optimal option was selected.

Upgrades included a change of facility layout with a new reception area, an improved shift system and the reallocation of staff. Two staff nurses are employed in the screening section. These nurses will also perform the screening function. This freed up an additional doctor, who is then allocated to the assessment section. An additional staff member is placed at the reception area to attend to patients as they arrive at the medical casualty.

The results from the optimal scenario are presented below. Figure 13 shows the total average time that patients spend in the improved public healthcare institution. Table 2 Appendix A shows the staff scheduling for the new system.

Figure 13 - Total Average Normal time

![Figure 13 - Total Average Normal time](image)

Figure 14 - New patients’ flow process

![Figure 14 - New patients’ flow process](image)
4. DISCUSSION
In the original site layout of the public institution, it was found that an average patient would need to walk at least 700 metres during the course of their stay at the hospital. This is obviously spread out over a considerable amount of time, but is nevertheless too great a distance for someone who is severely ill. A direct corollary to this is that hospital staff needs to walk similar distances on an ongoing basis. Clearly all of this leads to lost productivity as both doctors and patients spend a large amount of time and effort in moving between stations. The layout in the private institution is much more thoughtfully designed, with a central administration section and simplified flow that eliminates repeat journeys, for example by clustering functions and having necessary auxiliary services such as sample taking stations, radiology and the pharmacy close by. Although the unit is physically smaller, it attends to two thirds of the volumes of the government institution daily. The average patient walks less than 150 metres during their stay. Using the principles of design for patient flow, the public institution’s layout was redesigned. This results in fewer repeat journeys and crucially reduces the average travelled distance from 700 metres to 250 metres. This is a significant improvement, though we are mindful that this is still a significant burden on critically ill patients, and therefore not yet sufficient.

On average doctors in both sectors see virtually the same number of patients daily. Doctors in Private and public healthcare see an average of 19 and 17 patients per shift respectively. Because doctor scheduling in public healthcare doesn’t target peak loads, those doctors who are on duty during these times actually see as many as 35 or more patients per shift. This slows the system down and drives accumulation in the queues, consequently amplifying the backlog and sending the system further out of control. Conversely, during low activity times, more doctors are on duty and therefore they each treat fewer patients. The changes
that were made to the scheduling system anticipate high levels of patient arrival and staffs the emergency room accordingly. This was augmented by one human resource relocation. A doctor was moved from the screening area into the assessment area. The screening role falls within the skills set of a qualified senior nurse. As a result, the simulation shows much lower levels of queue accumulation, down from ~37 to ~21 patients in system. Critically the overall time in the system is reduced dramatically from ~408 minutes to ~217 minutes. In addition to this, the actual patient load is spread more evenly between doctors, and an individual doctor is unlikely to see more than 25 patients.

One will notice that by comparison with the public institution, private institutions spend more time adding value to the patients, 53 minutes compared to 49 minutes. This is probably caused in part by the greater need for customer satisfaction and in part by the fact that doctors in private practice see a lower number of patients during busy times, and can therefore spend more time attending to patients. This is an even starker difference when expressed as a VA/NVA ratio, in private the ratio is 41% whereas in public healthcare it is 12%. With the improvements discussed above, the ratio in the government institution would rise to 23%.

5. CONCLUSIONS
The aim of this paper was to prove that public healthcare is less efficient or less effective than private healthcare. After conducting an extensive study in both types of institution, the hypothesis was found to be true.

This is shown especially in the following:

- Average queue lengths are more than three times longer in public healthcare
- During peak hours, doctors see fifty percent more patients in public healthcare
- Ineffective layouts cause a great deal of waste, as doctors and patients need to walk long distances; which doesn’t add value.

We conclude that Lean is applicable to healthcare and if implemented, can improve healthcare efficiency; implementing Lean could yield the following:

- Reduced patient length of stay at hospitals
- Reduced patient queues
- Improved morale
- Improved patient satisfaction
- Improved patient safety
- Reduced operating costs
- Improved quality of service [13]

It can be concluded that the success of the new system lies in the implementation of the new facility layout together with an effective and efficient reception area, and the new doctor schedule. The facility layout and the reception area are responsible for eliminating most of the identified disturbances that affect the availability of the nurses and doctors. The doctor schedule was identified as the biggest contributor to the bottleneck in the assessment section. Therefore, it is concluded that the new system will eliminate most of the identified waste thereby improving efficiency for the public healthcare institution.

Inefficiencies were found in both institutions and it can be concluded that the healthcare sector can benefit greatly from further involvement of industrial engineers.
6. REFERENCES

Lean Healthcare: A Casualty Of Inefficiency  
Mandavha, Hartmann

7. APPENDIX
Appendix A

Table 2 - Staff scheduling for the new system

<table>
<thead>
<tr>
<th>Shift</th>
<th>Simulation From</th>
<th>Simulation Until</th>
<th>Current From</th>
<th>Current Until</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shift 1</td>
<td>07:00</td>
<td>16:00</td>
<td>07:00</td>
<td>16:00</td>
</tr>
<tr>
<td>Shift 2</td>
<td>07:00</td>
<td>19:00</td>
<td>07:00</td>
<td>16:00</td>
</tr>
<tr>
<td>Shift 3</td>
<td>19:00</td>
<td>07:00</td>
<td>08:00</td>
<td>19:00</td>
</tr>
<tr>
<td>Shift 4</td>
<td>19:00</td>
<td>07:00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shift 1A</td>
<td>07:00</td>
<td>15:00</td>
<td>08:00</td>
<td>16:00</td>
</tr>
<tr>
<td>Shift 1B</td>
<td>07:00</td>
<td>15:00</td>
<td>08:00</td>
<td>16:00</td>
</tr>
<tr>
<td>Shift 2</td>
<td>09:00</td>
<td>19:00</td>
<td>11:00</td>
<td>19:00</td>
</tr>
<tr>
<td>Shift 3</td>
<td>12:00</td>
<td>20:00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Shift 4</td>
<td>20:00</td>
<td>07:00</td>
<td>19:00</td>
<td>08:00</td>
</tr>
</tbody>
</table>
MICRO-MATERIAL HANDLING THROUGH THE MANIPULATION OF VAN DER WAALS FORCES
BY THE GENERATION OF A PREDETERMINED SURFACE TOPOGRAPHY.
S. Matope¹*, A.F. Van der Merwe², R. Nemutudi ³, M. Nkosi⁴, M. Cele⁵ and M. Maaza⁶
¹Department of Industrial Engineering
University of Stellenbosch, South Africa
smatope@sun.ac.za

²Department of Industrial Engineering
University of Stellenbosch, South Africa
avdm@sun.ac.za

³Department of Material Science
iThemba Labs, South Africa
rudzi@tlabs.ac.za

⁴Department of Material Science
iThemba Labs, South Africa
mlungisin@tlabs.ac.za

⁵Department of Material Science
iThemba Labs, South Africa
hmcele@tlabs.ac.za

⁶Department of Material Science
iThemba Labs, South Africa
maaza@tlabs.ac.za

ABSTRACT
This paper presents an investigation into the manipulation of Van der Waals forces in micro-
material handling by the variation of the root mean square (rms) surface roughness values
of the interacting surfaces. Electron beam evaporation (e-beam) is used in the generation
of the required surface topography for the interactive surfaces involved in the picking and
placing of micro-work parts. The Rumpf model of rms values of roughness is used to
determine the Van der Waals forces exerted by the generated profiles. The surfaces are
generated on a silicon base using e-beam deposition rates ranging from 0.6 to 1.2
Angstroms/second in a vacuum environment of 2x10⁻⁶ millibars. The surfaces with high rms
were found to be suitable for the pick-up position, and those with low rms for the
placement position in an effective material handling system.

Key words: Surface topography, Van der Waals forces, micro-material handling.

¹The author was enrolled for a PhD (Manufacturing Engineering) degree in the Department of Industrial
Engineering, University of Stellenbosch and is the corresponding author.
²The author is the Chairman of the Department of Industrial Engineering, University of Stellenbosch.
³⁴The authors are researchers in the Material Science Research Department.
* Corresponding author
1. INTRODUCTION

This paper seeks to explore the effective use of Van der Waals forces in the handling of micro-materials by the generation of a predetermined surface topography. The demand for micro-products has significantly increased in the twenty-first century (Feddema [3]). However, their handling during manufacturing is complicated because the handling laws change as one moves from the macro-manufacturing world into the micro-manufacturing (Feddema [3]). The natural forces, for example gravity, are largely experienced in the macro-environment; but are less dominant in the micro-region (Thoreson [11], Van der Merwe & Matope [12]). On the other hand Van der Waals forces, capillary forces (or surface energy forces) and electrostatic forces are predominant in the micro-environment (Rabinovich [8]). The capillary forces can be eliminated by decreasing the humidity and the electrostatic force by using an anti-static equipment. The greatest challenge is imposed by Van der Waals forces during the picking and placing of micro-materials. Fortunately surface roughness has a large bearing on the Van der Waals experienced by a body (Thoreson[11]). Hence this paper investigates the manipulation of the Van der Waals forces from the surface roughness perspective.

Surface roughnesses of interacting surfaces can be varied in several ways. Mechanical, chemical or laser beam machining are among the paramount methods of generating a predetermined surface roughness. However, they have the disadvantages of changing the physical, mechanical and chemical properties of the prepared surfaces. Hence this paper explores electron beam evaporation (e-beam) method in the preparation of a controlled surface topography for purposes of micro-material handling.

1.1 Micro-material handling.

Micro-material handling refers to the picking and placing of micro-materials in a manufacturing environment (Fukuda [4], Matope & Van der Merwe [7]). Micro-materials range in size from 1 micron to 1 millimetre. Micro-material handling system includes a micro-work part, the pick-up position, the micro-gripper and the placement position as shown in Figure 1. The mechanism in a micro-material handling system is such that micro-work parts are picked by a gripper from the pick-up position and released precisely on the placement position during the assembling of a micro-product. With the advent of micro-factories (Okazaki et al [4]), micro-material handling effectiveness and efficiency need to be improved.

![Figure 1 - Micro-material handling system: the pick-up position (A), micro-work part (B), the micro-gripper (C) and the placement position (D).](image)

1.2 The Van der Waals forces

Van der Waals forces are short-range forces, acting when surfaces are sufficiently close together, and are due to spontaneous electrical and magnetic polarisations that cause a fluctuating electromagnetic field within the medium and the gap between the surfaces.
involved (Debrincat [1], Zhang [13]). These forces are generally attractive in nature (Debrincat [1]). A releasing problem arises when a micro-work part is picked (Fukuda[4]). The challenge is how to release the micro-material since the Van der Waals forces have an adhesive effect when particles are in contact. This problem is addressed in this paper by the generation of a predetermined surface finish of the micro-work part and the interacting surfaces involved during material handling.

1.3 Root mean square (rms)

Root mean square (rms) is a measure of the roughness of a given surface topography. It is the square root of the arithmetic average of the squares of the asperity heights along equally spaced points on a trace (Rabinovich [9]). The heights are measures from a given datum.

2 ELECTRON BEAM EVAPORATION DESCRIPTION

The e-beam evaporation machine’s main parts include the vacuum chamber, crucible, electron emitter, magnetic electron deflector, and the sample holders. The target materials (materials to be evaporated) are put into the crucible, and the samples are mounted onto the sample holders. These are then assembled in the evaporator. This is then closed and evacuated for twenty-four hours to create the required vacuum pressure. The purpose of the vacuum is to prevent impurities from contaminating the specimens and to lower the vaporisation temperature of the materials to be evaporated. A tungsten filament electron gun produces, at a predetermined rate, electrons that are focused onto the crucible by a magnetic field. Upon impinging into the material, the electrons’ energy is transferred into the target material such that evaporation starts. The vapour of the material ascends at a predetermined rate and deposits onto the mounted specimens. The duration of the deposition is again predetermined in order to achieve a predetermined surface topography.

E-beam has several advantages which include controllable material deposition rate, impurity-free deposition, reduced mechanical and chemical deformations, and generation of a predetermined topography with uniform rms values.

3 EXPERIMENTAL DETAILS

The apparatus and materials used included the e-beam machine, silicon substrate, cleaning agents and the target material. The target material in this case was copper. The cleaning agents were methanol, acetone, trichlorethylene, de-ionised water and 20% hydrofluoric acid solution.

The silicon substrate was washed in methanol, acetone, trichlorethylene; and then in acetone and methanol respectively. Later it was washed in de-ionised water and then in 20% hydrofluoric acid solution. The cleaning process removed the oxidation layers and other impurities so as to achieve an effective e-beam deposition of the target material.

Copper was then placed into the crucible and the silicon substrates were mounted into the e-beam evaporator. The machine was closed and then evacuated for twenty-four hours. Thereafter the e-beam deposition of copper onto the silicon substrate was done. Six copper depositions of different durations were done on six silicon substrates and the experimental results are shown in Table 1 and in Figures 2 to 7. The deposition rates were within the range of 0.6-1.2 Angstroms/second (Å^2/s) and at a vacuum pressure approximately 3x10^-6 mbar and at a current less than 10mA. The following section details the experimental results.
4 EXPERIMENTAL FINDINGS

Table 1 shows the durations of e-beam depositions, the resulting copper film thicknesses and their rms values on the silicon substrates.

**Table 1 - The deposition of copper on silicon substrate**

<table>
<thead>
<tr>
<th>Holder</th>
<th>Duration of deposition, minutes</th>
<th>Thickness of deposition, Angstroms (Å)</th>
<th>rms (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>132</td>
<td>2.4</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>378</td>
<td>2.72</td>
</tr>
<tr>
<td>C</td>
<td>7</td>
<td>510</td>
<td>38.3</td>
</tr>
<tr>
<td>D</td>
<td>10</td>
<td>654</td>
<td>51.8</td>
</tr>
<tr>
<td>E</td>
<td>15</td>
<td>840</td>
<td>75.8</td>
</tr>
<tr>
<td>F</td>
<td>20</td>
<td>984</td>
<td>217</td>
</tr>
</tbody>
</table>

Figures 2, 3, 4, 5, 6, and 7 reveal the corresponding atomic force micrographs for the deposition times in Table 1.
Micro-material handling through the manipulation of van der Waals forces
Matope, van der Merwe, Nemutudi, Nkosi, Cele, Maaza

Figure 4 - Atomic Force Micrograph of copper deposited for 7 min.

Figure 5 - Atomic Force Micrograph of copper deposited for 10 min.
5 **ANALYSIS**

Figure 8 shows a plot of the thickness of the deposited film against the duration of deposition. It is deduced that the rate of film thickness growth decreases with time. However, the $rms$ value increases with time as shown in Figure 9. Therefore, the longer the deposition, the greater is the surface roughness.
5.1 Rumpf’s equation.
Rumpf’s equation (1) presents the relationship between surface roughness (in \( rms \)) and Van der Waals force intensity (Rabinovich [9]). This equation represents the interaction between a spherical particle and flat surface (Li et al [6], Komvopoulos [5], 1996; Suresh & Walz [10], Eichenlaub [2]).

\[
F = \frac{AR}{6D^2} \left[ \frac{1}{1 + R/1.48rms} + \frac{1}{1 + 1.48rms/D} \right]^2
\]

\[
\text{………………..(1)}
\]
Where $F$ is the Van der Waals force, $A$ is the projected area of contact, $R$ is the radius of the spherical micro-work part, $D$ is the separation distance between the micro-work part and the interacting surface as shown in Figure 10. It should be noted that $D$ varies around a spherical micro-work part.

Rumpf's equation has two main components in its brackets, namely:

The component which reveals the effect of surface roughness (in $rms$) in the noncontact region (shown in Figure 10) given by

$$\frac{1}{1 + \frac{R}{1.48rms}}$$

(2)

And the effect of surface roughness for a contact interaction given by

$$\frac{1}{[1 + \frac{1.48rms}{D}]^2}$$

(3)

Equation (1) is also referred to as the Rumpf-Rabinovich's equation (Li et al[6]).
Figure 11 - Van der Waals force intensity against surface roughness

An analysis of the limits of the two main Rumpf’s components reveals that the contact component supersedes the noncontact since it is a power function. Hence the effect of the non-contact component may be ignored in practical cases as in this case.

When $rms$ approaches a zero value the contact component approaches a unit value, giving a maximum of Van der Waals forces exerted between interacting surfaces as shown in Figure 11 (a natural logarithmic plot of the Van der Waals forces’ intensity against the $rms$ value). This would suit the pick-up action. Hence the interactive surface between the gripper and the micro-work part should have a less $rms$ value than that between the pick-up position and the micro-work part.

Conversely, when the $rms$ value approaches infinite the contact component approaches a zero value, leading to less exertion of Van der Waals forces as depicted in Figure 11. Therefore, for a releasing action to be realised, the interactive surface between the gripper and the micro-work part should have a higher $rms$ value than that between the placement position and the micro-work part.

It was then synthesised that the e-beam depositions shown in Figure 6 and Figure 7 form effective and efficient pick-up positions. The topographies in Figure 2 and Figure 3 are good for a placement position. Figure 4 and Figure 5 topographies are suitable for a gripper’s surface roughness in a micro-material handling system since they have medium $rms$ values.

6 CONCLUSIONS

Conclusively, the e-beam evaporation proved to be a viable method of generating a predetermined surface topography suitable for an effective and efficient micro-material handling system. The 2 minute and 5 minute depositions had low $rms$ and were observed to make better placement positions than the others because they exerted larger Van der Waals forces than the others (Figure 11). On the contrary, 15 minute and 20 minute depositions proved to be suited for the picking place since they had greater values of $rms$ which made them exert less Van der Waals forces.

The 7 minute and the 10 minute depositions had medium $rms$ values (Table 1, Figure 11), hence they were suited for the micro-gripper’s surfaces. Since their $rms$ values were less than that of the picking position, they exerted more Van der Waals forces resulting in the
Micro-material handling through the manipulation of van der Waals forces
Matope, van der Merwe, Nemutudi, Nkosi, Cele, Maaza

picking up of a micro-work part. Conversely, as they reached the placement position, the micro-work part was released because the gripper exerted comparatively a less Van der Waals force than the former. It should be noted that the micro-work part should have a uniform surface topography on all sides for an effective material handling system to be realised.

6. REFERENCES

WAREHOUSE PRE-POSITIONING FOR DISASTER RELIEF IN SOUTHERN AFRICAN DEVELOPMENT COMMUNITIES

E. Meyer¹, W.L. Bean² and V.S.S. Yadavalli³

¹Department of Industrial and Systems Engineering
University of Pretoria, South Africa
eliznameyer@gmail.com

²Logistics and Quantitative Methods
CSIR Built Environment, South Africa
wbean@csir.co.za

³Department of Industrial and Systems Engineering
University of Pretoria, South Africa
sarma.yadavalli@up.ac.za

ABSTRACT

At any given time people around the world are adversely affected by the impact of current or recent disasters. This is increasingly true as increase in population density, population migration, and technological development amplify the severity, and in some cases the frequency, of disasters. Developing countries, such as those that constitute the Southern African Development Communities (SADC) are especially vulnerable.

This paper aims to improve disaster response in SADC countries through the pre-positioning of emergency supply warehouses. A Maximal Covering Location Problem (MCLP), that includes spatial objects rather than single points, partial coverage, and weights assigned to disaster areas, will be used to suggest potential locations for warehouses that will maximise the coverage of disaster-prone areas in the SADC regions.

OPSOMMING

Daar kan op enige tydstip iewers in die wêreld rampe plaasvind. Hierdie bedreiging raak net meer relevant soos die groei in bevolkingsdigtheid, bevolkingsverskuiwings en tegnologiese ontwikkeling dit waarskynlik maak dat rampe meer dikwels aangetref sal word en dat hulle meer ernstige gevolge sal hê, veral in ontwikkelende lande soos die Suider-Afrikaanse Ontwikkelings Gemeenskappe (SAOG).

Hierdie artikel fokus op die pre-positonering van nood voorraad pakhuise in SAOG lande om te help met die bespoeding van rampreaksie. ’n Maksimale Dekking Plek Probleem (MDPP) wat ruimtelike voorwerpe in plaas van enkele punte, gedeeltelike bedekking, en gewigte aan rampgebiede toeken insluit sal gebruik word om plekke te vind vir die pakhuise in SAOG lande wat die dekking van ramp sensitiewe gebiede verhoog.

¹ The author is enrolled for a B Eng (Industrial) degree in the Department of Industrial and Systems Engineering, University of Pretoria.
² The author is a researcher at the Logistics and Quantitative Methods Competency Area, CSIR Built Environment.
³ The author is the Head of the Department of Industrial and Systems Engineering, University of Pretoria.
*Corresponding author.
1. INTRODUCTION

"Nations have passed away and left no traces,
Any history gives the naked cause of it -
one simple reason in all cases;
they fell because their people were not fit."

Rudyard Kipling

1.1 Background

At any given time people around the world are adversely affected by the impact of current or recent disasters. Each year hundreds of disasters occur at locations without sufficient local capability or resources to prevent the death and destruction of their communities [1]. The duration of these disasters range from a few seconds to many years. The severity of the impact of these events varies according to the degree to which man has created an environment vulnerable to damage. That is, an environment in which life and property are at risk [2].

A definition of disaster, including both the physical event and the social and economic consequences is given by Fritz [3]:

"...An event, concentrated in time and space, in which a society or community undergoes severe danger and incurs such losses to its members and physical appurtenances that the social structure is disrupted and the fulfilment of all or some of the essential functions of the society is prevented."

From as early as 1986, it was acknowledged that the concern about disasters is becoming increasingly relevant as growth in population density, population shifts, and technological development make it likely that disasters will be encountered more frequently and that they will be more severe [4]. The Intergovernmental Panel on Climate Change (IPCC) also declared that climate change significantly alters the geographic distribution, frequency, and intensity of natural disasters [5].

According to the United Nations Development Program (UNDP) [6], the development status and the disaster risk of a country is closely related. Statistics show that 11 percent of people in underdeveloped countries exposed to natural hazards, account for more than 53 percent of the total recorded deaths. This corresponds to Dilley et al.'s [1] finding that disasters represent a major source of risk for the poor and that, in developing countries, it wipes out development gains and accumulated wealth.

By pre-positioning emergency supply warehouses in Southern African Development Communities (SADC), disaster response will be improved during the initial days of deployment following a disaster. This will ameliorate existing vulnerability and reduce death and illness resulting from hazards.

1.2 Overview of Disaster Types

The definitions of various types of disasters have been widely discussed by, amongst others, Turner and Pidgeon [7], Richardson [8], the World Health Organization [9] and the Federal Emergency Management Agency [9, 10]. In these discussions, it is established that all types of disastrous events can be classified as either natural, man-made, or hybrid.

Shaluf [11] describes these three disaster classes as follows. Natural disasters are catastrophic events resulting from natural hazards. These natural hazards result from internal (beneath the earth's surface), external (topographical), weather related
(meteorological or hydrological) and biological phenomena. Natural disasters are beyond
human control and are often termed ‘Acts of God’ and include events such as earthquakes,
tsunamis, landslides and so on.

Man-made disasters, on the other hand, are those catastrophic events that result from
human involvement and can either be sudden or long-term disasters. Sudden man-made
disasters include structural, building and mine collapses, occurring independently without
any outside force. Long-term man-made disasters tend to refer to national and
international conflicts.

The third class, hybrid disasters, is a compound of human decisions and natural forces. An
example of a hybrid disaster is the extensive clearing of jungles causing soil erosion, and
subsequently heavy rain causing landslides. The common denominator of all types of
disasters is the severity of their impact on people, property, and the environment [12].

1.3 Typical disasters in Southern African Development Communities

The vulnerability of SADC countries to disasters are exacerbated by factors such as high
levels of multidimensional poverty, the spread of HIV/AIDS, environmental degradation,
seismic activity, climate change, floods, drought, food shortages, civil strife, internally
displaced persons and refugees, slow economic growth and political instability, to name a
few [13].

The frequency and magnitude of disasters in SADC differ from country to country and,
except for a few of the above-mentioned disaster types, almost all of them occur in at least
one SADC member. These disasters claim thousands of lives, cause material losses in the
billions of dollars, and inflict a terrible toll on SADC countries, particularly as they divert
attention and resources from development needed desperately to escape poverty [14].

The disaster risk of SADC countries is increased by the countries’ development. Economic
growth and social improvement generate new disaster risks. The growth of informal
settlements and inner city slums, caused by rapid urbanisation and population growth,
leads to the increase of unstable living environments. These settlements are often located
in ravines, on steep slopes, along flood plains or adjacent to noxious or dangerous industrial
or transport facilities [6].

Because the development status and the disaster risk of a country is so closely linked, it is
important to find ways to improve disaster relief in developing countries such as SADC
countries. Widespread disaster-related mortality can affect households and communities for
years, decades, and even generations [1].

The UNDP [6] found that natural disasters can be held directly responsible for the
destruction of infrastructure, erosion of livelihoods, damage to the integrity of ecosystems
and architectural heritage, injury, illness and death. These disaster losses then interact
with and aggravates other stresses and shocks such as financial crises, political or social
conflict, disease (especially HIV/AIDS), and environmental degradation. In turn, these
stresses and shocks set back social investments intended to ameliorate poverty and hunger,
provide access to education, health services, safe housing, drinking water and sanitation,
protect the environment and enable economic investments that provide employment and
income [6].

By decreasing the time it takes to distribute aid to disaster victims, the severity of the
impact of these disasters can be reduced.

2. THE FUNCTION OF PRE-POSITIONED WAREHOUSES IN DISASTER RELIEF

Beamon [15] described the life cycle and relative resource requirements for a relief mission
by means of four distinct phases, of which the last three are corroborated by Thomas [16]
Warehouse pre-positioning for disaster relief in Southern African development communities – Meyer, Bean, Yadavalli

(Refer to Figure 1). The first phase is assessment. Minimal resources are required to identify what aid and skills are needed based on the disaster characteristics. Secondly, during deployment, resource requirements are ramped up to meet the identified need. The third phase, sustainment, describes the period of time when relief efforts are sustained at a certain level. Finally, reconfiguration is the fourth phase during which operations are reduced and then terminated.

![Figure 1 - Relief Mission Life Cycle](image)

Supply Chain Management System (SCMS) [17] reported the following, subsequent to the catastrophic earthquake in Haiti on 12 January 2010:

"Within 48 hours following the catastrophic earthquake, the SCMS team in Haiti...identified emergency requirements and...began delivering emergency kits of medicines and supplies to hospitals in Port-au-Prince [capital city of Haiti]...By January 25 [thirteen days after the earthquake], more than 67,000 pounds [about 30,400 kg] of medicines and emergency medical supplies from warehouse stock had been distributed to more than 40 health facilities...Emergency kits came from existing supplies in the warehouse and included antibiotics and other essential medicines, blood transfusion sets, first aid supplies, infusion solution and oral rehydration salts and syringes."

The challenges of global procurement in the post-disaster environment stem primarily from the time-consuming processes involved, which include competitive bidding and customs clearance, and transportation capacity requirements for shipping large quantities of bulk supplies [18]. Therefore pre-positioned warehouses are crucial in the initial days of the deployment phase. These warehouses supply the immediate demand spikes for food, water, clothing, and medical supplies [19]. Time is of the essence during the deployment phase when a few hours' delay could cost many lives.

The pre-positioned warehouses not only provide relief items during the deployment phase but also serve as storage locations and distribution centres for relief items received by the various global relief organisations during the sustainment phase.

3. FACILITY LOCATION PROBLEMS IN LITERATURE

Operations Research techniques have for years been applied to a large variety of problems to determine the optimal geographical locations for facilities (See [20], [21], [22], and [23] for surveys on facility location research). Facility location problems obtain their importance firstly from their direct impact on the system's operating cost and secondly due to their effect on the timeliness of response to the system's demand [24]. Facility location models
Models with coverage-type objectives are used extensively in facility location research and applications, especially when the primary performance criteria is response time (See [28] and [29] for discussions and reviews of covering models). In these covering-type facility location models, a source of demand can be defined as covered only if it is located within a specified distance of a facility [27].

These models seek to choose facilities among a finite set of candidate sites in such a way that all demand sources are covered with a minimum number of facilities. In disaster relief, this would imply that each potential demand point must be within a specified target response time or distance of a facility in the relief network [27]. However, to cover the entire demand area of every potential disaster scenario from pre-positioned warehouses might not be practicable. Therefore, Balci and Beamon [27] found that a maximal covering-type model that selects facility locations (pre-positioned warehouse locations) to maximise the amount of covered demand (disaster areas), subject to resource limitations (number of warehouses and costs associated with building these warehouses) is more suitable for relief chain network design.

3.1 Maximal Covering Location Problem

The Maximal Covering Location Problem (MCLP), first presented by Church and ReVelle [30], maximises the total number of people served within a maximal service distance, given a fixed number of facilities or budget limitations [27]. The MCLP has a broad range of applications and has been studied extensively. Since its original formulation, the MCLP has also been modified and customised to cater for more realistic problem situations [31].

The original MCLP formulation defines a demand point as covered if it is within a certain distance $D$ from at least one server and not covered otherwise. However, this ‘black and white’ perspective that either considers a demand point as covered or not covered may be considered somewhat crude. Alexandris and Giannikos [31], explain it as follows: the original MCLP formulation implies that demand points at a distance $D + \varepsilon$ from a server are fully covered, whereas points at a distance $D$ are not covered at all, where $\varepsilon > 0$ may be arbitrarily small. This issue has been addressed with different approaches.

Church and Roberts [32] and Berman and Krass [33] used a stepwise coverage function with break points. Pirkul and Schilling [34] and Araz et al. [35], depicted the ‘quality of service’ function and Drezner et al. [36] the coverage equivalent of the cost function. Berman et al. [37] employed a coverage function that is neither convex nor concave. In general, different levels of coverage may either be included in the constraints or in the objective function [38]. Murray [39] approached the covered or not covered dichotomy as follows. He found that if an area is partially covered by a sufficient number of servers, it is highly likely that this area will be fully covered.

Alexandris and Giannikos [31] studied the effects of demand representation on maximal covering models. Specifically, they considered the general MCLP and demonstrated that it produced results which varied considerably according to the definition of the demand space and also implied considerable coverage gaps. They introduced an integer programming model based on the representation of demand through spatial objects rather than single points, which exploits the capabilities of Geographic Information Systems (GIS). Their model also includes Murray’s [36] model for partial coverage. Results of this model provided better, more robust coverage using fewer servers than the conventional maximal covering models. The coverage gaps also reduced significantly [31].

The model developed by Alexandris and Giannikos [31] to address the MCLP with spatial objects, was found to be the most suitable for the pre-positioning of warehouses in SADC.
By using a grid, disaster areas can be defined more correctly than just using single points. The inclusion of the partial coverage mechanism ensures better coverage, using fewer warehouses. The model also includes a weight that is assigned to each demand area, ensuring that the model maximises coverage of areas affected more severely by disasters.

4. SADC DISASTER DATA GATHERING

To pre-position emergency supply warehouses to aid in disaster relief, data on different disasters occurring in the SADC countries is required. This data is used to assign weights to the different areas.

The Disaster Events Database (EM-DAT, available at [40]) of the Centre for Research on the Epidemiology of Disasters (CRED), was developed in 1988 and contains records of disasters that required international assistance ever since 1900 to present [40, 41, 42]. The main objective of the database is to serve the purposes of humanitarian action at national and international levels. It is an initiative aimed to rationalise decision making for disaster preparedness, as well as providing an objective base for vulnerability assessment and priority setting [40]. EM-DAT uses various sources, including United Nations (UN) agencies, non-governmental organisations, insurance companies, research institutes and press agencies, as sources on disaster events, and includes information on fatalities, injuries, homelessness, persons affected and financial losses [40, 42].

Only data related to natural disasters is detailed enough to map within each country. However, as discussed in section 1.3, natural disasters are considered to exacerbate other disasters, and contribute to cause them, therefore the risk factor that will be used to determine where warehouses should be pre-positioned will consider only natural disaster data. By pre-positioning warehouses in this way, the aid provided after the occurrence of a natural disaster, will help in reducing the occurrence and the extent of other disasters.

There are six major natural hazards that cause disasters. They are earthquakes, volcanoes, landslides, floods, drought, and tropical cyclones. Volcano data is excluded from the data used to solve the model because, although data exist, the variability of volcanic hazards is too complex to be entered into a general model, according to the UNDP [6]. The UNDP [6] identified African states as having the highest vulnerability to drought, and a high concentration of deaths associated with it, but because it is a slow onset disaster that can occur anywhere [5] it is also excluded. The impacts of drought are very different from those of sudden impact disasters [5], and assessment by the UNDP [6] strongly reinforced field study evidence that the translation of drought to death is mediated by armed conflict, internal displacement, HIV/AIDS, poor governance and economic crisis (in other words, drought rarely leads directly to mortality). The warehouses are pre-positioned to enhance relief response during sudden impact disasters and therefore will not be of much use for slow impact disasters such as drought.

Peduzzi and Deichmann [5] analysed (using EM-DAT) the mortality and economic loss risk, experienced in past disasters, for multiple hazards, including earthquakes, landslides, floods, and tropical cyclones. They determined the disaster risk for a given location with the probability that a hazard event of a given magnitude will occur, the number of people exposed or the value of the assets exposed, and the vulnerability. They calculated risks for grid cells, rather than for countries as a whole, and thus were able to estimate risk levels at sub-national scales.

Human mortality data provides a clearer picture of the risk of disasters, especially when looking at developing countries [6]. Therefore the dataset (available for download at [43] in different formats), developed by Peduzzi [5] and used to create the global map of the distribution of multiple hazards’ mortality risk (GIS analysis and cartography [44]), is used to determine the weights assigned to the different disaster areas in SADC.
5. THE MAXIMAL COVERING LOCATION PROBLEM WITH SPATIAL OBJECTS MODEL

5.1 Notation

Parameters:

- \( S \) = number of warehouses to be located.
- \( b \) = minimum acceptable coverage percentage (by a warehouse).
- \( \theta \) = minimum number of partial coverage pre-positioned warehouses needed for complete coverage.

\( a_{ij} \triangleq \begin{cases} 1 & \text{if a warehouse located at } j \text{ can fully cover disaster area } A_i, \text{ where } i \in I \\ 0 & \text{otherwise.} \end{cases} \)

\( w_i \triangleq \) weight assigned to disaster area \( A_i \), where \( i \in I \).

Sets:

- \( I \) = set of disaster areas.
- \( J \) = set of candidate warehouse locations.
- \( N(i) \) = set of warehouse locations that can cover disaster area \( A_i \), where \( i \in I \).
- \( W(i) \) = set of candidate warehouse locations \( j \) partially covering disaster area \( A_i \) at least \( b \), but less than 100%, where \( i \in I \).

Decision variables:

- \( x_j \triangleq \begin{cases} 1 & \text{if a warehouse is located at location } j \in J, \\ 0 & \text{otherwise.} \end{cases} \)
- \( y_i \triangleq \begin{cases} 1 & \text{if disaster area } A_i \text{ is covered by at least one warehouse, where } i \in I \\ 0 & \text{otherwise.} \end{cases} \)
- \( v_i \triangleq \begin{cases} 1 & \text{if disaster area } A_i \text{ is partially covered at least } \theta \text{ times, where } i \in I \\ 0 & \text{otherwise.} \end{cases} \)

\( C \triangleq \) the percentage of the total weighted coverage achieved by the located warehouses.

Parameters \( b \) and \( \theta \) express the extend to which partial coverage by more than one warehouse may be considered as equivalent to full coverage. The values of \( b \) and \( \theta \) are normally such that \( b \cdot \theta = 100\% \).

5.2 Model Formulation

In the context of maximal coverage, a model that considers partial coverage may be formulated on the basis of spatial objects rather than single demand points [31]. Demand is represented as a union of areas \( A_i \) for \( i \in I \).

The complete model can be formulated as follows:

\[
\max \quad z = \sum_{i \in I} w_i y_i
\]
subject to

\[ \sum_{i \in I} a_{ij} x_j \geq y_i - v_i \quad \forall i \in I \]  

(2)

\[ \sum_{i \in I} x_j = S \]  

(3)

\[ \sum_{j \in W(i)} x_j \geq \theta \cdot y_i \quad \forall i \in I \]  

(4)

\[ C = 100 \left( \frac{\sum_{i \in I} W_i y_i}{\sum_{i \in I} W_i} \right) \]  

(5)

\[ x_j \in (0,1) \quad \forall j \in J \]

\[ y_i, v_i \in (0,1) \quad \forall i \in I \]

The objective function (1) considers all fully and partially covered disaster areas. Constraint (2) ensures that a fully covering facility is located to cover disaster area \( A_i \) when it is not partially covered \( (v_i = 0) \). When \( v_i = 1 \), \( y_i \) will be equal to 1, which means that area \( A_i \) is covered. Constraint (3) ensures that the specified number of warehouses are located, whereas constraint (4) imply that if \( v_i = 1 \), disaster area \( A_i \) is partially covered by at least \( \theta \) warehouses (Figure 3).

Coefficient \( w_i \) expresses the desirability, or benefit, of providing full coverage to disaster area \( A_i \). This coefficient is equal to the risk weight factor assigned to disaster area \( A_i \). Constraint (5) calculates the percentage of the total benefit achieved, through covering demand areas, by the specified number of located warehouses. This percentage is used to compare the cost of building a number of warehouses against the percentage coverage achieved. Finally, constraints (6) and (7) are included to ensure that the relevant decision variables are binary.

6. MODEL EXECUTION

The dataset [43] that Peduzzi [5] developed was used in GIS and layered with a grid of 100 km x 100 km cells that represents the disaster areas.

A total of 153 disaster areas and 21 potential locations for warehouses were identified, using the grid as reference, throughout the fifteen SADC countries. The potential locations for the warehouses were firstly selected so as to cover all disaster areas and then a number of locations were added to cover the more disaster prone areas. Figure 2 illustrates the disaster areas, warehouse locations and the coverage area of the warehouses. Examining Figure 2, location 15 in the Democratic Republic of the Congo, can be taken as an example of an area with a high risk to earthquakes. Madagascar on the other hand has a low to medium risk, that extend over the whole island, of cyclones.
Warehouse pre-positioning for disaster relief in Southern African development communities - Meyer, Bean, Yadavalli

The coverage provided by a pre-positioned warehouse is described by a circle. Figure 3 illustrates the fully and partially covered concept. A disaster area is also considered fully covered if it is partially covered by \( g = 2 \) warehouses, with the coverage \( b \geq 50\% \).

![Figure 2 - Illustration of disaster areas, warehouse locations, and warehouse coverage radius](image)

![Figure 3 - Full and partial coverage](image)

### 6.1 Assumptions

Some assumptions had to be made. They include:

- The positions of the pre-positioned warehouses are assumed to be at the centre of the 100 km x 100 km grid cell,
- the maximum distance a warehouse can be positioned from a disaster area is 400 km,
- the cost to build a warehouse is assumed to be R350,000,000, and
- a disaster in a region does not affect the warehouse(s) in that region.
6.2 Results

The model was executed 21 times in LINGO 8.0 optimisation software to determine the relative coverage that can be achieved by increasing the number of warehouses located. The percentage weighted coverage achieved for different numbers of warehouses located is depicted in Figure 4, while Table 1 shows the order in which the warehouses were located.

Table 1 - Order in which warehouse locations are chosen

<table>
<thead>
<tr>
<th>Number Located</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>X X</td>
</tr>
<tr>
<td>4</td>
<td>X X X X</td>
</tr>
<tr>
<td>5</td>
<td>X X X X</td>
</tr>
<tr>
<td>6</td>
<td>X X X X</td>
</tr>
<tr>
<td>7</td>
<td>X X X X</td>
</tr>
<tr>
<td>8</td>
<td>X X X X</td>
</tr>
<tr>
<td>9</td>
<td>X X X X</td>
</tr>
<tr>
<td>10</td>
<td>X X X X</td>
</tr>
<tr>
<td>11</td>
<td>X X X X</td>
</tr>
<tr>
<td>12</td>
<td>X X X X</td>
</tr>
<tr>
<td>13</td>
<td>X X X X</td>
</tr>
<tr>
<td>14</td>
<td>X X X X</td>
</tr>
<tr>
<td>15</td>
<td>X X X X</td>
</tr>
<tr>
<td>16</td>
<td>X X X X</td>
</tr>
<tr>
<td>17</td>
<td>X X X X</td>
</tr>
<tr>
<td>18</td>
<td>X X X X</td>
</tr>
<tr>
<td>19</td>
<td>X X X X</td>
</tr>
<tr>
<td>20</td>
<td>X X X X</td>
</tr>
<tr>
<td>21</td>
<td>X X X X</td>
</tr>
</tbody>
</table>

Figure 4 - Comparison between the percentage weighted coverage achieved and the number of warehouses located.
It is clear when comparing Table 1 and Figure 2 that the more disaster prone countries, like Madagascar, Mozambique and the northern parts of Tanzania and the Democratic Republic of the Congo are covered first. This is expected as the weights associated with these areas are much higher, therefore the ‘benefit’ achieved by covering those areas are much higher than the ‘benefit’ realised when covering less disaster prone areas, with lower weights.

When evaluating the cost associated with building a warehouse it can be seen, portrayed in Figure 5, that at some point the percentage weighted coverage achieved by locating another warehouse might not be reasonable when considering the cost of building that extra warehouse.

![Figure 5 - Comparison between the cost of building a warehouse and the percentage weighted coverage achieved](image)

**7. FUTURE WORK**

The MCLP with spatial objects model proves to be a good way of finding locations to pre-position emergency supply warehouses to aid in disaster relief. However, there are opportunities for improvement.

**7.1 Data Limitations**

“In God we trust; all others must bring data.”

W. Edwards Deming

Data limitations combined with the unpredictable and unique nature of hazards imply that much uncertainty remains. Given the limitations and uncertainties, the estimates of exposure and risk provided can only be taken as indicative [6]. Therefore, the data should only be used in identifying higher risk disaster areas.

It is of paramount importance to gather, maintain and manage credible data-records of disastrous events for an effective risk assessment and mitigation of disaster impacts [45]. If more effort is focused towards the collection of sub-national disaster data, it will help build datasets and indicators with a national level of observation and a local scale of resolution that can enable the visualisation of complex patterns of local risk [6]. This will help capture data on secondary disaster events and also highlight the ways in which natural, man-made and hybrid disasters interact, giving a clearer view of reality.
7.2 Warehouse Locations

Future work would include research on the specific locations of the warehouses. According to Whybark [46], the issues of distribution are complicated by political relationships between countries of the warehouse location and countries of the inventory destination. Whybark [46] found that often despite the need, the government in the country of the disaster refuses the disaster relief because of the source of the inventory that is part of the relief. This issue of acceptability is related to the decision of the pre-positioning of warehouses, locations must carefully be considered before being chosen.

7.3 Generic Application of Model

Although this paper only focus on SADC, this model can be used for other countries as well. The disaster data used to determine the risk factors can be used for any country in the world (Peduzzi and Deichmann's [5] map of the distribution of multiple hazards' mortality risk), therefore this model can be applied globally.

8. CONCLUSION

The number of people killed in disasters is estimated to be three to four times higher in developing countries [47] such as SADC countries. It is said that for every person killed, 3000 more are exposed to natural hazards [6], but in spite of this, very little has been invested in the prevention, mitigation and preparedness of disasters in SADC.

The aim of this paper was to find an appropriate way to pre-position emergency supply warehouses in SADC countries to help aid in disaster relief. This was achieved using a MCLP that includes spatial objects rather than single points, partial coverage, and weights assigned to disaster areas.

It was found that beyond a critical number of warehouses, the coverage achieved might not increase significantly enough to justify the cost associated with locating another warehouse.

Future work includes the collection of detailed disaster data on sub-national levels, and thorough research on the countries and specific locations for warehouses.

Although living with risk is the order of the day, we must learn to reduce these risks through appropriate measures to reduce contemporary risk with compensatory disaster risk management, such as disaster preparedness and response [6]. Let us do as the Global Assessment Report of 2009 [5] clearly stated, invest today for a safer tomorrow.

9. REFERENCES


Warehouse pre-positioning for disaster relief in Southern African development communities - Meyer, Bean, Yadavalli


EXPLORING WAYS TO REDUCE THE GREENHOUSE GAS EMISSIONS IN THE TEXTILE SUPPLY CHAIN

S. Nagel1*, C. Schoeman2, V.S.S. Yadavalli3
1Department of Industrial and Systems Engineering
University of Pretoria, South Africa
verstanienagel@gmail.com

2Built Environment Unit
CSIR, South Africa
cmschoeman@csir.co.za

3sarma.yadavalli@up.ac.za

ABSTRACT

With increasing awareness of global warming and the consequences of economic activities on the environment, enterprises are re-examining their business practises in search of eco-friendly alternatives. The paper explores ways to reduce greenhouse gas emissions of supply chains, incorporating suggestions from the GreenSCOR model and using the textile industry as a case study. Supply chain processes of organic T-Shirts at a major South African clothing retailer were analysed (including extraction of raw materials, manufacturing and distribution) and the carbon footprint determined (in kilograms carbon dioxide equivalent units). The impact of improvement suggestions were calculated in terms of reduced emissions.
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

1. INTRODUCTION

The term ‘Carbon Footprint” has become a buzz word in the action against global warming. According to Wiedmann & Minx [1], a carbon footprint is a measure of the total amount of greenhouse gas (GHG) emissions that is directly and indirectly accumulated over the life cycle of a product. With climate change high up on the political and corporate agenda, carbon footprint calculations are in strong demand.

A study in the UK found that more than 70% of customers would like more information regarding the carbon footprint of the products that they want to purchase, and 63% of these UK customers admit that they would switch to brands with smaller carbon footprints (McDermott) [2]. As the world is becoming increasingly carbon-constrained, businesses have to meet consumer need by taking a pro-active approach toward lowering their products’ carbon footprint in an attempt to mitigate climate change. Reducing carbon footprints is of utmost concern.

“Taking action on climate change delivers energy cost reductions, enhanced reputation, satisfied and motivated employees and stakeholders, the ability to proactively manage regulatory compliance, and competitive leadership through planning new business innovations.” (Carbon Trust) [3]

2. CONTEXT: ENVIRONMENTAL IMPACT OF SUPPLY CHAINS

2.1 Supply chains’ contribution to carbon footprinting

The following classifications of supply chains carbon footprinting will be considered, (Sundarakani et al.) [4]:

- **stationary sources**, i.e. emissions from raw material processing, product manufacturing and warehousing
- **non-stationary sources** which include carbon emissions from inbound and outbound logistical activities.

Both stationary and non-stationary activities of a supply chain release greenhouse gas emissions and thus contribute to the carbon footprint of a product. The green house gas emissions are measured in kilograms carbon dioxide equivalent units (kg CO₂e), which includes all the greenhouse gasses, namely carbon dioxide (CO₂), methane (CH₄), hydro fluorocarbons (HFCs), per fluorocarbons (PFCs) and sulphur hexafluoride (SF₆).

2.2 Carbon Footprint Calculation

Several methods, approaches and tools have been developed in order to calculate carbon emissions, e.g. the input-output technique (Wang & Watson) [5], the ecological footprint assessment (Wackernagel & Rees) [6], the Hybrid Life-Cycle-Methodology (Kejun et al.) [7] and the life-cycle assessment (LCA) method. The LCA method, systematically assesses the energy use across a products’ supply chain (Reich-Weiser & Dornfeld) [8]. Consumers want to know what the emissions are of the products they consume (Carbon Trust) [9]. Thus by doing such carbon calculations; organisations could increase customer satisfaction by publically disclosing the carbon footprint details of their products (Defra) [10]. By quantifying the carbon emissions of products, new opportunities to reduce these emissions at a low cost can be identified and developed in the design, manufacturing and supply of products (Pekala et al.) [11]. Section 3 and 4 discusses the method of carbon calculation employed for the case study and the reasons for its suitability.

Businesses are increasingly confronted with the implications of global warming and climate change. Organisations that actively partake in the fight against global warming derive numerous benefits from doing so. According to Srivastava [12], the advantages include:
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

- **Reduced Operating Costs**: increasing the efficiency of the supply chain can result in decreased resource employment and therefore lower operating costs.
- **Increased customer satisfaction**: by capitalizing on the important marketing and global communication and media opportunities, customer satisfaction increases as the companies’ climate awareness increases.
- **Regulation**: understanding the potential carbon related taxes, tariffs and other legislation.
- **Improved reputation**: distinguishing a companies’ offering and branding the company as a green leader, enhances the goodwill of the brand.

2.2.1 The Carbon Label

A carbon label is a public display of a product’s total GHG emissions throughout its life-cycle up until the stage where the consumer buys the product. There is a need to inform customers on the carbon footprint of the products they purchase, to enable them to choose brands that are less damaging to the planet and environment (Carbon Trust) [9]. The British grocery store, Tesco is currently busy determining the carbon emissions of all 70 000 of their products and attaching labels containing the carbon breakdowns of each product (Finch) [13]. These labels can be compared to food labels displaying the nutritional values and are attached to the product label. The value on the label will be given as CO$_2$e per product i.e. how many grams of GHG the product emitted during the raw material sourcing, production and distribution, thus enabling customers to make environmentally informed product choices.

2.3 Green Supply Chain Management (GrSCM)

Srivastara [12] defines green supply chain management (GrSCM) as:

> “integrating environment thinking into supply chain management, including product design, material sourcing and selection, manufacturing processes, delivery of the final product to the consumers, and end-of-life management of the product after its useful life.”

2.3.1 GreenSCOR

The Supply Chain Council’s SCOR model integrates best practises and metrics into the SCM process and places emphasis on each step in the supply chain of a product. An environmental approach has been applied to the SCOR structured model, resulting in a GreenSCOR model illustrated in Table 1. The GreenSCOR framework is an instrument for communicating and structuring green supply chain management (GrSCM) programs to get faster, repeatable, collaborative results (GreenSCOR, Supply Chain Council) [14].
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

Table 1 - SCOR process environmental impacts

<table>
<thead>
<tr>
<th>SCOR process</th>
<th>Potential Impact</th>
</tr>
</thead>
</table>
| Plan         | • Supply chain partners collaborate on environmental issues  
              | • Plans created to minimize energy consumption |
| Source       | • Select suppliers with positive environmental records  
              | • Select materials with environmentally friendly content  
              | • Minimize delivery distance |
| Make         | • Minimize energy consumption  
              | • Minimize packaging material |
| Deliver      | • Route to minimize fuel consumption  
              | • Retrieve packaging for re-use |
| Return       | • Minimize delivery distance |

The GreenSCOR model is used to identify opportunities for carbon emissions reduction of the T-shirt supply chain.

3. ANALYSIS OF THE T-SHIRT SUPPLY CHAIN

The carbon footprint calculation for a clothing company’s T-shirts reflects the stages of the T-shirt life-cycle from growing and extracting organic cotton in India up to the stage where the consumer purchases the product in South Africa. In calculating the carbon footprint of a textile product, it must be clear which set of emissions will be included in the boundary of the quantification. The retail, use and disposal of the T-shirts do not fall within the boundaries of the carbon calculation for this paper. An LCA will be taken for the calculation of the carbon footprint of a 100% organic T-Shirt. The method enables one to do an evaluation throughout the life cycle of the product and then focus on the prominent areas of environmental impact.

The process map of the T-Shirt supply chain is illustrated in Figure 1. The life-cycle stages of the paper include: extraction of raw materials, energy used in manufacturing of the T-Shirts, storage and transportation. The T-shirt consists of 100% organic cotton. The farms where organic cotton is grown are situated in Maharashtra state in India. The cotton is grown with non-genetically modified seed and no fertilizers or insecticides are used in the process. The seeds, leaves and other non-cotton particles are removed by a ginning process.

Once the cotton reaches the textile factory, the manufacturing process begins. The organic cotton is spun and transformed to yarn. The yarn is then steamed, knitted, dyed, cut and made to yield the final products. As the wastage of textile products is not constant and usually falls between 6-25%, the wastage for a T-shirt will be assumed as 6% as it is a simple garment to manufacture. It is assumed that waste is discarded. All of the manufacturing processes require high temperatures and lubricants. Chemicals are also needed for the dyeing process. Currently the energy source of the manufacturing factory is grid electricity. The carbon calculations of the lubricants and chemicals as well as energy used for infrastructure and buildings does not fall within the ambit of the research. The final products are then packed in boxes and into a container which is transported by truck to the Mumbai port of India. The containers are shipped to South African ports and are then transported by trucks to five distribution centres across South Africa. The T-Shirts are then distributed to clothing stores.
4. MODEL TO MEASURE CARBON EMISSIONS

4.1 System Specification

Carbon calculator models were designed to measure the carbon emissions throughout the supply chain of a T-Shirt. Carbon conversions were obtained from the Carbon Trust [15] as well as the UK Department for Environment, Food and Rural Affairs (Defra) [10]. Both these sources provide accessible online guides as to how organisations can measure the carbon footprint of their products. The carbon calculator models which were designed for the purpose of this paper focus on calculating the carbon emissions of a single product, a 100% organic T-Shirt. See Table 2 below. The information for the carbon emissions calculation was gathered from the various suppliers and contributors in the supply chain. The greenhouse gas emissions are calculated to the extent of the attainability of data. Data on the energy emissions from harvesting and extracting of organic cotton is unavailable and was extrapolated from other recorded data sources. (The basis for this assumption is discussed in more detail in section 4.2.1) A high quality T-Shirt weighs approximately 292g. For this assessment an assumption will be made that an organic T-Shirt weighs 250g, i.e. 4000 T-Shirts per tonne. The emission calculations in the model are done per tonne of T-shirts (measured in kg CO$_2$e), which is then divided by 4000 to result in the carbon footprint per T-Shirt.

4.2 Carbon Emissions in the T-Shirt Supply Chain

Table 2 summarises the carbon emissions of each stage in the supply chain and results in a total of 21,922.8 kg CO$_2$e per tonne T-shirts, which converts to 5.455 kg CO$_2$e per T-Shirt (21,922.8 / 4000).
Table 2 - CO$_2$e-emissions sheet of a T-Shirt

<table>
<thead>
<tr>
<th>Process</th>
<th>Kg CO$_2$e per tonne T-Shirts</th>
<th>Kg CO$_2$e per T-shirt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Materials</td>
<td>2300.00</td>
<td>0.575</td>
</tr>
<tr>
<td>Farming and Ginning</td>
<td>2300.00</td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>19080.40</td>
<td>4.770</td>
</tr>
<tr>
<td>T-shirt manufacturing</td>
<td>19080.40</td>
<td></td>
</tr>
<tr>
<td>Storage and Transportation</td>
<td>244.50</td>
<td>0.110</td>
</tr>
<tr>
<td>Storage</td>
<td>59.61</td>
<td></td>
</tr>
<tr>
<td>Transportation of Cotton</td>
<td>131.11</td>
<td></td>
</tr>
<tr>
<td>T-shirt Transport to Port of India</td>
<td>46.22</td>
<td></td>
</tr>
<tr>
<td>T-shirt Shipment to South Africa</td>
<td>132.33</td>
<td></td>
</tr>
<tr>
<td>Transport to Distribution Centre</td>
<td>24.58</td>
<td></td>
</tr>
<tr>
<td>Transport to Store</td>
<td>49.17</td>
<td></td>
</tr>
<tr>
<td>Total Emissions</td>
<td>21922.80</td>
<td>5.455</td>
</tr>
</tbody>
</table>

A carbon emission of approximately 5.455 kg CO$_2$e is equal to (Berners-Lee, M) [16]:

- Driving 35 km in a medium sized car
- Taking 6 showers of 11 minutes each in a typical electric geyser shower
- Flying 30 km in an aeroplane

The processes contributing to the total emission figure is discussed in greater detail below.

4.2.1 Raw Materials

According to Ferrigno et al. [17], organic cotton represents only 0.76% of the global cotton production. A large number of producers employ a variety of production factors (farming approach, cultivation method and irrigation) in the cultivation and ginning of organic cotton used to produce a T-Shirt in Maharashtra, India. In this supply chain, the data required to calculate the farming of raw materials' carbon emission contribution is unobtainable. Therefore, a figure of 2300 kg CO$_2$e per tonne of organic cotton is used in the paper – as is advised by the 77th International Wool Textile Organisation Congress held in Beijing (Moris) [18]. Thus the carbon emissions for the organic cotton extraction and ginning is equal to 0.575 kg CO$_2$e (i.e. 2300 / 4000).

4.2.2 Manufacturing

The carbon emissions model for the manufacturing part of the supply chain is illustrated in Table 3. The energy types used for manufacturing include:

- fossil fuelled power used for machinery (in kilo watt hours (kWh)),
- diesel for standby Gensets,
- petrol for vehicles and
- energy used in the production of packaging materials.

The energy values per tonne of T-shirts are converted by standard conversion factors supplied by the Carbon Trust [15] to yield the emissions kg CO$_2$e per tonne of T-shirts. The total T-Shirt Manufacturing emissions added up to 19 080.4 kg CO$_2$e per tonne T-shirts, i.e. 4.77 kg CO$_2$e per T-shirts (19 080.4 / 4000).
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

Table 3 - T-shirt manufacturing Emissions

<table>
<thead>
<tr>
<th>Manufacturing Process</th>
<th>Energy Type</th>
<th>Value per tonne T-shirts</th>
<th>Emission factor per unit (kg CO₂e)</th>
<th>Emissions per tonne T-shirts (kg CO₂e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinning</td>
<td>Fossil Fuelled Power</td>
<td>5140 kWh</td>
<td>0.54</td>
<td>2775.6</td>
</tr>
<tr>
<td></td>
<td>Diesel</td>
<td>10 litre</td>
<td>0.27</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>Packaging materials</td>
<td>10 kg</td>
<td>0.87</td>
<td>8.7</td>
</tr>
<tr>
<td>Knitting</td>
<td>Fossil Fuelled Power</td>
<td>470 kWh</td>
<td>0.54</td>
<td>253.8</td>
</tr>
<tr>
<td>Dyeing</td>
<td>Fossil Fuelled Power</td>
<td>29 000 kWh</td>
<td>0.54</td>
<td>15660.0</td>
</tr>
<tr>
<td></td>
<td>Diesel</td>
<td>10 litre</td>
<td>0.27</td>
<td>2.7</td>
</tr>
<tr>
<td></td>
<td>Petrol</td>
<td>10 litre</td>
<td>0.25</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Packaging materials</td>
<td>80 kg</td>
<td>0.87</td>
<td>69.6</td>
</tr>
<tr>
<td>Cutting &amp; Sewing</td>
<td>Fossil Fuelled Power</td>
<td>500 kWh</td>
<td>0.54</td>
<td>270.0</td>
</tr>
<tr>
<td></td>
<td>Packaging materials</td>
<td>40 kg</td>
<td>0.87</td>
<td>34.8</td>
</tr>
<tr>
<td>Totals:</td>
<td></td>
<td></td>
<td></td>
<td>19080.4</td>
</tr>
</tbody>
</table>

Thus, from Table 3 the total emission contribution per energy type is calculated to be as follows:

- Fossil Fuelled Power = 18 959.4 kg CO₂e per tonne T-shirts or 4.74 kg CO₂e per T-shirt (18 959.4 / 4000) or
- Diesel = 5.4 kg CO₂e per tonne T-shirts
- Petrol = 2.5 kg CO₂e per tonne T-shirts
- Packaging materials = 104.4 kg CO₂e per tonne T-shirts or 0.026 kg CO₂e per T-shirt (104.4 / 4000)

4.2.3 Storage

To determine the emissions emitted by storing a tonne of T-shirts, a basic calculation is done. Parameters included in the model formulation are as follows:

e = site energy consumed (kWh p.a)

a = site energy allocated per tonne T-shirts (%)

fₑ = emissions factor for electricity used in store

Tₑ = total emissions from storage (kg CO₂e per tonne T-shirts)

Equation 1  

\[ Tₑ = eafₑ \]

\[ = (10 000 000)(0.001\%)(0.6) \]

\[ = 59.61 \text{ kg CO₂e per tonne T-shirts} \]

Site studies proved that an average of 10 000 000 kWh of energy is used on site per annum, of which 0.001% is allocated to a tonne of T-shirts. The emissions factor is 0.6 kg per kWh according to the (Carbon Trust) [15]. The calculation shown in Equation 1 yields a total emission of 59.61 kg CO₂e per tonne of T-shirts in storage.

4.2.4 Transportation

A model (Table 4) was designed to find the kg CO₂e per tonne T-Shirts transported from India to South Africa. The table below lists the distances, modes of transport as well as the resulting kg CO₂e per tonne of T-shirts. A total distance of 10 684 km is covered in the supply chain and the total carbon emissions for the transportation sections of the supply chain add up to 383.41 kg CO₂e per tonne T-shirts i.e. 0.095 kg CO₂e per T-shirt. For the purpose of the paper two assumptions were made for transportation throughout the supply chain.
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

Chain: The mass loaded on trucks and ships are 100% en route to South Africa and 50% on return.

Table 4 - Transportation data of T-shirts from India to South Africa

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Mode of Transport</th>
<th>Distance (km)</th>
<th>Emissions per tonne T-shirts (Kg CO(_2)e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton plantation (India)</td>
<td>Manufacturing (India)</td>
<td>Truck</td>
<td>1600</td>
<td>131.11</td>
</tr>
<tr>
<td>Manufacturing (India)</td>
<td>Port of India</td>
<td>Truck</td>
<td>400</td>
<td>46.22</td>
</tr>
<tr>
<td>Port of India</td>
<td>Port of South Africa</td>
<td>Ship</td>
<td>7784</td>
<td>132.33</td>
</tr>
<tr>
<td>Port of South Africa</td>
<td>Warehouse</td>
<td>Truck</td>
<td>600</td>
<td>49.17</td>
</tr>
<tr>
<td>Warehouse</td>
<td>Store</td>
<td>Truck</td>
<td>300</td>
<td>24.58</td>
</tr>
<tr>
<td>Totals:</td>
<td></td>
<td></td>
<td>10684</td>
<td>383.41</td>
</tr>
</tbody>
</table>

The carbon emission values for the road transport leg are determined by entering the distances into a basic model with the following parameters:

\[
T_s = \frac{\sum (Df_{50\%}p + Df_{100\%}p_r)}{t}
\]

\[
= \frac{((1600)(1.04)(100%)+(1600)(0.87)(50%))}{18}
\]

\[
= 131.11 \text{ kg CO}_2\text{e}
\]

The distances are converted by the carbon conversion factors supplied by Defra [10] to determine the emissions per tonne of T-shirts. An articulated 18-22 ton truck is used for all road transport. An assumption is made that the trucks are filled with 18 tons of weight. To determine the carbon emissions/tonne of T-shirts, the total kg CO\(_2\)e of the truck is divided by 18 to yield the total CO\(_2\)e per tonne of cotton. The calculation in Equation 2 illustrates how the carbon emissions per tonne of cotton transported is calculated to be 131.11 kg CO\(_2\)e. The carbon emissions for all truck transport were calculated in this manner by substituting the parameters into Equation 2. The results are listed in Table 4.

The amount of carbon emitted for the sea-leg shipping of a tonne of T-shirts works on the same principle as the road transport emissions. The parameters are listed below and the equation is illustrated in Equation 3. According to (Defra) [10], a small container vessel emits 0.012 kg CO\(_2\)e per kilometre if it is fully loaded and 0.01 kg CO\(_2\)e if it is 50% full. Thus, travelling a distance of 7784 km from India to South Africa emits 132.33 kg CO\(_2\)e per tonne of T-shirts.

\[
T_s = \frac{\sum (Df_{50\%} + Df_{100\%}p_r)}{t}
\]

\[
= \frac{((1600)(1.04))(100%)+(1600)(0.87)(50%))}{18}
\]

\[
= 132.33 \text{ kg CO}_2\text{e}
\]
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

Equation 3 \[ T_s = \sum(D_f s_{50\%}P + D_f s_{100\%}P) \]
\[ = ((7784)(0.012)(100\%)+(7784)(0.01)(50\%)) \]
\[ = 132.33 \text{ kg CO}_2\text{e per tonne T-shirts} \]

Thus from Table 4 it can be calculated that the total emission contribution per transport mode type is as follows:

- Road = 251.8 kg CO$_2$e per tonne T-shirts or 0.063 kg CO$_2$e per T-shirt (251.8 / 4000), and
- Sea = 132.33 kg CO$_2$e per tonne T-shirts or 0.033 kg CO$_2$e per T-shirt (132.33 / 4000)

4.2.5 Summary of Current Carbon Footprint

The breakdown of carbon emissions for 100% Organic T-shirts, according to the values in Table 2, is summarized in Figure 2.

![Figure 2 - Kilogram CO$_2$e per ton T-Shirts](image)

It is clear that the largest contributor to carbon emissions in the supply chain is the manufacturing process and it therefore presents the greatest opportunity for efforts to reduce the carbon footprint.

5. REDUCING EMISSIONS IN THE SUPPLY CHAIN

Once the carbon footprint is calculated, several actions may be taken. Since the objective of this paper is to reduce carbon emissions across the supply chain, the next step is to identify carbon reduction opportunities throughout the T-shirt supply chain. This is done according to the GreenSCOR model since it indicates how the processes in the supply chain may be improved in an environmentally sound way.

5.1 Raw Material

The plan and source SCOR processes were considered in the harvesting and extraction of raw materials. The most important factors regarding cotton farming are the farming approach, cultivation method and irrigation.
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

5.1.1 Farming approach

The main greenhouse gas emissions emitted from growing and harvesting cotton are due to insecticide, pesticide, fertilizers and other chemicals used in the farming process. According to Pollheimer [19], conventionally farmed cotton emits between two and four kilograms CO$_2$e per T-shirt. For organic cotton, this figure is reduced to 0.575 kg CO$_2$e per t-shirt.

5.1.2 Cultivation Method

Organic farms make use of manual labour, which removes the need for heavy machinery that has high carbon emissions.

5.1.3 Irrigation

Organic cotton farms do not make use of artificial irrigation methods. The crops are rain-fed, which eliminates the emissions from high-energy irrigation pumps.

5.2 Manufacturing

Figure 2 illustrated that the manufacturing process is the most significant contributor to the carbon emissions of the T-shirt supply chain - at 19080.4 kg CO$_2$e emissions per tonne of T-shirts (see Table 3) - it is responsible for 88% of the carbon footprint.

The make process in the SCOR model suggests that the energy consumption should be minimised in manufacturing. GreenSCOR [14] The carbon footprint of manufacturing processes could be reduced by making use of low carbon technologies. These could include alternative and renewable energy sources such as biomass, photovoltaics (PV), wind, nuclear and hydro energy sources. Table 5 lists the emission factors for various energy sources, as well as the emissions per tonne of T-shirts for each of these. Postnote [20]

Table 5 - Emissions from Various Energy Systems

<table>
<thead>
<tr>
<th>Energy Source</th>
<th>Energy usage per tonne T-shirts</th>
<th>Emissions factor (kg CO$_2$e/kWh)</th>
<th>Emissions per tonne T-shirts (kg CO$_2$e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fossil Fuelled Technologies</td>
<td>35110 kWh</td>
<td>0.54000</td>
<td>18959.40</td>
</tr>
<tr>
<td>Biomass</td>
<td>35110 kWh</td>
<td>0.09300</td>
<td>3265.23</td>
</tr>
<tr>
<td>Photovoltaics</td>
<td>35110 kWh</td>
<td>0.03500</td>
<td>1228.85</td>
</tr>
<tr>
<td>Marine Technologies</td>
<td>35110 kWh 0.02500-0.05000</td>
<td>877.75-1755.50</td>
<td></td>
</tr>
<tr>
<td>Hydro</td>
<td>35110 kWh</td>
<td>0.00500</td>
<td>175.55</td>
</tr>
<tr>
<td>Wind</td>
<td>35110 kWh</td>
<td>0.00464</td>
<td>162.91</td>
</tr>
<tr>
<td>Nuclear</td>
<td>35110 kWh</td>
<td>0.00500</td>
<td>175.55</td>
</tr>
</tbody>
</table>

The current main source of energy in the supply chain is grid electricity (generated through fossil fuel burning), which has the largest carbon emission factor of all electricity generation systems (at 0.54 kg CO$_2$e/kWh as can be seen in Table 5). Within the manufacturing process, the fossil fuel power’s contribution to emissions was determined from Table 3, to add up to 18 959.4 kg CO$_2$e per tonne T-shirts.

In this paper, an emissions factor of 0.005 kg CO$_2$e will be used for renewable energy. If the grid electricity energy component of the manufacturing process is replaced by renewable energy the consequent emissions per tonne of T-shirts will be 175.55 kg CO$_2$e per tonne T-shirts (as is calculated in Table 5). The revised manufacturing carbon footprint is calculated by adding the emission contributions of diesel (5.4), petrol (2.5) and packaging (104.4) - as
Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli

derived from Table 3 - and amounts to 296.55 kg CO₂e / tonne of T-shirts. This translates to 0.074 kg CO₂e per T-shirt (296.55 / 4000) compared to the original manufacturing carbon footprint of 4.77 kg CO₂e per T-shirt - see Figure 3. Therefore the manufacturing component of the supply chain’s carbon footprint can be reduced by 18 783.85 kg CO₂e per tonne T-shirts by replacing coal generated electricity with greener energy sources.

![Graph showing carbon footprint of grid electricity vs. renewable energy]

**Figure 3 - Carbon footprint of grid electricity vs. renewable energy**

5.3 Transportation

The transportation of goods is considered in the deliver and return SCOR processes. In transportation there are few carbon emissions reduction opportunities. Road transport is responsible for the largest part of carbon emissions in the transport segment. Four variables that should be considered when looking at the greenhouse gas emissions from transport (Environmental Audit Committee) [21] are:

- the carbon content of the fuel used,
- fuel efficiency,
- distances travelled, and
- transport mode

Diesel trucks are the most efficient means for road transportation. Since present battery technology is not yet suitable for electrically powered trucks, drivers could be trained to drive in an energy efficient manner (Board of Academic Advisors to the Federal Minister of Transport, Building and Urban Affairs) [22].

Current carbon emissions from transportation equal 0.0958 kg CO₂e per T-shirt (calculated from Table 4 as follows: 383.41 / 4000). If a local supplier is used, the transportation distance would be reduced significantly. For instance, a South African cotton producer could be used at a distance of 100 km from the manufacturing plant - which in turn is 100 km from the Distribution Centre, which again is located 400 km from the store. The carbon emissions associated with the suggested local transport configuration is calculated using Equation 2 and the results are displayed in Table 6.
Exploring ways to reduce the greenhouse gas emissions in the 
Textile Supply Chain - Nagel, Schoeman, Yadavalli

Table 6 - Revised Transportation Emissions

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Mode of Transport</th>
<th>Distance (km)</th>
<th>Kg CO₂e per tonne T-shirts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cotton plantation</td>
<td>Manufacturing</td>
<td>Truck</td>
<td>100</td>
<td>8.19</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>Distribution Centre</td>
<td>Truck</td>
<td>100</td>
<td>8.19</td>
</tr>
<tr>
<td>Distribution Centre</td>
<td>Store</td>
<td>Truck</td>
<td>400</td>
<td>46.22</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>62.60</td>
</tr>
</tbody>
</table>

The improved transport carbon emissions equals 0.015 kg CO₂e per T-shirt (calculated from Table 5 as follows: 62.6 / 4000). This results in a reduction of 84.34 % as can be seen in Figure 4.

Figure 4 - Carbon reduction in transport of switching to a South African cotton producer

5.4 Summary of Reduced Carbon Footprint

Table 7 is a revised version of Table 2, with new figures based on all the improvement suggestions - to illustrate the reduced emissions of each component of the supply chain.

Table 7 - Revised CO₂e-emissions sheet of a T-Shirt

<table>
<thead>
<tr>
<th></th>
<th>Current Kg CO₂e per T-shirt</th>
<th>Revised Kg CO₂e per T-shirt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw Materials</td>
<td>0.575</td>
<td>0.575</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>4.770</td>
<td>0.074</td>
</tr>
<tr>
<td>Storage and Transportation</td>
<td>0.110</td>
<td>0.015</td>
</tr>
<tr>
<td>Total Emissions</td>
<td>5.455</td>
<td>0.664</td>
</tr>
</tbody>
</table>
6. RECOMMENDATIONS

Section 1 asserts that customers’ attitudes are changing towards global warming and climate change. By implementing the carbon label, companies will have to start developing strategies to reduce their carbon emissions.

The LCA was used to determine the emissions throughout the supply chain for organic T-shirts at a major South African clothing retailer. Considering emissions from raw material extraction, manufacturing processes, storage and distribution, new opportunities could be identified to lower the emissions by using a GreenSCOR model. Seeing that raw materials extraction, storage and distribution of T-shirts have a low carbon footprint, it seems clear that in order for this supply chain to reduce the carbon emissions, two key opportunities exist:

- moving to a T-shirt supplier that makes use of renewable energy, or
- urging the current manufacturing company to move to low-energy sources.

These recommendations would drastically reduce the carbon footprint of T-shirts in the supply chain as illustrated in Figure 3.

Additionally, by shortening road travel distances and making use of energy efficient transport methods, the carbon footprint can be further reduced.

Consequently by switching to a local supplier and producer of cotton, that makes use of renewable green electricity the carbon footprint of a T-shirt could be reduced by 87.82%, illustrated in Figure 5.

![Figure 5: Carbon footprint of the current supply chain vs. the reduced footprint](image)

7. CONCLUSION

The paper presented an analysis of the carbon footprint of 100% organic cotton T-shirts using a life-cycle approach. Suppliers and other contributors in the supply chain were engaged to provide the necessary data for the carbon calculations. The unavailable data were gleaned from journals, reports and other data sources. Emissions from raw materials, manufacturing processes, storage and distribution were considered and resulted in a total carbon footprint of 5.455 kg CO$_2$e per T-shirt.
Opportunities to reduce the carbon footprint were identified by making use of the GreenSCOR model. Recommendations were made and it was concluded that energy efficiency is the key to a lowered carbon footprint. Thus, switching to a manufacturer that makes use of renewable energy sources could diminish this component of the carbon emissions by 98.44%. By implementing energy efficient ways of transport, the carbon footprint of the transport segment could be lowered by 84.346%. By implementing both transport and manufacturing suggestions, the carbon footprint of T-shirts could be reduced from 5.455 kg CO$_2$e to 0.664 kg CO$_2$e.

From the study, it can be concluded that by determining the carbon footprint across the supply chain and identifying the processes that are the main sources of carbon emissions, new opportunities could be discovered to lower the carbon footprint of products. By attaching a carbon label to products, customers can make informed decisions regarding which brand they would support according to their environmental conscience.

8. REFERENCES

Exploring ways to reduce the greenhouse gas emissions in the Textile Supply Chain - Nagel, Schoeman, Yadavalli


DEPLOYING LEAN THINKING TO IMPROVE GRADUATE QUALITY AND SERVICE DELIVERY IN HIGHER EDUCATION

Rodger L. Nkumbwa
Copperbelt University, School of Engineering
Department of Mechanical Engineering
Email: Rodgers.nkumbwa@cbu.ac.zm

ABSTRACT

This paper presents the results of the investigation into how the higher education Service Delivery and Graduate Quality can be improved, using the principles of ‘Lean Thinking’, which were previously dominant in engineering firms like Toyota. The feasibility and rewards of deploying Lean Thinking in a higher education set-up are explored. Lean is a way of thinking in terms of delivering value to the customer, improving institution business processes through waste operations identification and elimination. In addition people: both customers and employees, are being respected. The business of higher education today faces greater competition owing to rapid changes. The focus now is on organization, bottom line, and also has a competitive edge for its clients, including students, sponsors, staff, donors and government agencies. Higher education quality and education product are defined. Current higher education calls for greater accountability, which implies establishing quality improvement strategies to give the required resource leverage. Strategic quality improvement programs such as lean and six-sigma can improve the performance of higher education institutions. The quality in higher education may be perceived to be political with underlying economical, social and technological implications. Times are continuously changing which leaves all institutions with no option but to evolve if they are to sustain their market share and remain relevant. The results of the study are that, yes, lean strategy can be applied in higher education environment to achieve higher graduate quality, organizational excellence and improved bottom line. All this with no massive monetary investment required, but only the ability to accept change and adapt. The goal of this study is to edify and convey quality management insights to higher education stakeholders. The scope of this paper is restricted to lean deployment in a higher education environment. Research methods used, include: interviews with experts, probing case studies, author reflection and experience as well as literature content review.
1. INTRODUCTION

Governments in general, and the higher education sector in particular, face a productivity imperative. Growth in new national priorities and citizens’ demand for improved higher education service delivery requires government to do more and to do it better. The need to increase productivity in the higher education sector is obvious to, education stakeholders as well as critical. If higher education delivery is to improve, valuable insights could be gained by looking at what is driving the corporate sector and global world class organizations, namely innovative management practices such as Quality Management, Lean and Six-Sigma Strategies. Today’s policymakers and higher education institution managers need to pursue a campaign for productivity improvement so as to leverage the available resources to achieve sustainable development. Performance transformation and transparency should be a priority in all government agencies, with education being no exception. However, the main problem is that we have very highly qualified educators working on broken institution processes, which add less value to the customers’ (students and employers) education experience.

While significant progress has been made in the quest to make higher education better, the motivation for the need to change the current model in higher education today, lies in the significant number of challenges that need to be addressed, including: the education financial deficit in most countries, capacity constraints, political concerns about operational costs, less motivated workforce, staff shortages and low quality of graduates. However, it is possible today to improve graduate quality, to make working lives less stressful and more rewarding for staff and boost efficiency and productivity, thereby pleasing politicians and taxpayers, all at the same time. It is also possible to do all these things without painful restructuring, cash injection or massive new investments in infrastructure, simply by applying the principles of Lean Thinking to the higher education service delivery processes. Examples of Lean deployment in higher education have been recorded at various institutions around the world, including; Cardiff University, Coventry University, Park University, Michigan State University, University of New Orleans, University of Iowa, Colorado Public Schools and Malaysian Schools Reed and Berry [2], Inozu et al. [3], Cermak [4].

2. BACKGROUND

In today’s globalised competitive market environment, in which the needs of the customers are constantly and rapidly changing, with the choices increasing and the need of higher quality graduates and services rising, strategic higher learning organizations can no longer afford to lose out on profit opportunities, having less customer satisfaction and delayed graduates. The ever rising cost of teaching materials, equipment and labour are other major pressing issues for various Universities and Colleges around the world. The driving force for improved graduate quality and service delivery is the regional and international competition pressure. Many organizations are striving to meet these challenges by adopting the new management philosophies to deliver greater productivity, less process waste and higher education value delivery to the stakeholders. Gish [5] suggests that,

"By applying Lean principles, routine business operations could be simplified, more rational procedures established, and repetition reduced (if not eliminated), thereby accelerating core business processes and responding more quickly to customer needs."

As competition gets tougher and the aftermath of the economic downturn, lean, agility, green and six-sigma techniques emerge to be the solutions and performance strategy for most innovative higher learning institutions. These techniques look at the entire organization value chain in a holistic way rather than hierarchical, that is including suppliers, customers (students and employers), stakeholders, curriculum development, information systems, finance, all employees and the leadership to drive the change. It ought to be understood that lean is a systematic approach for organization process waste
Deploying lean thinking to improve graduate quality and service delivery in higher education - Nkumbwa

identification and elimination, value addition from the customer perspective, understanding and removal of all non-value adding operational activities in a quest to satisfy the customers and other stakeholders.

3. METHODOLOGY

Research methods used in this paper include: literature content review being the main research method where distinguished scholarly works and sources provided insights into lean thinking deployment in higher education system, as well as interviews with other lean thinking experts and probing case studies as well as the author’s reflection and experience. In addition to these methods, there is an ongoing trial implementation project headed by the author at the Copperbelt University under the School of Technology.

4. EXPLORING LEAN IMPRESSION FOR HIGHER EDUCATION IMPROVEMENT

Lean is a way of thinking that was introduced by Toyota in the 1960’s as a systematic approach for identifying and eliminating institution process waste or non-value-adding operations through continuous improvement with the goal of creating value for all the stakeholders. However, initially it was not referred to as “Lean”, but as the Toyota Production System (TPS). John Krafick, then a graduate student at Massachusetts Institute of Technology (MIT), Sloan School of Management and a researcher, was the first to use the term “Lean Production System” in his research, where he highlighted the fact that lean business uses less of everything, compared to traditional production or operations system. He further outlined the fact that lean operations do the following: they optimize human resources effort, institution space is optimized leading to savings, the need for capital investment funding is minimized, they also decrease the amount of supplies consumed and uses less time to produce and deliver the required services or products Murman et al. [6]. The key principles of “lean” are based on identifying “wastes” from the customer perspective and eliminating the wastes.

Waste is explained as any activity that a customer would not want to pay for and that which does not add value from the customer perspective. It should also be noted that customers can be both internal and external, with internal customers those that are present in the next step in the value creation process in the organization. It can therefore be stated that lean focuses on eliminating waste in organization processes. Lean is not about eliminating people, but considering them as team associates and educating them as value creators for the stakeholders. Lean is about expanding organization capacity by reducing costs and shortening production or service delivery cycle times. It is also about understanding what is important to the customer experience and delivering it on the right time in the right amount all the time. Balzer [7] outlines that,

“Skeptical audiences within higher education, however, may prefer or even demand direct evidence that Lean Higher Education principles and practices can be successfully applied at colleges and universities before pursuing and embracing their implementation.”

To illustrate the magnitude of the benefits that may come with the implementation of lean thinking in a higher education system, the following case studies as summarized by Balzer [7] are worth noting:

“University of Central Oklahoma: 16,000 students and a budget cut of 15% in 2002. They utilized Value Stream Mapping, Kaizens on Process Work

---

1The MIT Sloan School of Management, based in Cambridge, Massachusetts, is one of the World’s leading Business Schools - conducting cutting-edge research and providing management education to top students from more than 60 countries. The School is part of MIT’s rich intellectual tradition of education and research. http://mitsloan.mit.edu/
Orders to reduce backlog from 3000 work orders to 300, cut lead time by 89%, number of touches by 82%, thereby saving $14,000 dollars on just one project. They identified 13 additional projects going forward.”

“University of New Orleans: 11,000 students. They targeted the “Personnel Action Form 101” reducing process steps by 85%, errors by 89%, and reduced cycle time by 99%.”

“Bowling Green State University: Working on their Counselling Center they reduced waiting time for an appointment from 5 days to less than 1 day or an 80% improvement, and waiting for a doctor appointment from 16 days to less than 1 day or a 94% improvement. The students requesting services increased from 761 to 950 an increase of 25%, which represents a much more effective service to the ultimate “customers.”

5. QUALITY STRATEGY FOR HIGHER EDUCATION BUSINESS

Higher education institutions, nowadays, operate in a competitive environment due to rapid changes in the past decades. The focus is now on the bottom line and the need to get a competitive edge for all their customers throughout the value chain. That is, from students and their sponsors, staff, donors and government agencies. Current higher education business models call for a greater accountability, which implies establishing quality improvement strategies across the board to provide leverage both for the business and academic side. A strategic quality improvement program can improve the performance of all areas of higher education institutions. It is now an opportune moment for institutions of higher education to demonstrate that they can offer what others can not. Rozsnyai [8] states that:

“...the quality of the University is similarly to love: intangible, but existent; perceptible, but not quantifiable, ephemeral, so that one has to endeavour to it (Muller-Boling, D., quoted Kappler et al, 2000, p180).”

The quality in higher education may be perceived to be political with underlying economical, social and technological implications. The quest for quality may be attributed to: an increase in regional and international competition, changing market forces, dissatisfaction from clients such as students and their sponsors and increased demand for accountability from higher institution stakeholders like government agencies. Expansion projects with limited funding and changing context such as internationalisation of higher education is another factor. Most higher learning institutions, today, are struggling to deliver valuable services, while on the other hand maintaining a better financial position. To balance this, the current trend is to pass costs along to customers. The customers here are the students and their sponsors, funding bodies, agencies and the government. This has resulted in an increased cost burden on the sponsors. However, with the rate at which markets are changing, due to globalisation and changes in legislation, this opportunity for passing the cost to the customer will not go on forever.

Challenges will soon be emerging that will force change in higher education institutions. Emilian [9] suggests that, amongst the challenges to be faced by higher education institutions, administrators will, in the next few years, include: an oversupply of capable higher education institutions, study programs (degrees) that are not differentiated from the competition, growth of for-profit education institutions and competition based on the pricing. In addition, education standards will become standardised around the globe because of the ever rising quality assurance and monitoring bodies, within regions as well as internationally. The above challenges could be addressed by adapting, as well as the adoption and effective application of such philosophies as Six-Sigma and Lean Thinking Management. This will help to reposition the institutions by reducing the operational costs, improving the quality of service delivery and graduates, simplify the organisation business processes, gain the leverage for increased market share, give employment stability and
Deploying lean thinking to improve graduate quality and service delivery in higher education - Nkumbwa

provide better customer satisfaction. Emilian [9] further suggests that people should overcome the bias that Lean Thinking is a ‘manufacturing thing’ only and that,

“Higher institutions administrators, faculty and staff must avoid the trap of viewing higher education as a special case where Total Quality Management, ISO Standards, Six-Sigma and Lean Thinking does not apply.”

According to Raifsnider [10] areas noted for enormous wastes in higher institutions include; admission processes, document management, individual courses, degree programs curriculum, student services and library services. Quality Management may perhaps only be achieved if value is delivered to the customers at minimal operational cost. Let’s now define the education product, customers for education and their expectations.

6. DEFINING QUALITY IN HIGHER EDUCATION

Quality is a relative subject related to the extent of the business processes’ results and the desired outcomes. It may be stated as, ‘exceptional or excellent’, ‘perfection or consistence’ (zero defect), ‘fitness for purpose’, ‘value for money’ and ‘transformational’. The focus of education quality should be on what students have learnt, what they know, what they can do and what their attitudes are, as a consequence of their interactions with their teachers, departments and higher education institutions. Higher education quality must be about scholarship and learning. Quality in higher education may be said to be multi-dimensional, multi-faceted, embracing three aspects, namely: goals, processes and the level of goal achievement. Recent trends in quality management have shown slowly increasing adoption of Quality Management, Lean and Six-Sigma concepts in higher education institutions around the world, for improved performance and service delivery. Many agitate for the application of the Plan Do Check Act (PDCA), Deming’s concepts of quality management that would provide a guiding principle for the overdue educational reforms. Experience and studies reveal that many Universities provide a number of Quality Management Courses or Training to their customers (students and enterprises) while they have not applied the same methodologies and techniques to improve the performance of their systems and quality of their service delivery. Quality management philosophies are applied to business processes to improve their performance, which can be in manufacturing processes for tangible goods or non manufacturing for non-tangible products such as services. Examples include banks, insurance companies, healthcare and education. The application of quality management methodologies and techniques in higher education institutions of learning is slowly getting roots, as the competition to scramble for the same high quality students gets tougher. Among the methodologies being adopted recently are included; Lean Thinking Management, Total Quality Management and Six-Sigma Operations. The adoption of the techniques has already shown great results for the early adopter, as well as in similar services like healthcare and financial institutions. Emilian [9], President of the Centre for Lean Business, in his article entitled “Lean in Higher Education”, states that:

“...the time is right for higher education administrators, faculty and staff to begin applying Lean Management to their business. The consequences of not doing so could be fatal.”
7. THE DIFFERENCE BETWEEN A TRADITIONAL EDUCATION SYSTEM AND A LEAN EDUCATION SYSTEM

Traditionally, organization operations have been driven by naïve sales or a demand forecast. Similarly, in the education system, Universities and Colleges would produce graduates in specialized areas, based on raw forecasted demand or speculation. This forecasted demand may be achieved or not, depending on the global business dynamics. However, in Lean Education, this is not acceptable. The degree programs delivered and education curriculum adopted, should be based on the concept that production, which is training of graduates, should be driven by customer demand in the upstream of the value chain and not naïve forecasts or interpolations. Below, table 1 demonstrates the differences between traditional education institutions and lean education institutions on various themes.

Table 1 - The Differences between a Traditional Education Systems and a Lean Education System

<table>
<thead>
<tr>
<th>Argument Theme</th>
<th>Traditional Education System</th>
<th>Lean Education System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organization Culture</td>
<td>Culture of royalty, favour, obedience and labour friction</td>
<td>Harmonious culture of involvement based on long-term development of human resources</td>
</tr>
<tr>
<td>Institution Operations</td>
<td>Division of labour and no problem solving skills for the workforce</td>
<td>Smart tools that assume standardized work and procedures, strength in problem identification and experimentation</td>
</tr>
<tr>
<td>Customer Relations Management</td>
<td>Produces Graduates needed by employers in large quantities at acceptable quality locally</td>
<td>Produces Graduates according to employers requirements with zero defects or to world class standards. Continuously get feedback</td>
</tr>
<tr>
<td>Education Business Strategy</td>
<td>Graduate-focused strategy on exploiting economies of scale</td>
<td>Customer/Employer-focused strategy on identifying and exploiting shifting competitive edge and sustainability</td>
</tr>
<tr>
<td>University Management Organization</td>
<td>Hierarchical structures that encourage orders and discourage the flow of information that unveils errors and deficiencies</td>
<td>Flat structure that encourages initiative and flow of information, that highlights defects, errors and deficiencies, as well as promotes innovation and continuous improvement in all activities</td>
</tr>
<tr>
<td>Graduate Inventory Levels</td>
<td>Large number of part-time and/or repeater students</td>
<td>Small number or no repeater students. Right first time all the time</td>
</tr>
<tr>
<td>Educator or Teacher Empowerment</td>
<td>Little input into University Operations</td>
<td>High responsibility for identifying and implementing quality</td>
</tr>
</tbody>
</table>

Source: Author
8. THE EDUCATION PRODUCTION LINE EXPOSED

8.1 The Higher Education Product Exposed

It should be noted that the student is not really the product; instead the product is the education of the student. A student need to be considered as a co-worker, who is actively participating in the design, execution and creation of the product and hence, is the co-manager of his/her education. Therefore, the student should be involved in the continuous improvement of the institution processes in quest for quality creation of the product, namely the Graduate. The student is the host of the education product.

8.2 Education Process versus Engineering Process

“Management of scholars” defines a business process as a collection of activities that takes inputs and creates output that is of value to customers. It should be noted that education, like any other business process, is also made up of various process operations or activities. Let’s take a typical engineering process like the manufacturing process illustrated in figure 1 below, and you will see that the work-in-progress moves from one work station to another until a finished product is created. Quality checks are also carried throughout the process in the work stations. At each stage the raw material is slowly being converted into a finished product.

Figure 1 - Engineering Manufacturing Process

The education process is no different from the engineering manufacturing process described above in figure 1. Graduates with knowledge, character, know-how and wisdom are the products. The raw material in education is the knowledge they have when being admitted into higher education. The work-in-progress is the students in different years of their degree program towards the final year. In each year students take a number of courses which can be referred to as the work centres where educators, who are the operators, add value by teaching them new knowledge, character and skills. Over the years, for instance: three, four or five, the students are converted into finished products, called the ‘Graduates’. Then these products are ready to be delivered to the customers who can be referred to as the potential employers, the student sponsors and the society at large, as well as provide raw materials for further higher education. This enlightenment is illustrated in figure 2 below.

Figure 2 - Higher Education Process
The quality of an engineering product depends on many factors, such as the quality of the raw materials, skills of the operators, plant and equipment capabilities, the production methods adopted and measurements or instrumentation being used to check for quality precision, accuracy and consistence. In a similar sense, the education product quality, namely the graduate quality, can also be attributed to a number of factors, such as the degree program curriculum, course content and quality of raw materials. This refers to the incoming students, teaching methodology and learning ads adopted. The quality of the operators, who in this case are educators/lecturers and the measurement tools adopted for assessments are also critical. Quality sourcing of raw materials in higher education continues to be more crucial. Today, the graduates select the universities on various factors such as institution reputation, potential for placement at the end of the degree program, teaching and service excellence, as well as tuition and other costs.

8.3 Customers for the Education Product Revealed

The primary customers for higher education are the students, since they are the hosts and the employers. The employers choose graduates based on the knowledge and know-how they have acquired over the years of training. It can be stated that the employers are the ultimate customers for the higher education product and in most instances today, they are not happy with the quality of the current graduates. However, the customers for higher education product may be segmented as follows:

a) **Student** - this is the host of the product and hence, the co-manager of the education production line and should always be considered first when defining quality in education.

b) **Student Sponsors** - parents, family members and organisations who pays for the students costs.

c) **Potential Employers** - these are organisations relying on the education of the student upon graduating to achieve their organization goals.

d) **Society at Large** - pays substantial costs of education through taxes and requires future participation of the student as an educated citizen and expects them to contribute positively to the general welfare of society.

8.4 Society Expectations from Higher Education Product

The education customers’ expectations may be clustered into four as follows:

a) **Knowledge** - this enables students to continuously learn in relation to what they already know

b) **Know-how** - this enables students to apply knowledge to work environments and this should be from different areas of learning

c) **Wisdom** - this is the ability to distinguish what is vital from what is a not vital and set priority to resource management.

d) **Character** - this may be said to be a combination of knowledge, know-how and wisdom, coupled with motivation to deliver value for the stakeholders.

From the foregoing, it can be said that, quality in education should look at the improvement of teaching, which is basically showing how to solve real world problems and developing the ability for continuous learning, which is figuring out how to solve real world
Learning to see what is wrong in higher education environment is the beginning of change. If you start seeing trouble or anything unusual in the institution, then you have reached the threshold required to begin change. Below are the ten forms of higher education wastes that are true for most, if not all Universities and Colleges of today. Waste is defined as any institution activity or operation performed that does not add any value to the customer satisfaction (student, employer and society) or experience. Organization customers may be broadly defined as any person or activity next in the organization process or value flow chain. Japanese scholars Taiichi Ohno and Shigeo Shingo of Toyota Motors Corporation initially proposed the first seven for the engineering environment, however, today they stand true to any other business environment, including the higher education Liker and Maier [11]. They are:

1) **Motion** - this involves movement of people (educators, students, technicians, administrators, etc) and/or teaching equipment that does not add value to the student learning experience. Examples include looking for information, forms, materials, educators and equipment that is located far from the point of operation. Taking files from one office to another, documents/forms requiring a number of signatures for approval, unnecessary moving around of students, due to disorganization and transportation of teaching resources or information. It is common to find operators (educators/lecturers) or front desk staff and the like, leaving workstations to fetch required information. This is true waste and does not add any value to the customer (student and society) satisfaction in any way.

2) **Waiting Time** - this is idle time created in the University processes when people, information, equipment or materials are not at hand for use. Examples include; waiting for long procedures, lecturers, approvals, signatures, forms, results and reports. Also when students are waiting for appointments, procedures and expert guidance. Others are invasion of teaching time, class interruptions, presenting conflicting information causing delays, poorly scheduled meetings and late arrivals by parties involved. Causes of all these may include poor understanding of the standard time required to do a task and lack of accountability for delivering on time. Students and other clients do not come to University and pay for waiting!

3) **Uncertainty** - this is when educators doing the work are not confident about the best way to perform the tasks. Examples of this may include; same course curriculum being taught in different ways by different educators, unclear teaching methodology, unclear course curriculum, unclear laboratory procedures and management orders. Causes of this may be due to the lack of standardized specification of activities or procedures of work. A common language for various activities in the University, that is referred to as standardization, should not be an option.

4) **Defects** - These include University activities that contain errors or lack something of value. Examples include: teaching errors, presentation errors such as at seminars, graduation ceremonies, lectures, etc., document errors, data entry errors, variation of same task outcomes, service delivery errors, product (graduate) errors and lost records. Others are failure to meet deadlines and deficient online information or facilities. Causes of these may include lack of understanding of what ‘defect-free’ University processes are and lack of standardization in work processes and quality management.

5) **Processing** - this includes the activities in the processes of service delivery that do not add value from the customer perspective. Examples include extra unnecessary steps, too many approvals, requirements confusion, charting during working hours,
missing procedure requirements and too much regulatory paperwork. Causes of these may include poor work area design that does not promote smooth value flow, complex and multiple data forms, as well as use of obsolete procedures and forms. Others include creating reports no one reads, unclear roles and responsibilities and repeated manual entry of student data or results.

6) **Over-Production** - this refers to redundant work. Examples include: duplicate work, charting during office hours, multiple forms with same information, re-creating already existing knowledge, teaching previously taught curriculum or subjects/topics, creating a new report when the data exists in the different department, uncoordinated or unplanned review meetings and creating departmental silos as well as protectionism. Also include the use of University facilities faster than is necessary without regard to whether the services are required downstream in the value chain by the ultimate customer (employer). Causes of this may include misinterpretation of regulations, as well as poor communication between departments and staff.

7) **High Inventory** - this is when there is more material at hand than what is required to do the work. Examples include overstocked and/or outdated books, misused of the available materials, poor understanding of supply and demand, outdated supplies not deleted, obsolete equipment not discarded, unread emails, unfinished projects, files waiting to be worked on, unresolved challenges and negative public and international image. Others are unnecessary work-in-progress (repeating or part-time students) and finished products beyond what is needed on the normal basis (producing more graduates in degree with less or no demand). The “evils of inventory” result in wasted space, administrative costs, storage costs and security costs. The institution image is also dented if the graduates' knowledge cannot help them find work to contribute positively to the general society.

8) **Under-utilized Resources** - these include the organization workforce, time, facilities and equipment available, which are not used to get the optimum benefit. Examples include minimal hours of operation for the library, computer facilities and laboratory for the students. Note that this is not referring to resource stressing, but just maximizing on the benefits that can be offered by the available resources. In most institutions it is common to find multi-skilled employees who are just using a single skill and then the rest are in idle mode. This idle potential ought to be tapped and rewarded so that it contributes to value creation for the organization. Closed University culture to innovation and change, untapped areas of passions for staff and choosing short-term cost reductions that do not motivate staff are among other critical factors in higher learning institutions.

9) **Poor Communication** - this involves information and data waste. This is common in most, if not all of today's institutions. All universities experience problems nearly every day, which need to be solved by different people in the organization. To solve most of these problems, information is required and now when the information, which is useful to solving the problem at hand, is not available or is difficult to retrieve, inefficiency in problem solving cannot be avoided which will have significant impact on the value creation and service delivery for the stakeholders such as the customers. Note here that problem solving is part of the work and therefore it can be stated that poor communication or lack of information is also a serious waste in Universities and this should be avoided if excellence is to be achievable. Information should be availed to all people in the organization at the earliest possible time, whether positive or negative. Examples of good information practice include annual audited financial reports which can be posted on websites, annual key performance indicators (KPI's), such as profit margins or losses, should also be published to all employees for review and improvement.
10) **Misused Resources** - this is when resources in the University are allocated for a particular project, with particular objectives and then on the way, some stakeholders divert the resources to other purposes, owing to some special or personal interests. This, in most cases, cannot be recovered and is not replaced at all. This must be known to be a complete waste and does not add value to organizational rationale and stockholders’ expectations. Stick to the annual budget and adopt the long-term philosophy over short-term or personal gains.

### 10. DEPLOYING LEAN INITIATIVE FOR HIGHER EDUCATION EXCELLENCE

To ensure that the above identified wastes are eliminated, there is need to deploy quality management strategies, such as ‘Lean Thinking Initiative’, which may be deployed through creation of the following:

a) **University Business Transformation Council** - this will involve institution wide deployment of the quality improvement strategies that will focus on core university business processes. Strategic projects should be identified and problem solving teams established, including the project charter and leader. This may initially be set as a two year project but will continue to evolve to meet the market changes and requirements.

b) **Strategic Quality Improvement Charter** - Universities should establish a quality improvement charter which will act or work as a ‘Quality Improvement Centre’. This is where clients like students, staff and employers can report quality issues and make suggestions for improvements. This should be headed by a Quality Professional whose main duty is to look into the quality issues all around the University.

c) **Quality Problem Solving Teams** - this will involve the creation of teams for strategic opportunity or operational weakness determination and analysis. The focus here should be on the key customers and analysis of the critical business process through the various faculties or departments. Problem solving projects should be established which will be linked to the strategic initiative. This can be done at faculty and departmental levels, where a critical few projects will be identified to highlight the current state of the University performance and establish where resources may be focused. The initial time line may be one year but will continue to evolve as change happen. The initial project should be quality awareness sensitisation and training of staff within the faculty and other departments. This may last for three to six months, while simultaneous implementing is being learnt.

d) **Strategic Environmental Policy** - with today’s ways of doing business, moving towards sustainability, it’s no option that Universities need to develop and put in place, including on websites, their ‘Environmental Policy’ to show their eco-commitment and how it will address issues concerning its interaction with the environment. Green stewardship needs to be planted in higher learning institutions so that it grows in the minds of all graduates and employees. Lean thinking is part of the sustainability strategy.

e) **Teaching Excellence Rewards** - administrators, academics and even students should be rewarded for performance excellence such as scholarships for discovery and innovation, etc. This should be handled by the Quality Improvement Charter.
Deploying lean thinking to improve graduate quality and service delivery in higher education - Nkumbwa

11. CONCLUSION

This paper has presented the quality view of service delivery and graduate quality in higher education environment. The author believes that better understanding of society expectations, from higher education institutions and academic processes, is vital for performance, re-engineering of most higher education institutions today, if excellence is to be achieved. It is hoped that, with the areas covered, most education stakeholders will be edified about the need for change of the current practice and the new meaning of quality in today and tomorrow’s higher education institutions. Among the expected benefits of deploying lean thinking in higher education environment are included simplification of both the academic and administrative processes, elimination of waste activities or operations, creation of value flow in the institution, reduction of activity cycle time and graduate lead time, reduction of error rates and delays, as well as an increased institution bottom line due to massive savings, improved productivity and efficiency in the value delivery to the stakeholders. Corporate image and sustainability are also enhanced.

12. REFERENCES


A SYSTEMS THINKING APPROACH TO ADVANCING GTL TECHNOLOGIES

Andrew Nyamvumba, Grace Kanakana and Prof John Trimble
1Department of Industrial Engineering
University of Tshwane, South Africa

ABSTRACT
A systems thinking approach is suggested in order to enhance South Africa’s competitive advantage in developing gas-to-liquid technologies by taking advantage of integration of information and communication technologies in the energy cluster (smartgrid technologies), need for cleaner technologies, plastics and recycling industry, and a clean development mechanism.

Gas-to-liquid (GTL) processes involve the conversion of natural gas into synthetic liquids that can be transformed into different useful hydrocarbon fractions including liquid transportation fuels. Such technology can also be used to convert worldwide scattered natural resources such as methane gas, coal, and biomass to fuels (coal-to-liquid (CTL) and biomass-to-liquid (BTL)). Major GTL technologies are the Fischer-Tropsch (FT) and ECLAIRS processes.

Based on a systems approach an analysis of the energy, polymer, recycling, and smartgrid industries, the authors share their proposal for advancing GTL technologies. The paper presents how the GTL, plastics and the recycling industry are combined on a smartgrid to stimulate a competitive advantage for integration of GTL in the energy portfolio-mix for most countries worldwide.
1. BACKGROUND AND HISTORY OF GTL

Gas-to-Liquids (GTL) fuel is claimed to be a pollution free alternative fuel, made from natural gas, which can be used in conventional diesel and jet engines, and has other long hydrocarbon chains as by-products (e.g., paraffin wax and naphtha). It can be used as a stand-alone fuel, or blended with diesel. GTL fuel is also claimed to deliver a superior environmental performance with significantly lower local emissions than ultra-low sulphur (50ppm) diesel.

The literature review indicates that GTL technologies have been in development for nearly a century (1922) developed by two German scientists, Franz Fischer and Hans Tropsch. The Fischer-Tropsch (F-T) GTL process was named after them. It was used in Germany during World War II to produce approximately 600,000 bbl per year of liquid transportation fuels from coal.

Carthage Hydrocol is known to have conducted further development in Brownsville, Texas, from 1948 to 1953 to convert natural gas rather than coal. The plant constructed had a production of 365,000 bbl per year but was shut down and dismantled when there was a dramatic rise in natural gas prices.

South Africa is documented to have seriously focused on the F-T process beginning in 1955 using coal as feedstock. By the early 1980’s, in Bintulu, Malaysia, Shell used the F-T process in a GTL plant to convert natural gas to fuels and by 1986 Mobil acquired rights to process natural gas from a giant gas field off New Zealand and built a GTL plant. The plant processed gas first into methanol, then into gasoline. What is significant in this historical path is that whenever oil prices collapsed plants were no longer profitable, and in the case of New Zealand the plant was sold to Fletcher Challenge Ltd. in 1993. In general, GTL processes were resumed recently in the late 1990s and the last two decades saw the number of announced projects and demonstration projects increase dramatically, especially in the past ten years.

Whereas various levels of policy makers continue to debate and find ways to meet the challenges posed by climate change, investments in technology development are guided by competitive advantage (a major driver) despite market failures associated with suggested market mechanisms. Considering climate as a public good raises a global commons problem. This is because externalities associated with increasing GHG emissions in the atmosphere which may lead to climate-related disasters and ecological imbalances. This can result in a “tragedy of the commons” because of the situation that is exhibited by the conditions of lack of incentives for action and the accruing public costs and as such relies on cooperation through global agreements as a solution because privatization is not a solution.

Other challenges are associated with knowledge of what causes climate change and the relationship of climate change to global. In general, improvements in increasing the level of certainty through science and economics prediction will increase the responsiveness and sensitivity facing global warming debate and enforcement of agreements. This, it is hoped, will better shape policy for stabilization of green house gases concentrations and consequently reduce feedback systems that are associated with current increase the global warming.

Therefore, in pursuit of how science, policy and politics should shape the policy options, a systems approach offers a methodology that guides how energy efficiency and emissions intensity are systematically studied as related to influence the trend of technological evolution. The challenge of integration of different dimensions of the problem for a practical strategy is the basis for a systems approach to determine what individual contributions can offer competitively and the context of the interaction with other contributions in terms of technology. Clearly such a problem involves political, social,
A Systems Thinking Approach To Advancing GTL Technologies
Nyavumba, Kanakana, Trimble

industry, firm levels which qualify the approach to be complex. This is because of several interacting systems that must act together to solve, or at least dissolve, a problem. Therefore, a systems approach is justified as a better approach. A systems approach was seen as a way of addressing complex problems and issues (Hitchins, 2007: 17) with three ways in which problems could be addressed. The three ways are: resolving (find an answer that is good enough), dissolved (change the situation in some way - move goalposts), and solved (find correct answer).

This approach is useful to deal with the ‘tragedy of the commons’ given the nonbinding process and failure of consensus (Rio (1992), Kyoto (1997), Bali (2007) to Copenhagen (2009)) compounded by uncertainty about future warming scenarios and impacts of this warming if no policy interventions as well as uncertainty of the extent to which the world population cares. The usefulness is most exhibited in the thinking about the system architecture and deciding climate policy instruments. These in turn justify investment in GTL technologies.

South Africa’s contribution to energy security in the context of climate change is considered along the advancement of GTL technologies integrated in a smartgrid environment. The focus of this research is to highlight exploration of the stranded gas reserves as bulk fuel for generators and for transportation in island distributed energy resources or/microgrids where natural gas/methane reserves are available. This effort is a contribution to global climate stabilization under the scenario where natural gas is expected to replace fossil fuels. Presently, there is no focus on this option as the integration of GTL and smartgrid technologies are mostly at pilot and testing phases.

2. STRATEGY AND POSITIONING

Opportunities for South Africa’s industries (PetroSA, Sasol, and others) are identified along energy security, consultancy and patents in catalysts improvement, polymer and chemical industries. With the economic growth of India and China driving increase in energy demand and consequently expanding global energy demand, stable supply of energy has become an important and urgent matter. GTL technology with which the available natural gas reserves can be converted into liquid fuel is an extremely effective technology for achieving the diversification of primary energy supplies. The fuel produced using GTL technology is also expected to be an environmentally friendly clean fuel.

Patterns of energy supplies and demand are indicated by energy flows worldwide which point to a growing mismatch between the regions where energy is needed and where natural resources are found. Figure 1 below shows projected regional flows for gas by 2030 as forecast by the International Energy Agency (IEA). Energy supply security challenge for some countries with stranded gas reserves lies in recovering the gas and integrating it in the national and regional market. Estimates by IEA (2005 -2030) indicate a need for world investment of over $8 trillion USD in the gas and oil sectors, and over $11 trillion USD in the electric power sector, and points out that much of this investment is likely to be in developing countries.
Natural gas proven reserves and stranded natural gas throughout the world indicate potential for investment in GTL technologies advancement. This potential is further explored in various simulations which tend to agree that natural gas, over the short term will replace fossil fuels in global climate change stabilisation efforts.

The world natural gas reserves by geographical region as of January 1, 2009 (MIT ESD 128s: energy in the context of climate change) are summarised in Figure 2 below.

GTL also must be recognized as much more than FT diesel, jet fuel, waxes, and naphtha. It’s a raw material for ammonia, methanol and dimethyl ether (DME), the latter product used for power generation and clean-fleet (modified diesel) fuelling in certain markets. Comparatively, GTL is a chemical process that differs from liquefied natural gas (LNG) as it is not simple refrigeration and is easily transportable to liquid fuel markets as shown in Figure 3 below.
3. ADVANCES IN GTL TECHNOLOGY

Currently the debate centres around how the carbon price associated with greenhouse gas emissions can be used as a policy instrument or implemented worldwide due to the global trend of strict environmental restrictions on emissions. In order to supply strategic GTL products in existing oil markets GTL technology requires the following improvements to be fully competitive and replace fossil fuels where possible:

1) Reduction of construction and operation costs for a GTL plant,
2) Enhancement of energy efficiency, and
3) Reduction of carbon dioxide emissions.

Review of some GTL technologies by Robertson (1999: vii) indicates at least six advances:
1) A basic Fischer-Tropsch (F-T) technology,
2) Sasol’s F-T technology,
3) Exxon’s Advanced Gas Conversion for the 21st Century (AGC-21),
4) BP’s compact steam reformer,
5) Syntroleum’s diluted nitrogen technology, and
6) An F-T process using DOE’s Ceramic Membrane.

In addition to the above advances, the Japanese JOGMEC’s Revolutionary GTL Technology uses innovative catalyst technology in both the syngas reformer, where conversion of methane takes place, and in the FT reactor, where carbon monoxide and hydrocarbon are converted to a range of hydrocarbon products. JOGMEC’s GTL technology offers the following features:

1. It can use around 20 to 40 mol percentages of carbon dioxide as a feedstock.
2. It may be able to provide an attractive, competitively priced option for marketing stranded gas, which cannot be brought to market due to lack of access to transportation infrastructures.
3. GTL products are environmentally sound because they are sulphur-free and aroma-free, and very efficiently combusted, particularly when used as transportation fuels.

These new technologies could drive the cost of GTL distillate fuel and petrochemical feedstock plants down and if eventually combined with cogeneration, petrochemicals production and (in some cases) excess hydrogen sales. The question of when they will achieve the target of competitiveness with oil refining, considering crude trading in the $15-18/barrel range, remains an R&D challenge that South Africa could exploit.
Simulation of the future of natural gas shows replacement of fossil fuels by natural gas (MIT Future of Natural Gas, Interim Report) under the context of climate change. South Africa has potential for advancing research in several dimensions of reducing cost through technology development and market design to develop the market for GTL and related industries. Figure 4 below shows the increase of share of natural gas under the emissions reduction scenario.

Figure 4 - MIT natural gas simulation projections under the emissions reduction strategy

Furthermore, projections by the MIT Joint Program Report #186 indicate the importance of natural gas both as a replacement of the fossil fuel market and a complementary option in ramping power, as well as meeting peak demand in the short and medium.

4. CHALLENGES AND SMARTGRID

Climate change is important and critical because it threatens the stability of the world’s climate, economy and population as a result of rising greenhouse gas emissions. It is commonly agreed amongst scientists in the world that most of the world’s carbon dioxide emissions come from the way energy is produced and used. Consequently, energy policy has to play a major part in meeting this challenge.

Policy is informed by several barriers ranging from technological to social and financial. Since there is no single ‘silver bullet’ solution to such a complex issue, integration of systems and technologies (called the smartgrid in this context - for which no definition is agreed upon), is important.

The smart grid is a collection of tools and technologies that are integrated to deliver energy at a minimal loss and in a better managed manner by using information and communication technologies. The need and applications of the smart grid are to develop a digitized infrastructure that can allow in renewable energy once policy and regulation have incentivised the development of these energy sources.

The smart grid provides the possibility of better energy management, allow for renewable energy to be incorporated into the grid to contribute to transition to a fully cleaner energy source, and as such improve the energy portfolio mix in which GTL would play a significant role. This is the important consideration for a systems thinking approach to establishing an energy policy.

Since the causes and consequences of climate change are global, and national governments have to take action, the ultimate solution must be collective global effort. Current trends
indicate that global emissions are set to reach double pre-industrial levels before 2050, with severe impacts on global climate and the global economy.

A key conclusion from the technical teams is that in the long-term the cost of inaction would be far higher than the cost of tackling climate change now. The teams also make it clear that the costs are lowest if nations act together. On the basis of present policies, global energy demand will be more than 50% higher in 2030 than today, with energy related greenhouse gas emissions around 55% higher.

The major challenges are related to regulatory and infrastructure costs. The International Energy Agency (IEA) forecasts that $20 trillion USD of investment will be needed to meet these challenges by 2030. The investment decisions that will be taken over the next two decades will be critical in determining the world’s climate and the security of its energy supplies. Figure 5 shows GTL activity and Figure 6 shows an example of a smartgrid.

**Recent worldwide GTL activity and announcements**

![GTL activity map](image)

*Figure 5 - Recent GTL activity worldwide. Source: Sasol*

The importance of the island and decentralized energy resources are necessary to offer a full solution where such an integration leads incentives to develop cleaner energy and in cases where tradeoffs for gas to other products is of concern.
5. CONCLUSION
There is potential for South Africa to lead in GTL technology advancement in combination with the smartgrid (in its general definition in this paper). This is critical for emission reduction. GTL offers opportunities for competitiveness and climate change stabilization. This, however, requires identifying areas of research and development focus including catalysts and chemicals. The transport market is generally likely to keep dominating the market and use of GTL as back-up for renewable sources has great potential.

Greenhouse gas emission targets have important implications for technology innovation given policy changes in the electricity, transport, and other economic sectors. The global nature of the appropriate solution to stabilisation of climate change creates complexity that is best achieved through a systems approach to problem solving.

Development of F-T process efficiency is critical as well as ECLAIRS (which is less capital intensive may be explored for small and medium scale GTL process). The smartgrid is a necessary infrastructure to consider as it will justify allocation and inform policy options.

6. REFERENCES
A CONCEPTUAL FRAMEWORK FOR THE MANUFACTURING OF TOOLING BY METAL CASTING IN REFRACTORY MOULDS AND SHELLS PRODUCED BY RAPID PROTOTYPING (RAPID CASTING FOR TOOLING)

Kasongo Nyembwe¹ *, Deon De Beer² and Shepherd Bhero³

¹ School of Mechanical Engineering and Applied Mathematics
Central University of Technology
dnyembwe@uj.ac.za

² School of Mechanical Engineering and Applied Mathematics
Central University of Technology
deond@vut.ac.za

³ Engineering Metallurgy Department
University of Johannesburg, Johannesburg
bshepherd@uj.ac.za

ABSTRACT

In this paper an alternative methodology for the manufacturing of metallic tooling referred to as Rapid Casting for Tooling (RCT) is proposed. It essentially consists in an integrated process chain comprising five steps including Computer Assisted Design (CAD), Casting Simulation, Rapid Prototyping (RP) Metal Casting and Finishing Operations. A theoretical assessment based on fundamental principles and available research on direct RP of sand moulds and shells as well as metal casting is conducted in order to understand the technological merits and limitations of RCT compared to the established tool and die making processes such as CNC machining and Rapid Tooling. The analysis of RCT focuses on manufacturing parameters amongst them cost and time as well as tool quality in terms of dimensional accuracy and surface finish. The study indicates that RCT offers potential for significant time compression, cost efficiency as well as good quality.

Key words: Tooling, Rapid Prototyping, Metal Casting, Rapid Tooling, CNC Machining

¹ The author was enrolled for a D Tech (Mechanical Engineering) degree in the School of Mechanical Engineering and Applied Mathematics, Central University of Technology

* Corresponding author
1. INTRODUCTION

Metallic tooling includes dies, sand casting patterns and permanent moulds used in mass production process such as forging, metal casting and injection moulding. They are essentially manufactured by CNC machining and rapid tooling. These two die and tool making processes change very rapidly as a result of the ever increasing demand to satisfy the stringent requirements of the tooling market. Users insist on superior quality of tools in terms of durability, surface finish and dimensional accuracy. Furthermore short delivery times and premium price are critical to increase profitability as well as quick turnover of new products on the market.

New challenges have emerged in the last ten to fifteen years. First the worldwide shortage of the tool making skills due to threat of digital revolution is one of the current problems facing the tooling manufacturing industry. Second, the environmental quest to reduce waste and elimination of pollution has contributed to the challenges. Finally the stiff competition from countries such as China and India that are able to produce very cheap products has given rise to an incessant demand for new tool manufacturing methods that are easy to learn, competitive, technologically advanced and green.

Rapid Tooling (RT) processes have appeared ten to fifteen years as a means of addressing some of the challenges mentioned above. RT is the application of rapid prototyping technologies to the manufacturing of tooling. RP itself is the additive manufacturing of physical parts in various materials including metal layer by layer directly from CAD files. Various rapid tooling processes are commercially available and extensively reviewed in the literature [1]. Their major benefits compared to CNC machining include the reduction of lead design time and the cost efficiency.

In reviewing the different applications of RP to tooling manufacturing, it is clear that that RP of sand mould has not been fully use for tool and die making. One rare case study in the literature reports the manufacturing of a metallic tooling for injection moulding by casting in a sand mould produced on an EOSINT S RP machine [2]. The reasons for the limited application of RP sand moulds for casting of metallic tooling are due to the disadvantages of sand casting as a manufacturing process including the difficulty to produce thin wall parts, segregation, coarse microstructure of castings as well as poor dimensional accuracy and surface finish of parts. These technical challenges have restricted the use of metal casting for the production of quality tooling.

The low thermal conductivity of sand moulds results in coarse metallographic microstructure of castings with inferior mechanical properties compared to metallic processes such as forging and machining. Furthermore shrinkage during metal solidification and cooling combined with mould wall movement due to the metallostatic pressure are at the origin of dimensional accuracy problems. Finally chemical reaction of metal and sand as well as the presence of interstices between sand particles at the surface of the mould are responsible for poor surface finish of castings. These metallurgical phenomena are explained in details in various specialised foundry technology manuals.

Major developments have taken place in the field of metal casting to improve the cast structure and the quality of castings. Worth to be mentioned is the use of advanced and powerful casting simulation software, synthetic refractory sand and superior mould coating technologies and products. Casting simulation is common practice in modern foundries and reduces scrap rate due to porosity and shrinkage defects. Synthetic sands with improved thermal properties allow the production of thin wall casting of superior mechanical properties. In addition new mould coating with fine suspension and high refractoriness are continuously being developed to improve the surface finish of sand casting. These coatings minimise metal penetration and reduce surface mean surface roughness of castings.
Taking into account the above sand casting developments, the use of RP moulds for the casting of metallic tooling deserves attention as a feasible tool manufacturing option. Its main benefits will possibly comprise the short manufacturing time and cost deriving from metal casting and rapid prototyping advantages.

In order for this proposed tooling manufacturing to be recognised amongst the tool making technologies, its conceptual framework needs to be formalised. This framework referred to in the present paper as Rapid Casting for Tooling (RCT) will validate tool design that can be successfully manufactured by casting in RP moulds. It will also describe the various steps that need to be performed in order to produce a sound and high quality tool.

The conceptual development and theoretical assessment of RCT is the object of the present paper. In a future paper RCT will be assessed with the aid of case studies. The next section is a background to the fundamentals and recent developments of the rapid prototyping of sand moulds and the metal casting.

2. BACKGROUND

2.2 Metal Casting

Metal casting is a manufacturing technique for metallic components that consists of the solidification of liquid metal in the cavity of a refractory sand mould or metallic die. The basic steps in the sequence of production with minor modifications depending on specific casting processes are shown in Figure 1. Some of these steps have experienced fundamental changes over the years in casting design, patternmaking and preparation of moulding materials.

The design of parts to be produced by metal casting is specific and critical to the process. Amongst several principles the casting drawings need to include contraction allowances to compensate for shrinkage during solidification. In addition the filleting of sharp corners is essential to prevent hot tears defects. Finally section transitions in the part should be optimal and modified if necessary to avoid hot spot defects or the formation of carbides compounds in thin sections. Drawing computer programs (CAD software) are nowadays used by the foundryman to speedily assist in the creation three dimensional models of cast parts [4]. These models are used for 3D visualisation and prediction of possible problems during production.
In addition to the physical dimensions and shape of the casting, casting design also provide accurate dimensions of gating and feeding systems. These features will be part of the mould and the die. The gating system constitutes the piping system of the mould to allow complete filling of the mould cavity in a manner that prevents casting defects such as formation of oxides, non-metallic inclusions, mould erosion, cold laps and short runs. The feeding system serves as reservoir of molten metal to compensate for liquid metal shrinkage and thus avoids macro-shrinkage defects. Modern metal casting methoding makes use of casting simulation computer programs to design and dimension feeding and gating systems. Many of these software packages are available including Magmasoft and Procast that are predominantly used locally. The main advantage of these simulation programs is the right first time (RFT) concept that leads to considerable reduction of casting defects during production and elimination of error and trial.

Patternmaking is an essential step of traditional metal casting to manufacture the foundry tooling that is used to create the mould cavity. This step is generally the rate determining step and cost driver of the metal casting process. The reason for the high cost is the required skill as well as the cost of technology involved such as CNC machining to produce a sand casting pattern or a tool steel die. Since mid 1990 several rapid prototyping (RP) technologies have been successfully used to manufacture foundry tooling [5]. The benefits have been the reduction of tool manufacturing time and cost. Lately new patternless processes based on the RP have been able to directly produce sand moulds and shells for metal casting thus eliminating the need for patternmaking. These processes are described in section 2.2.

Moulding materials are mixed together and mechanically or chemically processed to produce strong moulds and shells that can withstand the metallostatic pressure of liquid metal during casting. These raw materials essentially comprise various refractory sands and binders. They have a profound impact on the final properties of the casting in terms of

Figure 1 - Flow diagram of casting production [3]
minimum section thickness, mechanical properties, dimensional accuracy and surface finish. Developments in the field of refractory sands have seen the emergence of synthetic sand with better thermal properties to produced thin wall castings and high integrity with superior mechanical properties [6, 7]. The use of fine sands has tremendously improved the surface finish. New mould coatings based on high melting point fillers have reduce the incidence of surface defects such as metal penetrations and burn on. Finally modern catalysts and resins have improved the strength of moulds thus reducing the problem of mould wall movement that poses of dimensional accuracy challenges [8].

2.2. Rapid prototyping of sand moulds

Rapid prototyping is the manufacturing of parts, layer by layer in special machines directly from computer three dimensional model files. Rapid prototyping techniques employ five basic steps to produce a part [9]. These steps are:


b) Conversion: Conversion of the CAD file into Stereolithography (STL) format. This standard format represents a three dimensional surface as an assembly of planar triangles. It is a well defined and easy-to-handle format that enjoys wide support in the CAD fraternity. It is the preferred format for visualisation, analysis programs and rapid prototyping manufacturing.

c) Slicing: Preparation of the STL file to be built by carefully choosing the suitable build orientation and creation of support structure for the model during building.

d) Layer by layer construction: Manufacturing of the part using one of the several techniques listed below.

e) Clean and finish: curing to provide strength, surface treatment, etc.

New developments in the field of rapid prototyping have seen the production of investment casting shells or sand moulds directly from 3D-CAD files [10]. The manufacturing processes are based on Selective Laser sintering (SLS) and 3-D Ink Jet Printing technologies. In the SLS a laser beam is used to selectively fuse pre-coated foundry sand into a solid part that will become component of the mould or shell. In 3-D Ink Jet Printing the principle is similar to ink impression on a piece of paper from an ink-jet printer. In this case in a layer by layer fashion, an ink-jet printing head selectively deposits or “prints” binder fluid that fuses the powder particles together in desired areas. Commercial examples of technologies for sand mould and shell production based on SLS and 3DP include the Direct Croning Laser Sintering Process (www.actTech.com) and the Direct Shell Production Casting (www.soligentechnologies.com).

The main advantages of RP processes of sand moulds and shells have been the reduction of lead design time. Prototype casting for visualisation and metallurgical evaluation can be produced quickly because no pattern is required. In addition because of their patternless nature, RP processes have reduced design time cost. These processes have also been used for manufacturing low volume casting [11]. It has been shown that for low volume casting RP processes are more cost efficient than conventional sand casting processes.

The development of RP of sand moulds and shells as well as the achieved progress in metal casting to overcome previous technical limitations are opening new possibilities for this manufacturing method of metallic components as describe in the section below.
3. THE RAPID CASTING FOR TOOLING FRAMEWORK

3.1 Rapid Casting for Tooling (RCT)

In the present study Rapid Casting for Tooling (RCT) is proposed for the manufacturing of metallic tooling. Essentially dies, permanent moulds and sand casting patterns will be produced by metal casting in refractory sand moulds or shells obtained by RP processes. These RP processes have been introduced in section 2.2. As such RCT is approached as an indirect tool manufacturing process based on two pillars: Rapid Prototyping and Metal Casting. These two key elements characterise and differentiate RCT from the other tooling manufacturing processes. They will also confer to RCT the advantages of near net shape and time compression. RCT is further refined by the use of computer aided design and analysis techniques such as solid modelling and casting simulation. Benefits such as right first time of RCT will be directly derived from these computer aided techniques.

RCT does not deal with the initial design stage of the tooling nor does it concern itself with the selection of the suitable alloys for the dies or moulds. These considerations depend on the final application of the tooling as well as the type of finished parts that will be produced. The inputs to RCT are the drawings of the final metallic tooling accompanied with the exact material specifications.

3.2 RCT Process chain steps

The steps of RCT process chain are linked together and interact with each other in what is referred to in this study as the RCT framework. The framework is made up of two structural parts comprising the validation stage and the process chain. The validation stage determines if a tool is able to be manufactured by RCT based on its minimum wall thickness and its material chemical composition. Metal casting limitations in terms of difficulty to produce thin wall parts as well as segregation have been explained in the background section. Only the design that successfully passes the validation test criteria can be considered for manufacture by RCT process.

RCT process chain is made up of five steps including Solid Modelling, Casting Simulation, Rapid Prototyping, Casting and Finishing operations. The flowchart of RCT framework is presented in Figure 2. The functions and main characteristics of these processing stages are summarised in Table 1. They are described in the sections below.
A Conceptual Framework For The Manufacturing Of Tooling By Metal Casting
Nyembwe, de Beer, Bhero

Figure 2 - RCT framework diagram showing the two structural pillars: Validation and Process Chain

Table 1 - Functions and main characteristics of RCT process chain steps

<table>
<thead>
<tr>
<th>RCT Steps</th>
<th>Functions</th>
<th>Main characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solid Modelling</td>
<td>Casting and mould design</td>
<td>STL file format</td>
</tr>
<tr>
<td>Casting Simulation</td>
<td>Gating and feeding systems design</td>
<td>Rule-based software</td>
</tr>
<tr>
<td>Rapid Prototyping</td>
<td>Moulding and core-making</td>
<td>• Selective Laser Sintering</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Three Dimensional Printing</td>
</tr>
<tr>
<td>Metal Casting</td>
<td>Casting of tool</td>
<td>Gravity casting</td>
</tr>
<tr>
<td>Finishing Operations</td>
<td>• Minimal machining</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Polishing</td>
<td>• Annealing</td>
</tr>
<tr>
<td></td>
<td>• Heat treatment</td>
<td>• Homogenization</td>
</tr>
</tbody>
</table>

3.2.1 Solid Modelling

In the solid modelling step, 3D representations of the tool and the mould are created using suitable CAD software. This first step forms the backbone of the RCT process chain ensuring three main functions namely:
- The input definition for downstream steps
- The casting design of the tool
- The modelling of the shell or sand mould components.

The downstream steps such as casting simulation and rapid prototyping require solid models in order to function. The CAD software that will be used should therefore be able to save solid models in a compatible format throughout the process chain thus facilitating exchange
between steps and avoiding alteration of the accuracy during transfer. Stereolithography Language (STL) format can be one such format.

Casting design consists in the modification of the original tool design to incorporate contraction allowance and filleting. Casting section transitions are also carefully examined and altered if necessary to avoid contraction defects such as hot spot and hot tears. Finally the modelling of mould or shell follows soon after the casting design. Here shell and mould cavities are generated from the 3D drawings. This will include the drafting of the sand mould features that are cope and drag cavities, parting line, core with its support (or print), gating and feeding system.

The solid modelling of the metallic tooling and the shell and mould components are saved as STL files to be used at casting simulation and RP stages. The casting simulation step needs to be completed before the RP manufacturing of the mould can start. There is an iterative loop connecting casting simulation and solid modelling. Ideally this loop ends as soon as the casting simulation confirms that the design is satisfactory with no defect expected during the metal casting step.

### 3.2.2 Casting Simulation

Casting simulation is carried out for two reasons including ensuring defect-free casting, i.e. devoid of shrinkage cavities, solid inclusions and oxidation and a reduction of manufacturing time by eliminating the reliance on trial and error. These two elements of the casting simulation provide cost advantage for RCT because of the “right first time” approach philosophy.

The following steps are followed during RCT casting simulation to ensure product quality:

- Import the appropriate casting model from solid modelling
- Carry out solidification and mould filling analysis to determine the location of the last freezing region and the possible flow pattern likely to cause turbulence and splashing
- Select the correct riser unit and gating system to attach to the mould
- Repeat casting analysis to check efficacy of the procedure
- If turbulence, splashing and shrinkage persist modify the feeder unit
- Verify new feeding design until elimination of all problems
- Send STL file to solid modelling step for modification of mould and casting design. This is shown in Figure 1 by the double arrows between the solid modelling and the casting simulation steps.

### 3.2.3 Rapid Prototyping

RP step constitutes the moulding stage of RCT. No pattern is required as sand moulds and shells are produced directly layer by layer from the STL file obtained in step 1 and 2. Processes such as the Direct Croning Laser Sintering Process (DCLSP) or the Direct Shell Production Casting (DSPC) can be used to manufacture sand moulds and refractory shells for metal casting in RCT. The moulds and shell are coated in order to improve the surface finish of castings.

### 3.2.4 Casting

The casting stage of RCT combined three steps of the traditional casting production flowsheet (Figure 1) including metal melting, casting and fettling. The suitable alloy specified in the tool material is melted in a furnace. The liquid metal is poured in the sand mould or shell and feeds by gravity. The casting is fettled to remove the gating and feeding components then sand blasted to a rough tool.
3.2.5 **Finishing Operations**

In this last stage of RCT, the near-net cast tool is brought to the required final dimension and surface finish. This is achieved by minimal machining and polishing that form part of the finishing operation step. This step also includes possible heat treatment in order to improve the mechanical properties and durability of the tool by altering the internal metallurgical structure of the tool.

4. **ASSESSMENT OF RCT**

Now that RCT conceptual framework has been presented and described, it is necessary to critically discuss how it addresses the requirements mentioned in the Introduction. In particular, a theoretical assessment of the manufacturing time and cost, the overall implementation cost, the tooling quality in terms of the surface finish and the dimensional accuracy, the user friendliness of the process is conducted in order to establish if RCT can compete CNC machining and other existing rapid tooling processes for metallic tooling.

In the RCT analysis that follows below only the RP and metal casting steps are considered because they have the biggest impact on the manufacturing time and cost as well the tool quality when RCT is compared to other tool manufacturing processes such as CNC machining and Rapid Tooling. The other RCT steps that are solid modelling, simulation and finishing operations are also found in some form or another in other tool manufacturing processes.

4.1 **Manufacturing time**

RCT manufacturing time will essentially depend on the RP step which is the rate determining step generally measured in terms of days while the metal casting that consists of melting and pouring of casting is measured in hours. Pouring is controlled by the Bernoulli equation [7] and generally expressed in time of the order of second. The melting time will be a function of the type of the melting furnace used and will be controlled by conduction, convection and radiation equations of heat transfer. The RP manufacturing time is a function of the size of the component to grow and the scan speed of RP equipment.

RCT is an indirect tooling manufacturing. In this case a mould or shell is grown that will be used as the casting cavity. It can therefore be expected that it is quicker to build the shell or the mould than it is to grow the part itself. In this sense RCT will be quicker than direct tooling and machining. Correspondingly, RCT is expected to be faster than the investment casting because there is no need for pattern manufacturing. Even in the case where lost wax patterns are produced by additive manufacturing, RCT provides better time-saving because the shell is manufactured by rapid prototyping instead of the usual and time-consuming procedures of dipping in refractory slurry. This was demonstrated by a study [12] that compared the investment casting of shells in selective sintering to traditional investment casting using RP patterns.

4.2 **Manufacturing Cost**

The RCT manufacturing cost is mainly driven by the additive manufacturing stage that is superior to melting and casting. The melting cost factors are listed in the literature [4] and include energy and labour costs. Of importance is the energy cost that will be a function of the type of furnace used and the superheat. In general electrical furnaces are cheaper and more environmental friendly than fossil fuel furnaces based on coke or natural gas [8]. Regarding the superheat necessary for good metal fluidity, the larger the superheat is the more the energy required hence the higher will be the energy cost.

The cost of rapid prototyping is determined by the technology used, whether SLS or 3DP. Because of the use of a laser, the manufacturing cost with the SLS is generally higher than
with 3DP. Laser usage time is the costing basis used when using Selective Laser Sintering. The reason for this practice is that the laser components have limited lifetime and need to be replaced frequently.

Many studies on the benefits of direct or indirect Rapid Tooling [13, 14, 15] seem to indicate that additive manufacturing of tooling is cheaper than machining. Cost savings of 50% are often reported by companies offering tooling manufacturing services [1]. This advantage of additive manufacturing is likely to be extended to RCT over machining.

Studies have also shown that for lower number of parts direct manufacturing of mould and shell considerably decrease the overall cost of metal casting compared to casting processes that require patterns [11, 16]. RCT has been designed for the production of single component and therefore this benefit applied when compared to traditional metal casting processes for tooling.

4.3 Tool quality

The tool quality refers to the mechanical characteristics, the dimensional accuracy and the surface finish of the final metallic tooling. All the RCT steps from CAD to finishing operations will influence the tool quality.

4.3.1 Mechanical properties

The mechanical properties include ductility, elongation, hardness essentially developed during the solidification of the casting. The factors that affect the cast structure and subsequently the properties are well explained in foundry technology textbooks [3, 6]. They include the cooling rate, the grain size, the alloy chemical composition, the presence of inclusions and shrinkage porosity, etc.

The prediction of these defects and the necessary preventive measures in order to avoid their occurrence takes place during the casting simulation step of RCT. In particular shrinkage and oxide inclusion defects can be eliminated by proper design of the gating and feeding systems.

It is possible to produce sound castings of superior mechanical properties making the final tooling superior to tooling obtained by direct tooling that generally suffer of porosity causing low durability (Song et al, 1997). Machining will generally produced superior mechanical properties to casting.

4.3.2 Dimensional accuracy

This property will depend on solid modelling and rapid prototyping steps. In the RCT description, it was explained that contraction allowance are included during the solid modelling in order to compensate for the metal shrinkage during solidification. On the other hand the accuracy of additive manufacturing will be a function of the technology used. Studies have been conducted to compare Selective Laser Sintering of sand parts and the 3D Printing process found that SLS was in general more accurate than 3D Printing [17]. Hence dimensional accuracy will be comparable to rapid tooling processes but inferior to machining.

4.3.3 Surface finish

The surface finish of the final metallic tooling will be a function of the mould surface smoothness. In the study mentioned above [17] the surface finish of SLS parts are compared to 3DP parts. It was found that SLS was giving a better surface finish. With the use of mould coating, it is possible to improve the surface finish of RP mould and shells and
produce a final tool with an improved surface finish than a rapid tooling part. However the surfaces finish of machining will always be superior.

4.4 Ease of use and learning curve

Additive manufacturing process is easy to learn. Once the CAD modelling is completed, the file is sent to the RP machine where the part is grown with limited human involvement. RP set-up requires no programming. In addition the process is automatic and environmentally friendly. These benefits of RCT confer its advantages over CNC machining and ordinary metal casting.

4.5 The overall implementation cost

Since their introduction a decade ago in most developing countries, the cost of rapid prototyping processes has been in free fall and lower than complete installation of machining centres. As such RCT is expected to be cost effective to implement. An additional advantage of RP of refractory moulds and shells is that they can also be used for the production of casting thus having a dual role in foundries.

5. CONCLUSION

In this paper RCT framework has been conceptually described, explained and assessed. It is made up of a validation stage to determine if a tool can be manufactured by RCT based on its chemical composition and its minimal wall thickness. This is followed by the RCT process chain itself of five steps including CAD modelling, casting simulation, rapid prototyping, metal casting and finishing operations. The casting produced at the end constitutes the metallic tool that can be used in mass production processes of finished goods such as metal casting and injection moulding. As such RCT is an alternative tool and die making process along with to CNC machining and rapid tooling.

From the theoretical assessment, an appreciable understanding of RCT in terms of factors influencing the manufacturing time and cost as well the precision of casting has been gained. It was found that RP and metal casting were the two pillars of RCT with profound impact on the characteristic of the final tooling. Benefits of these two steps comprising amongst other time compression, low manufacturing cost and good quality of parts are extended to the entire RCT process chain thus providing it potential advantages compared to conventional tool and die making processes. In future publications, these advantages will be evaluated using experimental case studies.

6. REFERENCES


   a. [http://www.me.psu.edu/lamancusa/rapidpro/primer/chapter2.htm](http://www.me.psu.edu/lamancusa/rapidpro/primer/chapter2.htm)


PERFORMANCE ENHANCEMENT IN THE MILLING OF Ti6Al4V

G.A. Oosthuizen1, G. Akdogan2 and N. F. Treurnicht3
1Department of Industrial Engineering
University of Stellenbosch, South Africa
tiaan@sun.ac.za
2Department of Process Engineering
University of Stellenbosch, South Africa
gakdogan@sun.ac.za
3Department of Industrial Engineering
University of Stellenbosch, South Africa
nicotr@sun.ac.za

ABSTRACT

Ti6Al4V is one of the most widely used titanium alloys in aerospace applications, but its machining remains a challenge. This is partly due to the lack of understanding of the thermal- and mechanical demands this alloy exerts on tool materials. This paper discusses the prominent tool demands and introduces a tool wear map that enables operators to understand the cause of failure, in order to take remedial action. Optical microscopy imaging is reported for experiments to evaluate these wear maps for both rough- and finish milling operations. It is concluded that the developed maps are appropriate and necessary for the production optimization in the machining industry.

OPSOMMING

Ti6Al4V is een van die bekendste titaan allooie in die lugvaartbedryf, maar die masjinering van die allooi bly steeds 'n uitdaging. Dit kan grootendeels toegeskryf word aan die gebrek aan inligting rondom die termiese- en meganiese eise wat die allooi aan die snybeitel materiaal stel. Die dokument beskryf die belangrikste snybeitel uitdagings en stel 'n verweringkaart voor wat masjien prosesontwerpers help om masjineringenal谯ings te verstaan en hul prosesse te optimeer. Optiese mikroskopie fotografie is gebruik om die kaarte vir beide rofwerk en afrond freeswerk te valueer. Die gevolgtrekking is dat die innoverende verwering skaart toepaslik is en bydrae tot die optimering van masjinering prosesse.

1 The author was enrolled for an PhD (Industrial) degree in the Department of Industrial Engineering, University of Stellenbosch
2 The author was a lecturer in the Department of Process Engineering, Stellenbosch University
3 The author was enrolled for an PhD (Industrial) degree in the Department of Industrial Engineering, University of Stellenbosch
*Corresponding author
1. INTRODUCTION:

Machining is a major cost contributor, but at the same time, a differentiating factor for a competitive advantage. Ti6Al4V is one of the most widely used titanium alloys in aerospace applications. This is due to their good strength-to-weight ratio and superior corrosion resistance [1]. Despite the increased popularity of Ti6Al4V in aircrafts such as the Airbus 380 and Boeing 787, many of the same qualities that enhance Ti6Al4V’s appeal for most applications, also contribute to its being one of the most difficult materials to machine [2]. Figure 1 illustrates Ti6Al4V’s machinability by comparing it to the machinability ratings [3] of other materials, with 1018 steel as benchmark.

![Comparison of machinability ratings](image)

**Figure 1 - Comparison of the machinability ratings of some popular materials [3]**

The recommended cutting speeds ($v_c$) for titanium alloys of around 30 m/min with high-speed steel (HSS) tools, and around 60 m/min with cemented tungsten carbide (WC) tools, results in rather low productivity [4]. These machining challenges can be divided into thermal- and mechanical tool demands. The tool face temperature ($T_v$) is a function of the cutting speed ($v_c$) and exposure time to this thermal load. The longer the duration of exposure time, the larger the volume of the tool edge, that is exposed to the critical tool temperature [5]. The combination of titanium’s low thermal conductivity and the fact that approximately 80% of heat generated [6] is retained in the tool, results in a concentration of heat in the cutting zone. These issues cause complex tool wear mechanisms such as adhesion, oxidation and diffusion [7,8]. Researchers [9,10] also measured temperatures of 900ºC at a cutting speed of 75m/min. The mechanical demands are a combination of the work piece chip load on the cutting edge and the machining vibrations. The contact area of a serrated Ti6Al4V chip was found to be only a third [11] of the contact area of a regular continuous chip. This results in high pressure loads on the cutting edge. Catastrophic (earlier than scheduled replacement time) tool failure, due to vibration while in cut, is caused by self-excited chatter [12] and forced vibrations due to the formation of shear localization [13] and the fluctuating friction phenomena between the tool and chip flow.

According to published work [14] the chip segmentation phenomena significantly limits the material removal rates and causes cyclic variation of force. The combination of a low Young’s modulus (114GPa) [15], coupled with a high yield stress ratio allows only small plastic deformations [1] and encourages chatter and work piece movement away from the tool. Figure 2 illustrates the importance of different tool characteristics that is required from innovative tool materials to efficiently machine Ti6Al4V.

---

1. Performance enhancement in the milling of Ti6Al4V
Oosthuizen, Akdogan, Treurnicht

---

173
Performance enhancement in the milling of Ti6Al4V
Oosthuizen, Akdogan, Treurnicht

Figure 2 - The importance of the different tool characteristics required when designing tool material for the machining of Ti6Al4V

Ti6Al4V is associated with very much the same demands as Inconel 718, although Inconel 718 generates higher cutting forces [16] and is more abrasive [3]. The most challenging demands for a tool material to machine titanium alloys are the tribo-chemical and impact related wear mechanisms. Tribo-chemical wear is a combination of molecular-mechanical wear and corrosive wear; and may be considered a thermally activated process whereby the work piece material and tool material react in such a manner as to remove material from the tool on an atomic scale [17]. Titanium’s chemical reactivity becomes problematic at temperatures above 500°C. Apart from diffusion wear, it has a strong affinity to adhere which leads to chip adhesion (welding) onto the tool cutting surface. Once a built up edge develops, tool failure follows rapidly [18]. Tool failure can be initiated by one or a combination of several forms of wear, which in an advanced stage may lead to failure [17]. Note that flank and crater wear are gradual and more predictable than fracture, which occurs suddenly.

2. WEAR MAP APPROACH
Ground-breaking titanium alloys and part designs increase the pressure on aerospace component manufacturing teams to meet cost, quality and on-time delivery demands. High material removal rates (MRR), high product quality and non-catastrophic tool life without tool breakage are crucial to efficient manufacturing. Cutting tool manufacturers have improved tool capabilities by increasing the density of substrates, designing specific geometries, using precise edge honing techniques and developing new coating technologies to manage heat at the tool and work piece interface when machining Ti6Al4V alloys. Using the right combination of cutting tools and machining parameters, process designers can achieve effective production rates. As the demand for titanium alloys grow, so too will the metalworking technology that supports efficient machining. Together with the high demands that titanium places on machining capacity, machining industries and manufacturers that use the most efficient technology will be first in line to benefit from surging demand for titanium parts.

One way to select the most suitable machining conditions is to examine the trends of tool wear with varying machining conditions. Wear maps representing the rates of tool wear in a two-dimensional space defined by the machining conditions in metal cutting have been found to assist in such an optimisation process [19]. It was found that machining conditions (chip load and cutting speed) played a critical role in determining the extent of wear on these cutting tools. These two machining parameters were chosen as the map’s axes,
Performance enhancement in the milling of Ti6Al4V

because it had been suggested that the wear of cutting tools depends principally upon them [20]. These maps, which describe the global wear characteristics of cutting tools, can be used to achieve a certain degree of optimization of machining operations. Tool wear maps were taken a step further by Kendall [21], when a safe operating region was proposed for cutting tools by superimposing possible boundaries. It was suggested that as long as the machining parameters lay within this safety zone, the tools would not suffer catastrophic tool failure (failure before scheduled replacement time). The key to designing high performance cutting tools is to identify the most economical scheduled replacement time (SRT) used in the industry for the specific component. Thereby the tool suppliers can optimize the operational conditions, whereby the cutting tool will withstand the demands of the machining process for a longer period of time than the SRT. These optimum operational conditions are defined as the safe zone [22] as shown in Figure 3. According to the group’s research, literature [23] and tool suppliers [24] the SRT for milling Ti6Al4V components is around 30 min. Therefore 30 min was used as SRT for this study.

![Figure 3 - Tool wear map for milling of Ti6Al4V as a function of cutting speed and feed rate (Adapted from[22,17])](image-url)
Using the wear characterization map, which is developed specifically for the milling of Ti6Al4V alloys, improved milling parameters and strategies, could be achieved. The used tools were examined with an optical microscopy imaging and were characterized. Thereby they are categorized into a failure region. Only the maximum thermal- and mechanical load failure regions were considered in these remedial actions, as the industry find it important to machine at the highest material removal rate (combination of high cutting speed and chip load). After identifying the failure regions, it was possible to take remedial actions as illustrated in Figure 4.

Figure 4 - Remedial actions to different failures

Thereby, it was possible to compromise between the tool wear and the machining time relative to the material removal rate; and remedial actions could be considered to improve the process.

3. EXPERIMENT 1: ROUGH MILLING WITH CARBIDE TOOL MATERIAL
The following section examines the performance for rough milling with carbide tools.

3.1 Experimental Setup
Table 1 details the experimental equipment and the milling parameters used. The signal of the dynamometer was converted by three charge amplifiers (type 5011) into a proportional...
Performance enhancement in the milling of Ti6Al4V

Oosthuizen, Akdogan, Treurnicht

voltage signal (+/-10V). The experiments were conducted with cutting speeds ranging from 40-100 m/min and feed rates of 0.1-0.4 mm/z.

Table 1 Experimental equipment

<table>
<thead>
<tr>
<th>Machine tool</th>
<th>Heckert HEC 500 D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tool and inserts</td>
<td>HM90 E90A-D20-3W20</td>
</tr>
<tr>
<td></td>
<td>Cemented carbide: APKT 1003R8T-FF (TiAlN PVD coated)</td>
</tr>
<tr>
<td>Lubrication</td>
<td>Flood (Emulsion Hebro 565B, 10%)</td>
</tr>
<tr>
<td>Force measurement</td>
<td>Kistler 3D dynamometer type 9255B</td>
</tr>
<tr>
<td>Charge amp settings</td>
<td>Type 5011 (+/-10V)</td>
</tr>
<tr>
<td></td>
<td>Measuring range +/-2000 N</td>
</tr>
<tr>
<td>Cutting speed</td>
<td>40, 70, 100 m/min</td>
</tr>
<tr>
<td>Feed per tooth</td>
<td>0.1, 0.25, 0.4 mm/z</td>
</tr>
<tr>
<td>Axial depth of cut</td>
<td>2 mm</td>
</tr>
<tr>
<td>Radial depth of cut</td>
<td>20 mm (Full)</td>
</tr>
</tbody>
</table>

The axial depth of cut ($a_p$) was fixed at 2 mm, with a full radial depth of cut ($a_e$). Flood lubrication was used to reduce the thermal demands on the tool. Figure 5 illustrates the mapping of trials from literature [1,5,7,4] and industry and illustrates the experimental parameters for this study.

![Figure 5 - Wear mapping for rough milling of Ti6Al4V with carbide tools](image)

3.2 Results and discussion

The results of carbide tools that were used for rough milling operations are illustrated in Figure 6. The tools were examined with optical microscopy imaging and were characterized to categorize into a failure region.
### Figure 6 - Failure modes for carbide tools in rough milling of Ti6Al4V

<table>
<thead>
<tr>
<th>Speed (m/min)</th>
<th>Feed (mm/z)</th>
<th>Tool Life (min)</th>
<th>Tool Failure Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>0.4</td>
<td>&lt;30</td>
<td>Catastrophic tool failure: Not preceded by extensive crater wear on the rake face</td>
</tr>
<tr>
<td>70</td>
<td>0.4</td>
<td>&lt;30</td>
<td>Catastrophic tool failure: Not preceded by extensive crater wear on the rake face</td>
</tr>
<tr>
<td>100</td>
<td>0.4</td>
<td>&lt;30</td>
<td>Catastrophic tool failure: Preceded by extensive crater wear on the rake face and deformation</td>
</tr>
<tr>
<td>40</td>
<td>0.25</td>
<td>&gt;30</td>
<td>Safe zone: Non-catastrophic tool failure</td>
</tr>
<tr>
<td>70</td>
<td>0.25</td>
<td>&lt;30</td>
<td>Catastrophic tool failure: Not preceded by extensive crater wear on the rake face</td>
</tr>
<tr>
<td>100</td>
<td>0.25</td>
<td>&lt;30</td>
<td>Catastrophic tool failure: Preceded by extensive crater wear on the rake face</td>
</tr>
<tr>
<td>40</td>
<td>0.1</td>
<td>&gt;30</td>
<td>Safe zone: Non-catastrophic tool failure</td>
</tr>
<tr>
<td>70</td>
<td>0.1</td>
<td>&lt;30</td>
<td>Safe zone: Non-catastrophic tool failure</td>
</tr>
<tr>
<td>100</td>
<td>0.1</td>
<td>&gt;30</td>
<td>Safe zone: Non-catastrophic tool failure</td>
</tr>
</tbody>
</table>

Performance enhancement in the milling of Ti6Al4V

Oosthuizen, Akdogan, Treurnicht
As illustrated in Figure 5, results from industry and experimentation indicate that rough milling operations are still limited to cutting speeds of 30-60 m/min with feed rates of 0.1-0.25 mm/z [24]. As a complementary perspective the results proved that the cutting speeds (thermal load) for carbide tools in rough milling conditions are limited to 100 m/min under a low mechanical loading with \( f_z = 0.1 \text{ mm/z} \), while the mechanical loading is limited to \( f_z = 0.25 \text{ mm/z} \) at \( v_c = 40 \text{ m/min} \). Figure 7 illustrates the cutting forces for the different rough milling conditions.

![Cutting forces for carbide tools in rough milling of Ti6Al4V](image)

**Figure 7 - Cutting forces for carbide tools in rough milling of Ti6Al4V**

The cutting forces were generally influenced by the chip load, while the cutting speed showed no significant influence. The segmented chip formation and the frictional phenomenon caused cyclic variations in the force components.

4. EXPERIMENT 2: FINISH MILLING WITH PCD TOOL MATERIAL

The following section examines the performance of different tool materials for finish milling with PCD tools. Similarly, the wear map was used for finish milling experiments with PCD.

4.1 Experimental Setup

Table 2 details the experimental equipment and the milling parameters used. The experiments were conducted with cutting speeds ranging from 100-300 m/min and feed rates of 0.025-0.05 mm/z.

<table>
<thead>
<tr>
<th>Table 2 - Experimental equipment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine tool</td>
</tr>
<tr>
<td>Tool and inserts</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Lubrication</td>
</tr>
<tr>
<td>Force measurement</td>
</tr>
<tr>
<td>Charge amp settings</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Cutting speed</td>
</tr>
<tr>
<td>Feed per tooth</td>
</tr>
<tr>
<td>Axial depth of cut</td>
</tr>
<tr>
<td>Radial depth of cut</td>
</tr>
</tbody>
</table>

The axial depth of cut \( (a_p) \) was fixed at 2 mm, with a radial depth of cut \( (a_e) \) of 1mm. Flood lubrication was also used in these experiments. Figure 8 illustrates the mapping of
Performance enhancement in the milling of Ti6Al4V
Oosthuizen, Akdogan, Treurnicht

Experiments from literature [5,15,23,25] and illustrates the experimental parameters. According to literature finish milling of Ti6Al4V shows promising results at cutting speeds in the range of 175-200m/min [9] with $f_z=0.025-0.05\text{mm}/\text{z}$ [26]. The axial depth of cut is low ($a_p=0.5\text{mm}$) and the radial immersion ranges from an $a_e=0.5-2\text{mm}$.

Figure 8 - Wear mapping for finish milling of Ti6Al4V with PCD tools

PCD tooling has a relatively high cost. Catastrophic tool failure in the final finishing stages of the component carries a high economic risk. Therefore, finishing at a mechanical loading considerably lower than the fracture limit is sound practice. This pragmatic approach is mirrored in the experimental design.

3.2 Results and discussion
The tools were also examined with optical microscope- and scanning electron microscope imaging and were characterized to categorize into a failure region. Thereby, it was possible to compromise between the tool wear and the machining time relative to the material removal rate; and remedial actions could be considered to improve the process. The results in Figure 9 illustrate from the PCD tools that were used for finish milling operations.
Performance enhancement in the milling of Ti6Al4V  
Oosthuizen, Akdogan, Treurnicht

<table>
<thead>
<tr>
<th>$v_c$ (m/min)</th>
<th>$f_z$ (mm/z)</th>
<th>TL (min)</th>
<th>Tool wear</th>
<th>Safe zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.05</td>
<td>&gt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>200</td>
<td>0.05</td>
<td>&lt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>300</td>
<td>0.05</td>
<td>&lt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>100</td>
<td>0.0375</td>
<td>&gt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>200</td>
<td>0.0375</td>
<td>&lt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>300</td>
<td>0.0375</td>
<td>&lt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>100</td>
<td>0.025</td>
<td>&gt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>200</td>
<td>0.025</td>
<td>&gt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
<tr>
<td>300</td>
<td>0.025</td>
<td>&gt;30</td>
<td>PCD</td>
<td>Non-catastrophic tool failure</td>
</tr>
</tbody>
</table>

Figure 9 - Tool wear modes PCD tools during finish milling of Ti6Al4V
As found in literature [5] the mechanical loading for PCD should be in the range of 0.019-0.032 mm. Using the wear map approach, these experiments identified \( v_c = 200-300 \text{ m/min} \) with \( f_z = 0.05 \text{ mm/z} \) as the optimum parameters for finish milling with PCD cutting tools. As found in literature [23] a cutting speed of more than 300 m/min led to Ti6Al4V material burn. Figure 10 illustrates the measured cutting forces for PCD cutting tool material at different cutting conditions.

![Cutting forces for PCD tools when finish milling Ti6Al4V](image)

The cutting force during finish milling operation was generally influenced by the chip load, while the cutting speed had a small influence.

4. CONCLUSION
From the results recorded in the experiments the following conclusions are reached. The most important tool demands have been discussed. A tool wear map, specifically designed for the milling of Ti6Al4 is illustrated. The tool wear map serves the purpose of clarifying the complex interaction between the machining parameters to process designers. Thereby a knowledge based iterative process improvement is facilitated, simultaneously enhancing machining knowledge. Optical microscopy is utilised for experiments to evaluate these wear maps for both rough- and finish milling operations. The developed maps are appropriate and necessary for the production optimization in the machining industry. In a similar way these maps can contribute to enhance the experimental labs in developing innovative tools for the milling of Ti6Al4V.

5. ACKNOWLEDGEMENTS
This research is funded by the Advanced Manufacturing Technology Strategy (AMTS) programme, an initiative of the South African Department of Science and Technology (DST). The Fraunhofer IWU in Chemnitz, Germany is acknowledged for hosting the experiments and valuable advice.

The RPD lab at Stellenbosch University, under the leadership of Prof Dimitrov, is acknowledged for the organizing of the overseas experimental work.
Performance enhancement in the milling of Ti6Al4V
Oosthuizen, Akdogan, Treurnicht

6. REFERENCES

Performance enhancement in the milling of Ti6Al4V
Oosthuizen, Akdogan, Treurnicht


THE APPLICATION OF A RESILIENCE ENGINEERING MODEL FOR ORGANISATIONAL INNOVATION

René Pellissier* and Fernando de Sousa
Department of Business Management
University of South Africa
pellir@unisa.ac.za

INUAF, President of Apgico, Portugal
fernandosousa@gaim.pt

ABSTRACT

Becoming globally competitive has requires specific skills in developing economies as these environments are far more complex. Business as a complex system requires acknowledgement that we cannot control organisations to the degree that a mechanistic perspective will. Moreover, as the system’s environment changes, so does the behaviour of its agents. Thus, the behaviour of the system as a whole can change. Linear strategies and decision making techniques become irrelevant with a shift to patterns and relationships between entities. Developing economies, especially, are more prone to the implementation of non-linear solutions because of the nature of the variables, the changes and interplays between the variables, the significant human focus and the consequent organic nature of the competitiveness. These variables introduce an unavoidable element of unpredictability/randomness into any management decisions. Complexity allows for pattern recognition which requires focusing on competencies, activities, technologies or resources signaling patterns that will impact on innovation, especially with respect to organisational, management and technological forms. Technology as knowledge is an intrinsic part of the pattern recognition the implementation of the above forms of innovation. This paper discusses the role of non linear management theories in a complex environment with regards to these innovations. In this sense, a resilience engineering approach provides the space for innovation implementation and the focus on organisational and management innovation through complex adaptive systems. The literature abounds with research on product and process innovation but less is said about organisational, management or technological innovation and their implementation.

Key words; Complexity, pattern recognition, organisational and management innovation, resilience engineering, strategic management, management decision making.
1 PREAMBLE AND SCIENTIFIC JUSTIFICATION

This is a position paper presenting an ‘academic’ perspective on the question on what is beyond current innovation practices with regards to organisational renewal. It also redefines innovation in terms of the complexity of the new world of work. The question presented, raises issues of complexity and issues of linearity. The answer may be unconvincing to some and strange to others, but then, so is the very nature of innovation. John Naisbitt (in Hamel [24]), maintains that ‘Academics are afraid to go beyond data. Alfred North Whitehead said that a proposition doesn’t have to be right, it just has to be interesting. Academics don’t understand how liberating it is not to have to be right. When you have to be right, you become a prisoner.’ we will not try to defend my solution here; it is for the reader to prove me wrong. In this, we hide behind Stephen Hawking’s [26] understanding of scientific proof in that, if you cannot find us wrong, then yes, we could be right! Hawking argues that there are two requirements for any theory to be true:

1. It must accurately describe a large class of observations on the basis of a model that contains only few arbitrary elements and
2. It must make definite predictions about the results of future observations.

According to him (Hawking) any physical theory is always provisional in the sense that it is only a hypothesis - one can never prove it, only disprove it. Thus, no matter how many times the results of subsequent experiments agree with the theory, one cannot predict whether the next time, the result may not contradict the theory. On the other hand, to disprove a theory, one only has to find a single observation contradicting the theory. Thus, a new theory really is an extension of the previous theory. In practice, a new theory may be devised following upon add-ons to the previous theory, through modifications or even abandonments. (An example of such is that of Newton’s laws of gravity, which in essence, have been proved not correct by Einstein’s predictions. However, since Newton’s laws are simple and the difference between its predictions and those of general relativity is very small, Newton’s work stands.) Moreover, Heisenberg’s uncertainty principle allows uncertainty .. until that outcome or result can actually be observed to take place.

Therefore, you are invited to build on this theory by arguing, or disagreeing but be sure to add something that can make it grow. This paper redefines innovation in terms of the complexity of the world of work currently experienced and links resilience engineering and complexity for new forms of innovation in terms of its organisation and its management.

INNOVATION IS RENOVATION

Up until recently, we would have maintained that information technology (IT) is unambiguous. That is still true. But beyond that (e.g. IT), value lies in the information and knowledge obtained, distributed, shared, distributed and created. And, when that is done, innovation begins. And, when that is done, there is more innovation. To us, innovation is not invention or entrepreneurship on a grander scale. Innovation is knowledge used in a unique and different way. Innovation is new thinking. That thinking can be radical, disruptive or incrementally different. But it is not more of the same. It is renewal and renovation.

The rise of the ‘networked’ or ‘information’ economies (Castells [9]; Nonaka et al.[44]; Roos, Roos, Draonetti & Edvinsson,[53]; Volberda [69]) signified by terms such as ‘intensive’, ‘innovative’ and ‘flexible’ suggest that firm performance is increasingly predicated on the efficient and effective use of knowledge (Grant [21]). Acting knowledgeably, rather than repetitively, is becoming critical because not only does it prompt learning from experience and provide insight into possible commercial futures, it is hard to imitate and can be strategically distinctive (Kogut and Zander, [37]). Innovation is generally the result of cumulative dynamic interaction and learning processes involving many stakeholders. Here innovation is seen as a social, spatially embedded, interactive learning process that cannot be understood independently of its institutional and cultural
context (Cooke, Heidenreich, and Braczyk [12], Lundvall [40]; Fornaciari & Dean [20]). Since Roberts’[51] definition (of innovation) maintains that an innovation can only be seen as innovation if it is has implementation and commercial value, it is important to measure the impact of innovation. Ravichandran [49] believes that measuring the impact of innovation activities will depend on (1) the typology, (2) the degree of departure from the preceding product/service or process, (3) the extent of usefulness of the innovation and, (4) the volume of profitability generated. Smith [59] identifies four types of innovation based on the work done by Henderson and Clark [27]: incremental (refining and improving the existing design within an established architecture), modular (use is made of new technology and components, within an existing system); architectural (an established system links existing components in a new way), and radical (involving a total new design using new components). In the last case the innovation can be disruptive. Henderson and Clark’s framework shows that systems and components in innovation are inextricably linked. For instance, architectural innovation reconfigures an established system to link existing components in a new way.

We define innovation as a continuum of activities as below: Innovation as renovation is the outcome of a series of interrelated activities on a continuum, starting with creative discovery, then entrepreneurship, and, finally commercial exploitation. In this, leadership is redefined, processes, systems and culture may be redesigned and organisations search for and find new meaning. This definition allows for product/process innovation but also includes organisational/management/technological innovation activities.

![Innovation types](image)

**Figure 1 - Innovation as a continuum of activities**

The more well-known forms of innovation are process innovations (‘performing an activity in a radically different way’, Davenport & Pruzack [15]); service innovations (‘a new way of providing a service, often with a novel and very different business model, .... even an entirely new service’, Smith [14]) and product innovations (‘ ... a core design concept that performs a well-defined function’, Abernathy & Clark [11]. Product/process/service innovations thus comprise both systems and components, calling for an integrative model
for innovation beyond the either/or instrumentalist vs. radicalism approach of the past. Organisational innovation encompasses all of these whilst highlighting the the way businesses operate (Birkinshaw, Hamel, & Mol [23]). These authors maintain that organisational and management innovation is difficult as it questions existing practices and processes and our assumptions on the nature of the way things are.

While innovation concerns the processes of implementation, relying mainly on organisational communication and power, in the domains of production, adoption, implementation, diffusion, or commercialisation of creations (Spence [63]), creativity remains exclusive to the relation established between the creator and his product, where not even originality and usefulness are important, but only the notion of ‘trying to do better’. The latter is connected to cognitive and emotional processes taking place at the individual level (Sousa, Monteiro, & Pellissier [60]; 2008; Sousa, Pellissier & Monteiro [61]; and Sousa, Monteiro [62]). If we relate creativity to problem definition, and innovation to decision implementation, this last step requires a series of problem definitions, in order to carry out a decision or an idea, thereby making it difficult to separate these concepts at an organisational level. In fact, when we move from the individual level to the team and organisational levels, creativity and innovation become more and more difficult to separate, so that we must agree with Basadur (1997)(4), when he says there is no difference between organisational creativity and innovation. Therefore, the moment we move to other levels besides the individual, we will use these terms (creativity and innovation) as synonyms, and we refer to organisational creativity as a system devoted to enhance creativity in organisations, thus using the definition proposed by Basadur [4].

As to the several approaches to identify types of innovation, either by separating the adoption of products and processes from its development (Cebon, Newton & Noble [10]) or, in a more classical way, product and process innovation (Adams [1]), most authors agree that innovativeness, or organisational (and management) innovation, is a separate type of innovation, which represents the potential of the workforce to promote changes to benefit of the organisation. As Huhtala & Parzefall [30] mention, ‘...to remain competitive in the global market, organisations must continuously develop innovative and high quality products and services, and renew their way of operating’, and they also maintain that companies increasingly rely on the employees continuous ability to innovate. Also, even though innovation may take place through the adoption or development an existing product or service, through investments in R&D or in technology acquisition, it is only through developing and sustaining a creative workforce that the organisation will succeed in maintaining the necessary potential to overcome difficult problems and situations that cannot be solved through investments only (Cebon, Newton & Noble [10]). To this end, technological innovation is seen is created as a result of an innovation activity comprising research (scientific), technical, organisational, financial and market activities. Technological innovation means objective improvement of the properties of a product or a process or a system of delivery relatively to the already existing products and processes. Less significant, technical or esthetical modification of products and processes which do not influence the performance, property, costs nor materials consumption, energy consumption and components consumption are not considered technological innovations. As Desrochers [17]) puts it, technological innovation can manifest in any business activity, for example in a basic activity, as well as in secondary and other activity (as defined in the system of national accounts), and in the auxiliary activity of sales department, accounting department, IT department etc. (e.g. computerisation of a sales department or a finance department of the enterprise can be considered a technological innovation).

The creative workforce potential is both the ability to retain creative managers and employees (McAdam [42]) and to provide an environment where each one will feels free and willing to contribute to organisational success. Aspects like raising job complexity, employee empowerment and time demands, together with low organisational controls (decision making, information flow and reward systems), are said to raise employee creativity (Adams [2]). However, more elements are necessary in order to make people willing and able to contribute to organisational effectiveness. For instance, supportive
leadership, knowledge acquisition, and team work procedures favouring creativity (Unsworth [68]) can add to success. Creative people (either managers or employees), are committed to their work and organisation, and so they may bring in important issues, provided that top management values their work and ideas. In fact, according to a Gallup Management Journal (GMJ) survey (Hartel, Schmidt & Keyes [25]), engaged employees are more likely to ‘think outside of the box’ and produce creative ideas than disengaged people; they also are more receptive to new ideas. The research concludes that engaged people tend to find and suggest new ways to improve their work and business processes, which may lead to the assumption that creative people have a deeper understanding of the organisational processes, by being in a privileged position to identify, define and find organisational problems.

To a certain extent, most of these can be achieved by the implementation of complex systems and the concept of resilience engineering to the business fundamentals. This is attained by elevating the importance of creativity and entrepreneurship and providing a system through which current goals are realised by new ideas can flourish. What are required are freedom to create, content and process skills to be able to create, and a supportive human environment (peers and team leader). The issues surrounding the potential of an organisation to innovate are still in its beginnings, although Mclean [43], Puccio, Firestien, Coyle & Masucci [48] and especially, Basadur and others [3, 4, 5, 6], did some empirical research. The major challenges are to define criteria to evaluate the impact of organisational innovation on process and product innovation (Wolfe [71]).

Innovation within the framework of a knowledge-based economy goes far beyond the linear or chain linkage models that have long been used in innovation theory to explain innovation processes in high-tech knowledge industries. Here innovation is seen as a social, spatially embedded, interactive learning process that cannot be understood independently of its institutional and cultural context (Cooke, Heidenreich, & Braczyk, 2004 [12]; Lundvall [40]). Strambach [64] suggests that the interdisciplinary view of innovation systems is concerned with understanding the general context of the generation, diffusion, adaptation and evaluation of new knowledge, which determines innovativeness. It follows that the focus is on non-technical forms of innovation as defined above. Common characteristics of the different approaches to innovation, identified by Edquist [18], include (1) innovation and learning at the centre, (2) a holistic and evolutionary perspective, and (3) an emphasis on the role of institutions. The increasing interdependence of technological and organisational change is a significant feature of systems of innovation, which means that technological innovation and organisational innovation have become increasingly important. These are combined with more diverse knowledge requirements which include not only technical know-how, but also economic, organisational, and sociological knowledge and competencies. The second reason for the increased interest in non-technical innovations is associated with the connection between the organisational innovation and the corresponding learning capacity. The acceleration of change that is part of the globalisation process means that organisational learning processes are more and more important for creating and maintaining competitiveness.

In organisational innovation, the unit for innovation is the organisation itself (Wolfe [71]). Although the outcome of the innovation may be process, product or service, the innovation needs to be undertaken through the creative inputs of the individuals and/or the management. As to measures of innovation, Dalal [14] mentions that the qualitative measure of emotional and psychological impact the innovation produces on the users (the ‘aha!’ moments); the quantitative measures of the total population of end users using the new innovation (and even helping co-create it), and the net new revenue generated for the company that can be attributed to the new innovation. Thus, while innovation concerns the processes of implementation, relying mainly on organisational communication and power, in the domains of production, adoption, implementation, diffusion, or commercialisation of creations (Spence [63], creativity remains exclusive to the relation established between the creator and the product, where nor even originality and usefulness
are important, but only the ‘trying to do better’, connected to cognitive and emotional processes taking place at the individual level (Sousa et al. [60]).

Some innovation theorists like Smith [14], believe innovation is meaningless without technology. We do not think so. Technology is a great platform for innovation achievement, but it is certainly not the only one. Technology is a good enabler for certain types of innovation. But real innovation comes from the inner self and individual contributions (in figure 1 we used creativity and entrepreneurship in this regard) and thoughts need to be given a place in organisations and in society to breed. We maintain that innovation takes place in the specific domains of product, process and/or service. However, there is more: innovation also takes place in leadership, culture, processes and systems, design, products and technology. Innovation is a thinking skill more than a doing skill. It transforms our views of current reality and focuses on renewal and regeneration. Danah Zohar [72] believes that ‘Most transformation programmes satisfy themselves with shifting the same old furniture about in the same old room. Some seek to throw some of the furniture away. But real transformation requires that we design the room itself. Perhaps even blow up the old room. It requires that we change the thinking behind our thinking - literally that we learn to rewire our corporate brains.’ Zohar’s ‘real transformation’ is really innovation.

The question to be asked is how strong is the link between innovation and real transformation. Our best contribution after days of deliberation and sleepless nights is 42 (after Douglas Adams’ [37] famous computer by the wise name of Deep Thought). Deep Thought had an answer for a vague question, but innovation sometimes takes place in an intuitive domain where there is no specific question (see serendipity). Either that or we do not know. We deeply believe that we have not at all tapped into our inner genius and the exploration of our own innovative sensibilities. Maybe we are afraid. Maybe we are so cultured to behave rationally and according to preset patterns that we consciously simply cannot remember how. As informationologists, we believe in the advantage and power of information and knowledge. If we try really hard, it is possible to design appropriate systems and technologies in place to obtain these. Then we need to do something with them. That makes Information Science and Knowledge Management possible. Creating space for innovation is more complex as it is not derived out of systems and technologies and has no theory to guide it; it comes from the inner self and only to those who are not afraid of newness. We thus turn to Jaworski’s Synchronicity principles [32]. He says: ‘Because of our obsessions with how leaders behave and with the interactions of leaders and followers, we forget that in its essence, leadership is about learning how to shape the future. Leadership exists when people are no longer victims of circumstances but participate in creating new circumstances. When people operate in this domain of generative leadership, day by day, they come to a deepening understanding of ‘how the universe actually works’. That is the real gift of leadership. It’s not about positional power; it’s not about accomplishments; it’s ultimately not even about what we do. Leadership is about creating a domain in which human beings continually deepen their understanding of reality and become more capable of participating in the unfolding of the world. Ultimately, leadership is about creating new realities’ [Joseph Jaworski, Synchronicity: The Inner Path of Leadership]. Now substitute ‘innovation’ for ‘leadership’ and ‘innovators’ for ‘leaders’ in the above:

- Because of our obsessions with how innovators behave and with the interactions of innovators and followers, we forget that in its essence, innovation is about learning how to shape the future.
- Innovation is about creating new realities.

After all the definitions for innovation and innovators we have participated in constructing or have seen, this is by far the most encompassing in terms of complexity. However, it makes out innovators and their contributions to be different and it links leadership and people (more than technology) to innovation. This links innovation to quantum thinking and complexity science.
The Application Of A Resilience Engineering Model For Organisational Innovation
Pellissier, de Sousa

2 THE QUANTUM APPROACH TO MANAGEMENT

Innovation takes place in a quantum environment that is the antithesis of the clock-work based Newtonian world (thesis) we feel safe in. Although it is true that quantum (generally non-linear) thinking and Newtonian (generally linear) thinking are perceived to be on the Hegelian dialect of opposites, we do not believe there should be an either/or here but rather we should beg and borrow from both worlds to find synthesis from both. The principle of dialect is a known one and examples abound:

- In 1852, Riemann created a new geometry that did away with the Euclidean geometry. Riemannian geometry is about space, curved surfaces and multiple dimensions. It is only recently that the significance of his insights have been realised.
- In 1904, Einstein was working on his Theory of Relativity. However, he did not have the mathematical formulae to make his theory work, and it was not until someone showed him Riemann’s theorems that he had the answer. Space-time geometry is Riemannian. Einstein’s theories held until the emergence of Quantum Mechanics in the 1930s.

Let us juxtapose these worlds, and the thinking behind. The graphic below (figure 2) shows the evolution of different worlds not dissimilar to Toffler’s well known epochs (Toffler [65]; Pellissier, [46]). This one, however, is not based on needs and valued-added activity but on thinking and doing. Over time, and as knowledge grows, it takes us though a Plato-est ideas state of ‘true/false’ concepts (or on/off depending on the language we use) that converges into a mechanistic world of rules and laws that govern all decisions and behaviours, diverging into a quantum world of chaos (and order) and complexity. The quantum thinking paradigm is not a bad one as [quantum mechanics] ‘... does not predict a single definite result for an observation. Instead, it predicts a number of different possible outcomes and tells exactly how likely each of these is’. Thus, if one made the same measurement on a large number of similar systems, each of which started off in the same way, one would find that the result of the measurement would be A in a number of cases and B in a different number and so on.’ The reader will see the direct analogy to accepted practices like scenario planning and simulation activities. There are five main caveats of the quantum principle that are relevant to the modern world (although the analogy still needs to be scientifically proven). These are that:

- energy is continuous but comes in small but discrete units;
- the elementary particles behave both like particles and like waves;
- the movement of these particles is inherently random;
- it is physically impossible to know the position and the momentum of a particle at the same time. The more precisely it is known, the less precise the measurement of the other is and
- the atomic world is nothing like the world we are used to. Equilibrium can be upset by some force (external or internal).

These certainly make sense from an innovation (or regeneration) point of view as it presents a space for randomness and uncertainty.
In the figure above, we made the point of the mechanistic worldview converging and/or diverging to the quantum one. There is reason to believe that, although we all maintain to embrace non-linearity, we cling desperately to rules for comfort. Thus, in some unique way (of which the answer may be 44), these worlds should not diverge. Below are the opposing principles in Newtonian and quantum physics. The optimum solution would be to be able to tunnel, i.e. jump from linearity to non-linearity and back, depending upon the need. This of course provides us with far more competencies than before - the only requirement to be so non-linear as to be able to acknowledge a space for linearity!

<table>
<thead>
<tr>
<th>NEWTONIAN PHYSICS</th>
<th>QUANTUM PHYSICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>The dualistic separation of consciousness and matter</td>
<td>The monastic unification of consciousness with matter</td>
</tr>
<tr>
<td>The universe as the motion of inert molar objects</td>
<td>Universes as materialised by conscious participation</td>
</tr>
<tr>
<td>The space between molar objects as flat and empty</td>
<td>Space/time as curved and filled with matter and energy</td>
</tr>
<tr>
<td>The unique existence of only one absolute universe</td>
<td>The existence of many relativistic universes</td>
</tr>
<tr>
<td>The deterministic certainty of inert molar objects</td>
<td>The probabilistic uncertainty among self-motion monads</td>
</tr>
<tr>
<td>The fundamental separation of inert molar objects</td>
<td>The eternal connections between self-motion monads</td>
</tr>
<tr>
<td>The eventual death of the one absolute universe.</td>
<td>The eternal self-organisation of relativistic universes.</td>
</tr>
</tbody>
</table>

Table 1 - Newtonian and quantum physics principles
The analogies to Newtonian and quantum management thinking are be summarised in table 2 below.

<table>
<thead>
<tr>
<th>NEWTONIAN THINKING (ON)</th>
<th>QUANTUM THINKING (ON/OFF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The exclusion of consciousness in the design of formal systems</td>
<td>The inclusion of consciousness in self-designing systems</td>
</tr>
<tr>
<td>The organisation as passive jobholders following official procedures</td>
<td>Organisations as conscious participants in self-designing processes</td>
</tr>
<tr>
<td>The white space between passive jobholders as implicitly ignored</td>
<td>Cross-boundary processes as explicitly addressed and infused with information</td>
</tr>
<tr>
<td>The unconscious administration of a rigidly structured organisation</td>
<td>The conscious self-management of a flexibly designed organisation</td>
</tr>
<tr>
<td>The external control of passive jobholders</td>
<td>The internal commitment of active participants</td>
</tr>
<tr>
<td>The enforced segregation of passive jobholders</td>
<td>The empowered relations among active participants</td>
</tr>
<tr>
<td>The eventual self-destruction of a rigidly structured organisation.</td>
<td>The eternal self-transformation or self-destruction of flexibly designed organisations.</td>
</tr>
</tbody>
</table>

Table 2 - Newtonian and quantum thinking principles

Quantum thinking necessitates newness, like communities of practice; open and creative space; borderlessness; group actualisation; creativity, courage and renewal; innovative leadership; metaphors; appreciative enquiry and discontinuity.

From a modelling perspective, built on our understanding of science it is clear that there is a fundamental shift that needs to take place within organisations. Zohar [73] implies that through thinking differently, one can help create new ideas for organisational design; she states that, ‘The essence of quantum thinking is that it is the thinking that precedes categories, structures, and accepted patterns of thought’ (Zohar [73]). Thus quantum thinking, in line with complex environments, will lead to new organisational theories in design and subsequently leadership definitions to cope with these in organisations.

A Newtonian world view has structured organisational modelling based on the separation of parts within this fundamental science into an unconscious administration of a rigidly structured design that has forced the segregation of organisational parts and participants into divisions and separate structures (Kilman, 2001) [35]. The space between these internal participants and external organisations are implicitly ignored and are only relevant once a reaction to an external stimulus is required. Organisations today are rigid structures, controlled from outside the system and the white space between organisational participants is implicitly ignored (Kilman [35]). Wheatley [71], states that the machine imagery (of Newton’s theories) is captured by organisations in an emphasis on structure and parts and links this to the machine imagery of Newtonian-Cartesian physics: ‘Responsibilities have been organised into functions. People have been organised into roles. Page after page of organisational charts depicts (sic) the workings of the machine’ (Wheatley [71]). This separation of the system has led organisational designers to design organisations and value chains upon principles of parts interacting along a continuum managed through force and reaction, in line with Newton’s laws of motion. Shelton and Darling [57] relate this design directly to Newtonian-Cartesian scientific principles and the management theorists that utilised these principles to design organisations: ‘Newton’s thinking had enormous impact, not only on science, but on organisations as well. The founding fathers of industrialism were greatly influenced by his worldview. Newton frequently characterized the universe as a great clock-like machine and his machine metaphor was transferred to the workplace...Data were (sic) collected and analyzed (reductionism); prediction was highly valued (determinism); and what could not be measured simply did not exist (positivism).’ These structures and categories to design, manage and lead organisations, now seem to have been ‘formulated for life in simpler times when organisations were
viewed as stable entities that functioned in a logical, linear, predictable manner’ (Shelton et al.[57]).

In stark opposition to this philosophy, is the Quantum worldview that through a shift to a systems thinking paradigm, in line with quantum, thermodynamics, complexity and chaos theories, is a shift away from thinking of the parts above to a paradigm of a system as a whole, where fractals will determine new organisational constructs at the edge of chaos, within a system of bounded instability (Senge [54]; Zohar and Marshall [75]), through the interconnectedness of people (Kilman, 2001)[35], aligned to a central vision of ‘meaning’ (Zohar [73]). This is corroborated by Kilman [35], where he suggests that a principle for quantum organisations must be the ‘inclusion of consciousness in self-designing systems, which will lead to the eternal self-transformation of flexibly designed organisations’.

Once guiding principles (guiding visions, sincere values and organisational beliefs) have been agreed, Wheatley [71] highlights two fundamental shifts that organisations need to make to become quantum thinking organisations, that operate in this new structure. The first shift must be a shift to systems thinking and the second shift moves structured organisational dynamics into the organising dynamics of a living system. The future of integrated extended organisations and value chains is inherent in the value or ‘meaning’. Wheatley sees this dynamic organisational system as a relationship of networks. The integration of this dynamic system, writes Wheatley [71], is through the three elements of (1) identity (ability to manage the shapes in motion of an organisation around a central tenant of meaning), (2) information (the integration of organisational learning) and (3) relationships (beyond the traditional boundaries to establish relationships with people anywhere in the system), as supported by Jaworski [32] and Kilman [35], where Kilman states that in quantum organisations, active participants must have empowered and empowering relationships.

The perpetuation of Newtonian thinking has continued to be expressed within the human resources field. Many academics and organisational theorists are of the opinion that management unconsciously administers a rigid inflexible organisation, that employees have been reduced to passive jobholders who are externally controlled and employed to administrate and not to think, which is furthermore structured by the segregation of jobholders into organisational roles and responsibilities (Wheatley [71]; Senge [54]; Zohar [73]; Kilman [35]; Shelton et al.[57]. The publication of Taylor's Principles of Scientific Management in 1911, reinforced the idea that organisations should be managed according to scientific principles. Taylor's management philosophy was directly tied to the prevailing Newtonian worldview, to bring more ‘predictability and control to the management of organisations...at the same time in France, Henri Fayol developed a set of management skills that further reinforced the Newtonian worldview. Fayol's management skills (planning, organizing, commanding, coordinating, and controlling) have been widely used for almost a hundred years’ (Shelton et al.[57];). Zohar [73] contends that a paradigm shift in the way we think is needed before we can change organisational models, and that this thinking is based in the Quantum realm of conscious participation, she contends that: ‘Quantum thinking is the link between the brain’s creativity, organisational transformation and leadership’ (Zohar [73]).

The quantum approach to modelling organisations sees organisational participants involved in continually re-designing processes and systems (Kilman [35]) dynamically controlled through responsibility and alignment of values (Wheatley,[70]), taking into account the changes in the external environment. This paradigm relies on the internal active participation of people. Kilman [35], sees these trusted participants as the enlightened people within current organisations, trusted with the ability to make fundamental decisions on behalf of the organisation. This is aligned to Maslow's 'peak experiences' and therefore implies a spiritual context to personal values and meaning (Zohar, [74]). Kilman [35] further states that, in the past people have been trained and rewarded according to a Newtonian-Cartesian paradigm and have been changed into inert molar objects, controlled by external forces of reinforcement and coercion. Kilman [35] continues that a shift in
thinking to ‘enlightened participants’, will ‘...generate quantum thinking, which is the basis for self-designing, implementing, and improving formal systems and value-added processes. As members enhance their innate capabilities for creativity, collaboration and commitment, they will perpetually transform themselves and their organisations.’ Intervention is required in assisting people to develop their self-aware consciousness to become self-motion monads within the organisational structure.

The basic premise of the economic motive for companies is to make profit, which aligned to Newtonian thinking, has created businesses that are focused on their own individual ‘dimensions’ of profit making, within a larger dimension of separated and fragmented supply chains, driven by shareholder value. Beyond the basic assumptions of capitalism there are unconscious assumptions ‘each agent or corporation is an island unto itself whose actions have no unwanted consequences, and whose interests are under its own self control...without regard to or concern for wider issues’ (Zohar, [74]).

The concern with a deterministic and one dimensional approach to organisational design, is that the values associated become ‘selfish’ (Zohar and Marshall [75]) and lead to a continual search for more and more finite measures within the current system or paradigm leading to learning inefficiencies, as ‘problems are addressed and errors are corrected using only past routines and present policies’ (Robbins, [52]). Werman, [70] states that; ‘all along we had been trying to control the outcomes by forcing artificial structure and measures upon the discrete aspects of the plan’. Wheatley [70] translates Newtonian-Cartesian thinking into deterministic values within an organisation, when she states that ‘In organisations, we are very good at measuring activity. In fact, that is primarily what we do. Fractals suggest the futility of searching for even finer measures that concentrate on separate parts of the system. There is never a satisfying end to this reductionist search, never an end point where we finally know everything about even that one small part of the system’

Values associated with measuring organisations from a quantum world view, suggests that the world operates within multiple dimensions and not linear ones telling us that the parts are not independent but dependent on the whole. All participants are connected therefore and that there is no space between them. Thus, suggesting a movement towards a holistic view of the organisation, within its environment. As such, Heylighen [28] states that a complex system cannot be separated into a set of independent elements; hence a reductionist methodology cannot be applied to the study of complex adaptive systems. To study perturbations to the system in context requires an understanding of sets or patterns within complex environments. The study of thermodynamics and fractal dynamics suggests a systems thinking approach where parts are dependent on the whole (Prigogine, [47]; Kauffman [36]; Mandelbrot in Heylighen [28]). This system can in general be seen as a cyclically closed system of processes (self-organisation). Thus, through understanding the metaphysical aspects, where the shift is to the conscious self-management of a flexibly designed organisation within a multidimensional environment (Kilman [35]), the move away from a pure shareholder value system to an integrated value system, is in line with systems thinking and would include the entire value chain of participants having the same vision, mission, aligned strategies and objectives, and measured by greater set of norms and values.

Fornaciari and Dean [20], state that there is an emerging stream of research exploring aspects in organisations relating to collaboration and participation, and social scientists have started to give attention to religion and spirituality in organisations that opens up organisations to intelligences beyond IQ (intelligence quotient), and EQ (emotional quotient) to SQ (spiritual quotient), that move in dimensions beyond our limits of space and time. We believe this will move us into transcendental territory, which allows us to align both quantum physics (nuclear) and psychology of the unconscious (Jung). On the premise that innovation requires multiple dimensional perceptive thinking to envision new worldviews, SQ has wider implications for innovation within an organisational context,
beyond forming the strange attractor of meaning between individuals and organisational outcomes.

3 COMPLEXITY IN A QUANTUM ENVIRONMENT

There is no denying that the future world of work will be different. There are numerous signs of environmental changes, for instance, (1) politics: the threat of India and China as new world powers, and possible demise of Africa, the war on terror, the issue to secure oil and other resources and human rights violations; (2) economic turbulence caused by globalization and de-regulated markets and new technology; (3) globalisation and de-regulated markets that will remove barriers, increase free trade, create more and more consumer choice (companies can no longer rely on regulations to protect their business, the most significant competitor will not come from current industry, governments are stripped of power and increase the power of the consumer); (4) technology: the decreasing price boom, raw material prices and commodities will fall and the computer and telecommunications industry have provided the platform for e-commerce; (5) the knowledge economy: primary resources have become far more intangible and difficult to contain. Knowledge and information have no value until it is used for a specific purpose; (6) global warming and sustainability [76]; international terrorism and piracy.

Greene [20] believes that [quantum mechanics] is the ‘The framework of laws governing the universe whose unfamiliar features such as uncertainty, quantum fluctuations, and wave-particle duality become most apparent on the microscopic scales of atoms and sub-nuclear particles.’, while Hawking [26] famously proposed that quantum mechanics does not predict a single definite result for an observation. Instead, it predicts a number of different possible outcomes & tells how likely each of these is. Thus, if one made the same measurement on a large number of similar systems, each of which started off in the same way, one would find that the result of the measurement would be A in a number of cases and B in a different number and so on.’

We agree with Lewin & Regime [39] who state that the science of complexity has to do with structure and order and not random chaos, especially in living systems such as social organisations, the development of the embryo, patterns of evolution, eco systems, business and nonprofit organisations and their interactions with the technological environment. It provides a more accurate view of reality. The role of complexity science (and complex adaptive systems) in business is well explained (Elliott & Kiel [19] and Shaw [58]). Simply defined, complex adaptive systems are composed of agents that interact with each other and, in doing so, generate new behaviours for the systems as a whole (Lewin & Regime [39]). Moreover, the patterns of behaviour in these systems are not constant since, as the system’s environment changes, so does the behaviour of its agents and, as a result, so does the behaviour of the system as a whole. Complexity science is the study of complex adaptive systems. A complex system is a system having multiple interacting components, of which the overall behavior cannot be inferred simply from the behavior of components. Complexity science spans scales from particle fields to information mechanics (physical analysis of the dynamics of information transmission) and adaptive systems (learning and consciousness, including neural systems), to human society, ecosystems and extraterrestrial space. These phenomena all share the qualities of a self-organizing network. From their study, new methodologies and concepts of the nature of reality have emerged. In international relations, the emergence of an interconnected global civilization manifests this sort of complexity. In knowledge creation, so do the cross-fertilization and merging of academic specializations into ever newer and more numerous interdisciplinary subfields (Cutler [13]).

There are five primary principles to operate in a complex reality [39]. These are: (1) agents interact and mutually affect each other in a system: This focuses on relationships between and among people, teams and companies, (2) agents’ behaviours in a system are governed by a few simple rules: In business, rules become practices. These practices are
The Application Of A Resilience Engineering Model For Organisational Innovation
Pellissier, de Sousa

guided by shared values and beliefs (3) small changes can lead to large effects, taking the system to a new attractor: Multiple experimentation on a small scale is the most productive way to lead change rather than to attempt to leap too quickly to a perceived desired goal on a large scale (4) emergence is certain, but there is no certainty as to what it will be: Create conditions for constructive emergence rather than trying to plan a strategic goal in detail. This includes nurturing the formation of teams and creativity within teams and evolving solutions to problems (not designing them). Hierarchical and central control should give way to distributed influence and a flat organisational structure. (5) the greater the diversity of agents in a system, the richer the emergent patterns: Seek diversity of people in terms of culture, expertise, age, personalities and gender, so that people interact in teams (thus creativity has the potential to be enhanced).

When we relate business to a complex adaptive system (also called a learning system (Robb [50]), we look for ways to successfully adapt to changing environmental conditions. Complexity science focuses on relationships between individuals, teams or between organisations and businesses. Accepting business as being a complex system requires that we acknowledge that we cannot control organisations to the degree that a mechanistic perspective will imply but only that we can influence where the organisation is going and how it will evolve. From this view, organisations are complex adaptive systems nested in larger complex adaptive systems (for instance the economy or the country it is based in, or the industry it operates in). Lastly, complexity science allows an organic view of organisations and its resources. Resilient organisational structures, in focusing on the skills, culture and architecture, address this matter and will be discussed in a separate section.

Complexity allows a two tiered focus in business: (1) its performance system (responsible for the performance of current goals and tasks for immediate survival), and (2) its adaptation system which is responsible for the long-term sustainability through the generation of new ideas, operations and behaviours. It generates possible futures for the total systems. Successful resilient organisations should be robust in terms of both subsystems but tend to concentrate on only one (Robb [50]).

4 STRUCTURE AND LINEARITY IN MANAGEMENT

Twenty first century enterprises functioned in a relatively simple, stable and predictable environment. As a result, managers were able to make decisions based on intuition or by repeating procedures successfully used by other executives or used in the past. These methods did not always approach the problem systematically and did little to improve or advance the managerial decision making process. However, early examples of systematic approaches do exist. For example, in the fifteenth century, Venetian shipbuilders used an assembly line of sorts to outfit their vessels. In 1776, Adam Smith suggested a division of labour based on his analysis of straight pin manufacturing. And, in 1832, Charles Babbage, presented a number of concepts of industrial engineering, including a skill differential in wages. In the late 1800s, an American engineer named Frederick Taylor formally advocated the scientific approach to manufacturing. Using time studies, he analysed work methods, established standards and evaluated work performance. He believed there was one best or most efficient way to accomplish a task. Henry Gantt, a contemporary of Taylor’s, extended these concepts by including human behavioural factors. In summary, these early scientific approaches were limited to establishing or improving efficient performances for specific tasks in the lower levels of organisations.

World War II created unprecedented problems in resource allocation, production planning and scheduling, inventory and quality control, transportation and logistics. Past procedures proved inadequate and the leaders recognised the need for analytical reasoning. Multidisciplinary teams (consisting of physical scientists, engineers, mathematicians and military leaders) were formed to study the problems and make the necessary recommendations. Since mathematics is the language of the scientific disciplines, the
The Application Of A Resilience Engineering Model For Organisational Innovation

Pellissier, de Sousa

Problems were formulated mathematically. Industrial applications spread in the 1950s with computer technology developing. Advances in computer technology provided the means to solve the more sophisticated mathematical tools. More so, sophisticated software programs were developed for standard techniques. The advent of computer based management information systems helped some of the data required by the procedures.

It is important to view any organism or phenomenon in terms of its parts and as a whole. This greatly assists in managing, controlling and optimising the entirety and its constituent processes. When attempting to sell the quantitative disciplines, most analysts do not consider political and monetary factors, or the personalities of the managers. Moreover, it is important to remember that an organisation is a collection of interrelated parts (divisions, departments, machinery, people and so on) intended to accomplish specific objectives. As such, it constitutes a system. Hence, a decision in one part of the organisation could significantly affect the operations of other segments. For example, an inventory problem within a firm’s production department could disrupt marketing, finance, accounting and personnel functions. This calls for the systems approach, wherein a problem should be examined from the overall organisational point of view. This provides the key characteristic of management scientists. Furthermore the organisation exists within the environment in which it operates. Hence the firm’s actions may affect market, social and political conditions. Likewise, actions by unions, consumers, competing firms and the government affect the organisation’s operations. Incorporation of these environmental factors into the analysis should hence be part of the systems approach.

Peter Senge’s now classic, *The Fifth Discipline* [45], provides a great understanding in terms of moving from individual parts to the sum of all parts to create total synergy. He maintains that a systems approach is the result of individual personal mastery (he says: ‘I am my position’, i.e. to each of us rests the responsibility to take ownership of the part in society or in any structure that is allocated to us). If we all do this within the same entity or organisation, there will be shared vision, shared meaning and eventually, a systems approach where the sum of the past is larger than the whole. Turning this upside down, it also means that changes in individual elements of a system, will influence the total system. We will return to systems thinking in the last chapter. Logical, is it not? Until we remind ourselves how often leadership or individuals change in an organisation, but that same organisations still try to maintain its same values, beliefs and culture. If the individual elements change, the sum of the parts will be bigger than the whole, but a different whole. Make sense?

The past two decades has seen a radical shift in the basic foundations of how business is conducted. The globalization of markets and production resulted in national markets being integrated into a single global market trading in global products. The shift has been strengthened through the declining of trade barriers and fundamental developments in communication, information and transportation technologies. Globalization resulted in greater world output, foreign investment, greater imports and exports and immense competitive pressures both between nations and industries (Hitt, Ireland & Hoskisson [29]; Pearce & Robinson [45]). Advancements in information technology and related developments in communication technology has increased organisations’ ability to link global operations into sophisticated information networks, shrinking the time in which information is collected and enabling organisations to achieve tight coordination in worldwide operations (Hitt, Ireland & Hoskisson [29]; Laudon and Laudon [38] and Pearce & Robinson [45]). External factors influence the organisations’ direction, organisational structure and internal processes. These factors that exist in the organisations’ remote, industry and operating environments require constant monitoring for the formulation of strategies to optimise the organisations market opportunities and threats to allow them to survive in their competitive environment (Pearce and Robinson [45]).

All actions taken by organisations are intended to allow them to achieve strategic competitiveness and earn the organisation above-average returns. Strategic Competitiveness is achieved by an organisation when it successfully formulates and
The Application Of A Resilience Engineering Model For Organisational Innovation

Pellissier, de Sousa

implements strategy which creates value. If competitors are unable to duplicate, or find the strategy too costly to imitate, the organisation has achieved a sustainable competitive advantage (Hitt, et al. [29]). Hitt et al. [29] are of the opinion that the fundamental nature of competition in many of the world’s industries are changing, and the pace of this change is relentless and ever increasing. Conventional sources of competitive advantage are no longer effective, requiring managers to change their traditional mindset and adapt a new mindset which values flexibility, speed, innovation, integration, and the challenges that evolve from the constantly changing conditions (Hitt et al.[29]).

Organisations are aware that the rigid seldom repeated review of organisational strategy is no longer feasible to achieve competitive advantage. Ideas and mindsets that were prominent five years ago could now be archaic, requiring a rethink of organisational strategy on a more frequent basis. The pace of change and reform has increased the speed at which organisations are required to define strategy, but has not removed the need for a structured strategic management process. Strategic Management is an integral part of the running of an organisation, but require strategic decisions to be based on valid, and actionable intelligence, rather than the ideals, on which the organisations visions are based.

In the modern management environment, there are numerous examples of our focus on relationships and networking. Social networking sites like Twitter, Facebook, LinkedIn and Google Buzz; and search engines like Google, Bing, and the new natural language processing focusing on the comprehensiveness, accessibility, relevance and speed of information. The standard management theories are mostly based on a linear environment and provide linear solutions for the linear problems identified. For example, Table 3 summarises the extant management models indicating their application and evolution.
<table>
<thead>
<tr>
<th>PERIOD</th>
<th>FOCUS</th>
<th>CONTRIBUTORS</th>
<th>ENVIRONMENT</th>
</tr>
</thead>
</table>

Table 3 - Summary of evolution of the management models
5 LINEAR DECISION-MAKING PROCESS IN MANAGEMENT

Simon [55] originally identified a systematic decision-making process involving three major phases: intelligence, design, and choice. Simon later included an implementation phase, while Turban, Aronson, Liang & Sharda [67] suggest a fifth phase which includes monitoring and feedback (Turban, et al., 2007). Robbins & De Cenzo [52] identify the steps within the decision-making process as including identifying the problem, selecting a solution, and evaluating the effectiveness of the solution. They further identify a single problem as a discrepancy between two states, the existing state and the desired state of affairs. Furthermore, once the problem has been defined, decision criteria are to be identified. These decision criteria are factors that are relevant in the decision making process and could eliminate certain courses of action. The criteria are then weighted and the best outcome selected and implemented. The final step in their model includes the evaluation of the results to appraise whether or not their decision corrected the problem (Robbins & DeCenzo [52]). Simon’s model is seen to be the most concise and complete characterisation of rational decision-making. The model includes a continuous flow of activity originating at the intelligence phase moving through design to choice to implementation while a return to the previous phase is possible (Turban, et al.[67]).

![Image of the decision-making process]

Source: Turban et al. [67]

Figure 3 - The decision-making process

The decision-making process begins with the intelligence phase, during which the decision-maker examines reality which identifies and defines problems. During the design phase, a model that represents the system is constructed by making assumptions which simplify reality, and identify of all relationships among the variables. The proposed model is
The Application Of A Resilience Engineering Model For Organisational Innovation
Pellissier, de Sousa

validated and criteria evaluated which leads to alternative sources of action. The choice phase includes the selection of a proposed solution to the model, which is then tested to determine its validity (Turban, et al.[67]).

A major characteristic of decision-making is the use of a model, on which to analyse the effects of a decision on a model of reality rather than on the real system. Turban et al. [67] define a model as ‘a simplified representation or abstraction of reality.’ Models represent systems or problems utilising different degrees of abstraction. Models are classified based on their degree of abstraction (Turban et al.[67]).

- Iconic (scale) models are the least abstract models. They are physical replicas of the system, but on a smaller scale than the original. An iconic model can be either three dimensional, such as a model of a bridge, or two dimensional, like a photograph.
- Analogue models act like the real system but do not look like it. It is a symbolic representation of reality, and is most often two dimensional charts or diagrams. Examples include blue prints, organisational charts, animations, videos and movies.
- Mental models are descriptive representations of decision-making situations that people form in their heads and think about. The individual’s thought process works through the different scenarios to consider the feasibility and risks involved in each potential alternative.
- Mathematical models are utilised due to the complexity of relationships in organised systems which cannot be represented by iconic or analogue models.

The advances in computer technology have increased the use of both iconic and analogue models in conjunction with mathematical modelling. Combined, the models allow for easy manipulation, the compression of time, reduced cost, inclusion of uncertainty factors, enhance learning and training, most commonly through a web interface (Turban, et al.[67]). When the proposed solution is found to be viable, the process moves onto the implementation phase. Successful implementation would result in the solving of the real problem, while failure will lead to a return to the earlier phases of the process (Turban, et al.[67]). While the understanding of the decision-making process and the models that can be generated are critical components of decision-making within the management function, the foundation underlying organisational decision-making is the acquisition and use of information, intelligence and knowledge.

6 RESILIENCE ENGINEERING IN A COMPLEX SYSTEM

The concept of resilience has reached maturity over the past decade. Robb [50] defines a resilient organisation as one able to sustain competitive advantage through its capability to (1) deliver excellent performance against current goals, whilst, in paradox, (2) effectively innovating and adapting to rapid, turbulent changes in the environment. The first requires consistency, efficiency, elimination of waste and maximizing short-term results, whilst the second requires foresight, innovation, experimentation and improvisation, with an eye on long-term benefits (Johnson-Lenz [33]). The two modes require different skills sets and organisational designs (for example, move from JIT production to ‘just-in-case’ resilience). These organisations exhibit particular characteristics in the sense that they (1) can create structure and dissolve it; (2) provide safety in the face of change (although this is not necessarily security or stability); (3) manage the emotional consequences of continuous transformation, change, anxiety and grief; and, (4) learns, develops and grows. The resilience community agrees that resilience architecting (also called resilience engineering) occurs over the three phases of a disruption. In the pre-disruption phase the system should take steps to anticipate the disruption and avoid the disruption, if possible. In the survival phase the system should absorb the disruption so that it can recover in the recovery phase. In the recovery phase the system resumes some degree of its original goals, including the survival of the humans in it. Disruptions are the initiating event that may lead to a
The Application Of A Resilience Engineering Model For Organisational Innovation

Pellissier, de Sousa

catastrophic event. Human error is a common source of disruption. However, the resilience of the entire system will determine whether the system is prone to catastrophe. Disruptions may be either external, such as terrorist attacks or natural disasters, or they may be internal, such as human or software errors. The phenomenon in which systems fail when the components function as designed is discussed.

Resilience has four primary attributes: capacity, flexibility, tolerance, and inter-element collaboration. Capacity requires that the system be sized to handle the maximum and most likely events, such as terrorist attacks and natural disasters. However, a system cannot depend on capacity alone; the other attributes must be present to handle unpredicted events. Capacity includes functional redundancy. Flexibility requires the system to be able to reorganize. For example, plans must be in place to allow the command and control to shift upwards in the event of a serious disruption, such as a terrorist attack. Tolerance allows the system to degrade gracefully in the face of an attack. That is, all resources would not become inoperative after the first strike.

One of the most important resilience attributes is inter-element collaboration. This attribute allows all elements of the system to interact and cooperate with each other as in collaborative innovation systems.

We like Johnson-Lenz's [33] list of activities relating to resilient organisations. These are:

i. **Resilient organisations actively attend to their environments.** Monitoring internal and external indicators of change is a means of identifying disruptions in advance. Resilient organisations seek out potentially disturbing information and test it against current assumptions and mental models. They work to detect the unexpected so they can respond quickly enough to exploit opportunity or prevent irreversible damage. In short, they anticipate being prepared.

ii. **Resilient organisations prepare themselves and their employees for disruptions.** Attentive preparations build a team that imagines possibilities and displays inventiveness in solving problems. Managers know how and when to allow employees to manage them for focused productivity as well as adaptive innovation. Resilient organisations cross-train employees in multiple skills and functions. They know that when people are under pressure, they tend to revert to their most habitual ways of responding.

iii. **Resilient organisations build in flexibility.** Even while executing for lean and mean performance, resilient organisations build in cushions against disruptions. The most obvious approach is the development of redundant systems - backup capacity, larger inventories, higher staffing levels, financial reserves, and the like. But those are costly and not always efficient. Flexibility is a better approach.

iv. Engaging suppliers and their networks in devising makeshift solutions to temporary disruptions is a flexibility strategy. So are policies that encourage flexibility in when and where work is done. Employees who are used to telework and virtual workspaces adapt more quickly and are more productive following a crisis. In addition, research shows that flexible work practices contribute to greater employee resilience, productivity, and commitment, and to lower levels of stress.

v. **Resilient organisations strengthen and extend their communications networks - internally and externally.** A robust and redundant communications infrastructure holds up in a crisis. Social networks among employees at resilient organisations are rich, varied, and visible. People who have trust relationships and personal support systems at work and with friends and family are much more able to cope with stress and change. Good connections and communications also apply to external relationships with suppliers and customers. A key is to recognize what’s important to meet organisational goals and to listen to those with needed expertise and ideas wherever they are in the value web. Resilient organisations use networked communications to distribute decision-making. As much as possible, they push decisions down to where
they can be made most effectively and thus quickly. This in turn requires good access to information at all levels of the organisation.

vi. Resilient organisations encourage innovation and experimentation. In times of great uncertainty and unpredictability, the success and failure of small-scale experiments can help map a path to the future. Resilient organisations engage in market research, product development, and ongoing operations and service improvements. They invest in small experiments and product trials that carry low costs of failure.

vii. Resilient organisations foster a culture of continuous innovation and ingenuity to solve problems and adapt to challenges. A side benefit is that employees who believe they can influence events that affect their work and lives are more likely to be engaged, committed, and act in positive ways associated with resilience. Some organisations also have internal idea markets to surface new ideas and innovations, for example, ‘crowdsourcing’ to engage people externally in solving a given problem or Eli Lilly’s Innocentive Open Innovation Marketplace.

viii. Resilient organisations cultivate a culture with clearly shared purpose and values. When an organisation’s sense of purpose is shared by its employees, suppliers and customers, those networks can provide flexibility to help it through a disruption. Engaged employees will seek out opportunities to try new approaches, find creative solutions, and achieve great results.

To say that resilience is well-defined is not say that the work is finished. There is still much to be learned to accomplish the goals of resilience. Areas of major research include, but are not limited to, the application of resilience in both political and economic environments.

7 MAKING ORGANISATIONAL AND MANAGEMENT INNOVATION POSSIBLE

Increasing global instability and competitiveness impose the compelling need to identify new paradigms, methods, applications and technologies to support renewal and creativity. In order to operationalize innovation, organisms (e.g. organisation, entity or individual) should behave differently. The right hand column (table 2) embraces newness, renewal and develops a consciousness for them. There is more than talk and planning. There is energy and dialogue. The consequences are that the need for change unavoidably grows. Thus, the fundamental nature and soul of the organic organism should be transformed to make change the norm and accept this (change) as practice. These organisations could typically behave according to the following patterns: (1) employees collectively determine the direction and subsequently empower leadership to point the way, (2) leadership could help (collaborative) teams to realise they are off course and assist in the realignment within the whole, (3) frontline employees would be responsible for the movement of the organisation, not the managers and relationships become flexible and quickly adaptable. Thus innovation needs to happen at an organisational/management/technological level within a growingly unstable and unpredictable environment.
However we may make out that modern organic organisms (see how we defined them above) operate like this, the truth is that they remain hierarchical as the only way to retain control/order. This inhibits the creative space for organisational and management innovation as innovation takes place in an atmosphere of quantum thinking and thus requires resilience.

Organisations at the edge of chaos (i.e. where there is bounded instability), is the place for creativity and change, leading to innovation within organisations. This takes fundamental trust in the thinking ability of ‘self-motion monads’ (people within an organisational context). Such a thinking ability takes into account leaders as ‘being’. Jung maintained that attaining balance within the psyche established wholeness within the individual, he further stated that, ‘individuation means becoming a single, homogeneous (sic) being.’ (1959:181 in De Charon [16]). This has a correlation to the quantum skills model of Shelton et al. (1999), where the quantum skills are developed around a central tenant of being and supported by Kilman [35] and Wheatley [71], which is based on a chaos theory metaphor of fractals forming around a strange attractor, this strange attractor in an organisational sense, as per Wheatley (1999), is meaning or in this context ‘being’ and is the central link between the intelligence types of IQ, EQ and SQ as per Zohar & Marshall’s [74] work based on Maslow’s hierarchy of needs. This finding supports a systems view of the world as per a quantum paradigm and a quantum leadership propositions. Such management will lead organisational participants in developing individual creativity utilising their knowledge and technology and linking of these to add greater value to the world around them, thus becoming self-transcendent and ultimately driving dynamic innovative strategies in line with a continually changing environment.

Resilience engineering (as in figure 4) allows for two juxtaposed views of management and thereby allows for (1) the generation of conceptualisation (alternative ways to understand and define a problem or opportunity), (2) optimisation (alternative ways to get an idea to work in practice and uncovering all factors to successful implementation), (3) generation (new opportunities and problems that can be capitalized on) and (4) implementation (new actions and results to gain acceptance of a new idea) of options in order to implement
organisational/management/technological innovation. While this is being done, we will think about the really important question of what lies beyond innovation (other than more real innovation) if anything. And we think it might have to do with creativity and resilience in order to become globally competitive.

That great source of wisdom (Douglas Adams [01]) said that

‘There is a theory that states that if ever anyone discovers exactly what the Universe is for and why it is here, it will instantly disappear and be replaced by something even more bizarre and inexplicable. There is another theory that states that that has already happened.’

REFERENCES


http://www.provenmodels.com/571/transilience-maps/k.b.-clark--w.j.-abernathy


The Application Of A Resilience Engineering Model For Organisational Innovation
Pellissier, de Sousa


THE EFFECT OF TIME VARIATIONS IN ASSEMBLY LINE BALANCING: LESSONS LEARNED IN THE CLOTHING INDUSTRY IN SOUTH AFRICA

K. Ramdass 1* and D. Kruger 2

1 Department of Fashion
University of Johannesburg, South Africa
kramdass@uj.ac.za

2 Faculty of Management Sciences
Tshwane University of Technology, South Africa
davidkruger@gmail.com

ABSTRACT

South African clothing sector has undergone large-scale restructuring over the past 15 years. The global economy is pressurizing organisations to improve productivity of their business processes. Competition is forcing organisations to focus their energy on “core competencies.” Like many industries, the clothing industry is witnessing changes in technology, diversification of labour, managerial implications while competing on the global market. The South African clothing and textile industries have the potential to create jobs, but this potential has been steadily diminishing over the last decade. In this context, the clothing industry is regarded as a powerful engine for economic and employment growth. Nevertheless, the performance of the clothing industry, whether in terms of efficiency, working conditions or degree of social protection, is unstable. This paper aims to highlight some of the problems experienced by manufacturers’ through a semi-structured questionnaire and provide suggestions for improvement of the clothing industry through the application of line balancing as a means of productivity improvement. The research methodology employed in this paper is qualitative and exploratory in nature making use of applicable literature and appropriate case studies [10].

OPSOMMING

Die Suid-Afrikaanse klerasievervaardigings sektor het grootskaalse verandering ondergaan die afgelope 15 jaar. Die wereld ekonomie plaas verhoogde druk op organisasies om werker produktiwiteit te verhoog en werk prosesse te verbeter. Wedywerende organisasies moet hulle energie fokus op die organisie se kern bevoegdhede. Die klerasievervaardigings sektor aanskou tans tegnologiese veranderinge, diversifikasie van arbeid, bestuurstyl veranderinge terwyl hulle terselfdse tyd moet meding op werelde markte. Die sektor beskik oor die vermoee om grootskaalse werkskepping geleenthede te vestig. Ongelukkig het die werkskepping vermoee afgeneem gedurende die laaste dekade. Die sektor illustreer ‘n onstabile groei en is as ‘n geheel oneffektief in werkwyse en produktiwiteit vergelykend teenoor direkte mededingers van ander lande. Die referaat sal poog deur gebruik te maak van lyn balanceering te die sektor mededingend te maak. Kwalitatiewe en ondersoekende navorsings metodologie is gebruik met die opstelling van die referaat.
1. INTRODUCTION AND BACKGROUND

Line balancing is the distribution of work on the line in such a way that everyone gets the same amount of work in terms of time. With the clothing industry being labour intensive, machinists are pressurised for production output throughout the day. All workers should be productive during their working time and production should flow smoothly, thereby achieving planned targets. The objective is to accurately delegate workers to the various operations required to complete the product based on their skills and experience so as to achieve the highest level of productivity and delivery as per the planned targets. The standard minute value is generally used in the clothing industry as a predictor of sewing speed and production efficiency. The standard minute value that is derived through the application of work study methodology is generally assumed as a constant for line balancing. However, a lot of factors cause variations on operational time of the same task such as the fabrics and sub-materials, performance of the machinery, working environment and quality level of the product. With the aid of an illustrating example selected from a men’s shirt manufacturing factory, the effect of time variations for assembly line balancing has been studied in this paper.

The SA clothing industry shed over 67 000 jobs in the past three years [4]. There is a possibility that more jobs may be shed in Durban (Kwa-Zulu Natal). The Alexander Report mentions that between July 2006 and May 2007 there was a drop of 5275 in employment figures [18]. The cut, make and trim (CMT) industries find it difficult to negotiate wage increases as production costs escalate beyond proportion. If the lay-off of workers continues in the clothing industry, it would increase the unemployment rate, thus impacting on the economy of South Africa[17].

The problems faced by clothing producers are summarised below:

- “Throughput time per unit: The time to complete one unit of production vs. the time to complete the order.
- Inventory between processes: used as buffer stock, machine breakdowns, line balancing. It may extend completion time.
- Critical paths: need to be arranged in parallel or in serial. This could affect the production process.
- Bottleneck operations: need identification for action.
- Plant utilisation: estimates effective use of resources.
- Minimum order size: affects the production process in terms of set-up costs and has a bearing on cost per unit.
- Change-over/setup: the loss of production when there is a style change.
- Rejects and repairs: evaluates capacity lost through repairs and costs lost through rejects” [1;11].

2. LITERATURE REVIEW

The multitude of competitive priorities has been the subject of considerable argument by manufacturers around the world. A universal set of priorities does not exist for all the firms in the global marketplace. Porter [19] formalized the concept of competitive priorities into four different types namely, price; flexibility; quality and dependability. Ward et al [24] identified five different dimensions namely price; quality; dependability; product flexibility; and volume flexibility.

Hill [14] argued that a firm should identify those criteria or priorities that win orders against the competition in the marketplace. His “order-winning” criteria included price, delivery, quality, product design and variety. Hill also considered that “qualifying” criteria (or performance criteria) were also important for an organisation.

Glock and Kunz [13] put forth a more detailed list by differentiating four different aspects namely, cost; quality; time; and flexibility.
In the South African clothing industry context, competitive strategies that require implementation as a matter of urgency from the researcher’s perspective would be leadership that is people orientated with an insight into cultural diversity, financial management by understanding organisational costs, service delivery with quality management imperatives and last but not least, performance management initiatives [15].

The clothing supply chain continues to influence industrialisation throughout the developing countries [15]. Thus, consumer industries are struggling to maintain their share of the global market in terms of fashion design, production and distribution throughout the supply chain [8].

Stevenson [23] mentioned that an organisation can compete on three major issues namely:
- cost leadership - competing on the basis of price;
- differentiation - a unique product that is valued by customers; and
- focus - a niche market for its products developed on cost and differentiation strategy.

The use of technology in the clothing and textile industry is always a major factor in its competitive struggle. It is controlled by the established industries in Japan, Europe and the USA. However, technology is not the only way to achieve competitive advantage, but the understanding of systems methodology based on cost and productivity [16]. Information technology also plays a major role in processing information using real time technology, thus facilitating communication and feedback by the touch of a button. The information technology explosion is shaping business operations in all its contexts [12]. Although line balancing could be done manually, it would be advantageous to implement line balancing methodology using appropriate software.

The role of the supervisor is to ensure that the garment breakdown into the various operations is distributed as evenly as possible between workstations. The determination of the standard time value for each task is a critical component in the line balancing process which is obtained through the application of work study principles. However, this may be done using pre-determined motion time standards, which may oversee factors such as the property of the fabrics being used, performance of machinery, the effect of the environment and the quality level of the product may influence task times. There are numerous process improvement methodologies that could improve the status quo in the clothing industry. Qualitative results of the implementation of line balancing and its relative experiences at a South African clothing manufacturer are presented through a case history.

3. SURVEY EVIDENCE THROUGH CASE STUDY APPLICATION - (SELECTED VARIABLES FOR THE PURPOSE OF THIS STUDY)

3.1 DESCRIPTION OF THE INDUSTRY

The clothing industry is dispersed throughout South Africa, but is condensed in the provinces of KwaZulu-Natal and the Western Cape. In Kwa-Zulu Natal, the clothing and textile facilities are found in and around the Durban Metropolitan area with pockets in the north and south coasts, Newcastle, Ladysmith, and Qwa-Qwa. In the Durban metropolitan region (only) there are about 300 organisations employing over 15000 workers. The major part of the South African clothing industry is dominated by small and average sized organisations, i.e. those employing less than 200 employees which make up the majority of CMT manufacturers.

According to the National Bargaining Council statistics, it is estimated that there is approximately 827 formal (registered) clothing companies in South Africa. The clothing and textile industry in South Africa is predominantly South African owned. There are a substantial number of foreign-owned organisations which are located in non-metropolitan areas. The majority of CMT manufacturers, as well as micro-manufacturers receive their customer orders from the larger independent organisations.
Many industrialists from Chinese origin established clothing and textile operations in areas such as Hammarsdale, Mooi-river, Ladysmith, Isithebe and Newcastle. Clothing manufacturers appear to be medium-sized and focused on exclusive outer-wear for the middle to upper-end of the market. These manufacturers sometimes outsource their production to CMT (cut, make, trim) operations. The investigation focused on a limited number of “larger” organisations as these organisations restructured their facilities totally.

3.2 CLOSURES OF CLOTHING ORGANISATIONS

Clothing and textile organisations in South Africa that are unable to cope with the changes in industry sought closure as a resolution. According to the Inggs [15] approximately thirty clothing and textiles companies closed since July 2002. Many of these organisations sought legal advice on the process of liquidation. Closures are wide ranging, from textile manufacturers that produce standardised products to clothing manufacturers who focus on the fashion trends of the industry. Chinese imports proved to be intimidating force for the entire clothing value chain.

3.3 PARTIAL CLOSURES OF ORGANISATIONS

Organisations that are involved with standard products found the competition stifling. These organisations (especially textiles) are pressured to shut certain parts of the operations only. A very good example is Coats South Africa that closed its spinning and twisting departments. It imports yarn ready for the dyeing process from sister companies around the world. Such organisations streamlined the production process to focus on specialised products [2].

Organisations that focused on niche markets managed to become highly profitable. The National party supported a textile manufacturer in the production of parachute fabric manufacture and material for bullet-proof vests and airbags. Other organisations that maintain their competitive edge are those that supply the paper/pulp industry and the mining industry.

3.4 SOUTH AFRICAN ORGANISATIONS FOREIGN TAKE-OVER

Trade liberalisation prompted organisations to focus on its competitive abilities. In this regard an investment in the latest technology and human assets is required. Foreign investment is imperative to bolster the industry in its pursuit of becoming competitive. The latest example is Edcon, which was taken over by the Anglo-American group in April 2007. Approximately 30 percent of the clothing and textile firms in SA are under foreign ownership as organisations strained under competitive pressure [7].

3.5 RELOCATIONS OF SOUTH AFRICAN ORGANISATIONS

Many clothing and textile manufacturers, South African clothing industries (SACI) as an example, relocated to industrial districts such as Isithebe, Madadeni and Ladysmith. The organisations in urban areas are required to comply with the Bargaining Council regulations and unionised labour, thus escalating the cost of production. Rural areas constituted discretionary wage rates as determined by the employer and minimal union interference. The union has an important role to play in the South African clothing industry but stakeholders need to apply themselves in a harmonious relationship.

Many clothing and textile manufacturers abandoned their South African operations to set up facilities in neighbouring countries such as Lesotho, Swaziland and Botswana. The Botswana government provides incentives for clothing and textile manufacturers to invest in the country. In addition to this, there is skilled labour at approximately half the price of SA. These manufacturers are allowed to import raw materials according to the AGOA agreement [2].
The Botswana government provided incentives to draw organisations into the country. The major incentive is the payment of approximately 80% of the payroll costs for the first three years. Botswana provides excellent investment opportunities where the clothing industry could thrive. Taxes imposed on clothing manufacturers are much less than those imposed in SA. One textile manufacturer that set up its facilities in Botswana was Waverly Blankets, which relocated in June 2001.

3.6 PERFORMANCE OF THE INDUSTRY

The clothing industry in SA is performing in the region of 85% labour efficiency. Capital expenditure on new assets averaged approximately 1.4% of sales (Robbins, 2004). With clothing manufacture being labour intensive it is important to advance through the investment in capital requirement and technological innovations to overcome the competitive nature of the industry. Exports decreased, while imports increased. Labour efficiency of the SA clothing operators is approximately on par with the United Kingdom, European Union and the United States but it is the application of technology, working culture and productivity levels that are a matter of concern [3].

3.7 PERFORMANCE MANAGEMENT

Management of all organisations claimed that they had the skills that would improve their production levels. An amazing fact is that 100% of the organisations sampled did not measure their productivity levels, as they focus on labour efficiency. “We do not need training and development,” was mentioned by (5%) of the sampled organisations’ management. The fundamental aspects of work-study, production management and quality management are not implemented by CMT manufacturers. Thus, it can be presumed that output performance could be improved by approximately 30% by CMT manufacturers.

3.8 GARMENT COST BREAKDOWN

CMT manufacturers did not break down costs of the garment from a manufacturing and financial perspective. Garment costing provides information such as labour cost, the number of machinists required, line balancing requirements and production output. With a structured costing system and the estimation of production costs, the organisation would be able to determine delivery times, output performance and profitability of the organisation.

3.9 PLANT HISTORY AND CONTEXT

A qualitative approach using a case study is used in the implementation of line balancing methodology. This production facility manufactures men’s and ladies fashion wear and operates in a small town in Kwa-Zulu Natal. Currently, approximately 300 people work in the plant. The factory opened in 1970 and did not implement modern technology due to financial constraints. The facility had 16 supervisors and a plant manager.

The plant manager agreed to perform a pilot project on line balancing to determine its effectiveness. The sewing department used to the bundle system of manufacture. Work is passed to sewing machines in bundles of cut pieces. The number of cut parts in the bundle may vary according to weight or the complexity of operations required, but the principle remains the same: the operator unties the bundles, sews the cut parts together, re-ties the bundle, processes the work ticket and places the bundle into a bin or on a transporter system (a U shaped manual conveyor). The bundle then goes to another machinist who repeats a similar sequence; a bundle may be tied and untied several times before it completes its lengthy journey. Units move from operator to operator for completion of the respective operation. The bundle production system is a prominent production system used in the clothing industry. Manufacturers use it as a “buffer feeder” and fail to implement process improvement techniques to enhance production flow. Bundles of work-in-progress are found at workstations and sub-assemblies[5;22]].
The plant followed conventional management practices before the 1980’s but started changing as new management came on board. There was very little interaction among employees and management and there was an autocratic style of management.

Before the year 2000 the production facility was accustomed to lot sizes of between 2000 and 10000 units per order. Currently, there are lot sizes of approximately 100 units per order. The garments were not as complex in construction as the ones received currently. The factory was “flooded” with high lots of work-in-process throughout the plant. Employees who were loyal and employed for the last 30 years said that the environment in which they worked was hostile and they did the same operation for several years. It is important to note that there was no process improvement methodologies implemented at this organisation.

4. IMPLEMENTATION OF LINE BALANCING

4.1 LEADERSHIP QUALITIES

The workforce of the majority of the organisations sampled complained that management did not treat them as “assets” of the organisation. They claimed that they are often treated poorly and management would not consider their views on issues. Labour relations are considered “sub-standard” as management regard workers as another “input” for production. Workers mentioned that all management is concerned about is production, and didn’t care how it is achieved. The portrayal of an authoritative management style is common in the clothing industry due to its labour intensity. But the ability to improve the morale through the philosophy of total management could have a positive impact on the output performance of the industry[6].

Organisations are positive about their future in the industry and adopted an attitude of “survival of the fittest.” Managers are reluctant to believe that the implementation of process improvement methodologies within the production environment would enhance productivity levels [9].

Organisations realised the benefits of work-study principles, 45% complained that they did not have the capacity to apply the process improvement principles. Finally, the implementation of innovative practices with regards to production techniques, design and development of production, manufacturing processes, supply chain management and labour relations should enable clothing manufacturers to maintain and grow within the industry. The multitude and magnitude of challenges facing the SA clothing and textile industry are clear from the information presented.

4.2 PREPLANNING FOR LINE BALANCING

Line balancing is the distribution of work on the line in such a way that all machinists obtain an equal amount of work in terms of time value. All workers should be productive during the day and production should flow smoothly, thereby achieving planned targets. The objective is to accurately delegate workers to the various operations required to complete the product based on their operational skills and to maximise productivity and delivery as per the planned targets.

One of the major oversights on the part of management is that they do not indicate the link between labour costing, order scheduling, line balancing and control to the line staff. Line balancing has 2 stages:

- Pre-planning
- Control

Both stages are inter-complementary and extremely important. Pre-Planning: Planning includes theoretical calculations & assumptions and is generally guided by past practical experience (and in most cases reliant on the memory of the
Historically the method of allocating workers and operations was based purely on a 1 machinist to 1 operation basis. When bottlenecks were experienced, additional labour was allocated to minimise the constraint. Although this achieved the desired effect, labour utilisation was sacrificed. Thus, the productivity of the line was poor, and with lack of measurement practices, there was no factual information to address the situation. Supervisors used under-balanced workers for the preparation of the next order which was not required and thus over-stocked work in progress. The resultant effect was the production of products under stressful conditions for the workforce.

In order to address this situation, an operation bulletin was drawn up during the garment costing stage using pre-determined motion time studies such as GSD (general sewing data) or synthesis, indicating the sequence of operations, necessary machinery and equipment. Just before production could commence the operator allocation bulletin was handed to the supervisor for the allocation of operations. Follow-up procedures were adopted to verify the operation bulletin using work-study principles. Discrepancies were ratified and kept on record for future consultation. The follow-up is often related to forceful methods of target achievement and often unfair to workers allocated to critical operations. This causes discontentment among the workforce and leads to high labour absenteeism and labour turnover, with necessary skills being lost.

Where standard minute values are available, the balancing of the line becomes simple and can be completed via a simple spreadsheet exercise on a computer. This method is used by the proactive individuals who realise that one has to drive the balance rather than let it drive the individual.

Accurate worker to operation ratios were established based on time value expressed as SMV (standard time value). The results reflected that some workers were required for a full day while others were required for only a portion of the day to complete the tasks required. This entails the reliance on the number of different skills possessed by every worker and the ability of the supervisor to optimize the use of these skills which usually manifests itself in the “control” function of line balancing.

4.3. MONITORING AND CONTROLLING THE LINE

Control initially begins with the initial line set up and daily monitoring of operator efficiencies. It includes a thorough knowledge of worker skills and levels of efficiency and can be made available via an updated skills matrix. With a proper understanding of worker skills it is possible to make reliable decisions about operation combinations that would minimize equipment and labour resource requirements.

Thorough planning enhances the control function. A skilled supervisor has to have the knowledge of the appropriate levels of work-in-progress, an analysis of current machinery and its operational features, knowledge of replacement workers in the event of absenteeism or poor skills and full control on the quality of work to be produced. This is often the most difficult part of line balancing as it desires the knowledge, skills and experience of a person that is proficient in the clothing manufacturing industry.

Exploratory information has revealed that supervision is a major problem in the clothing industry. There is a lack of logical and common sense thinking and weakness in mathematical judgement. Supervisors rarely have the ability to foresee problems and are often reactionary, hence the term “fire-fighters.” In most cases the supervisor is the supplier of work and the other term with which supervisors are called are “glorified service-hands.” Observation skills, re-planning and adaptation skills are not used until the very last minute until the crisis arises.
In the process of controlling the production of each operator by recording outputs, it is common to view operations as requiring one worker or half a worker or 1 and half worker. It is important for supervisors to think in terms of time, that is, if a worker is required on an operation for the full day then it is for 60 minutes every hour, and if only half a worker is required then the worker should only spend half an hour on the job to achieve the required target. This is where most supervisors fail at their line balancing activity. They do not actually control the halves, or less than halves. Management should be guiding supervisor to an extent in which they can pre-plan this information so that supervisory control can improve. This would most certainly lead to improved line balancing[3].

An exercise was undertaken to highlight line balancing on a lounge shirt sewing line. The line balance exercise was completed on a simple and regular style with consideration given to controlling periods of 10 minutes. Each operation was studied using stop watch time study methodology. A 34% increase in production was achieved with outputs per hour rising from 81 units per hour to 109 units per hour. The in-line and off-line efficiencies were accurate for the first time in 15 years of efficiency control. The manager, supervisor and worker morale improved the team spirit of the line.

4.4 THE EFFECT OF LINE BALANCING

The following is a before and after example to illustrate the effect of line balancing on a clothing production line. It is important to review the efficiency expectancy, operation output and total output differences for the same style using the same number of workers.

Table 1 - Worker allocation to each operation before balancing
It is imperative to note that should an operation require an operator for a 15 minute period to complete an hour’s production, as defined by the line balance, the operator needs to be allocated another task on the line. Therefore, resource utilisation improves, with the optimisation of productivity and efficiency. This would enable the industry to improve their competitive status, quality, price and delivery dates.

5. RESULTS AND DISCUSSION

5.1 MANAGEMENT COMMITMENT

Any change in the organisation stems from top management. Commitment from management drives the process of change and nothing can be achieved if management does not support the initiative. Once management gives their approval any change is possible, but employees need to understand and support the changes for it to be successful.

Management realised that in order to counteract the threat of imports they would try out the team-working concept although employees wished for “the good old days.” They could not afford huge piles of inventory on the machine floor.

Employees were delighted that the plant manager supported the initiative and frequently visited to find out how they were performing. An employee of the team briefly summarized how he felt and mentioned that any project has to have the support of management and the acceptance from employees for it to be successful. The managing director of the organisation initiated the process of change as funding in terms of labour is involved. Support from management, especially in terms of funding is important for a project of this nature.

The planning, organising, leading and controlling of the project are important as it would benefit the organisation over a period of time. The clothing industry is in need of radical change that would be able to counteract the competition faced. Employees were thankful that they had commitment and the necessary expertise from the management team.

5.2 EDUCATION AND TRAINING

A number of training sessions were held with the team of employees to provide orientation with the objectives of the project. Employees held discussions regarding their concerns so that everybody understood their role in the project. The researcher explained that it was
important for employees to understand the basic principles under which businesses operate, and that the industry was under threat due to globalisation. It was explained that the objectives of the project was to improve their competitive position so that they could maintain their households.

The organisation invested in training and development of employees on an ongoing basis. It was mentioned that training of employees in the latest developments would enhance employee skills and workers would embrace changes in future. Another employee’s experiences was that people would be willing to change if they knew what the change was all about and how it would impact on their work. Mention was made that employee involvement from the very outset would clear any negativity that may be spread through the grapevine within the organisation. It was said that management discussions behind closed doors regarding changes are unhealthy for an organisation. Open communication and the building of trust among the people are extremely important.

An employee mentioned that learning can only take place by change in attitude and behaviour. She mentioned that training makes employees aware of the current occurrences, what employees should expect and the manner in which barriers could be reduced. This is a great motivator for the workforce.

Employees were enlightened on the benefits of team-work and line balancing and how it could change the entire organisation. An employee mentioned that training made them understand the current status in the clothing industry and how they could rise above the competition. It was mentioned that they are setting an example for the industry to follow and they could be the best manufacturer in the clothing industry. The advantages they experienced created team-spirit and they found that they were responsible for the making this project a success.

Another employee mentioned that the concept would be ineffective and that government intervention was the only way that the industry could be saved. The researcher interacted with the individual and convinced him of the way forward.

The employee admitted that he was sceptical and did not want change, but since there was communication with management and training of workers, he would “go with the flow.” The comments suggest that a project such as this needs education, training, communication and management support.

Open communication is important in a project. The sharing of information between management and employees enhances the success of the project. It was mentioned that the dissemination of too much information and the interpretation of the information could cause problems within the work environment. The “grapevine” misinterprets information and employees become despondent. It was mentioned that 15 years ago operators were not allowed to speak and at present communication is encouraged in team-work.

An employee mentioned that this was quite a change for “them.” It was mentioned that approximately 15 years ago the floor manager had an elevated office at a centralised point on the machine floor where there was a clear view of all employees. Supervisors were called into the manager’s office on a regular basis and screamed at while the operators observed his ruthless behaviour. “Management by walk about” (MBWA) has become a prominent feature in the clothing industry. It was mentioned that the manager should be a part of the team on the production floor, know the employees by name and understand the problems experienced. Much could be achieved if team-work is implemented throughout the organisation and all employees strive to achieve the mission and vision of the organisation. Human assets need to be appreciated to enhance their motivational level. Working together could “change a mountain into a molehill,” mentioned an employee [22].

It was mentioned that employees were often ignored and management made all the decisions. Issues such as product quality, customer expectations, productivity were never
disclosed to employees. A motivated workforce can achieve labour efficiency without the pressure from management. It was explained that communication among the employees and management improved quality of production and an empowered employee could definitely add value to the organisation, no matter what problems were faced.

It was mentioned that customer focus and expectations, together with quality and on time delivery is an organisational problem, and not only the responsibility of the floor manager’s. The team effort created a change in the working climate with information sharing that enables employees to better understand the operational aspects of an organisation.

The implementation process outcome indicates that active employee participation with knowledge sharing could improve the performance of the organisation. Sharing information about the costs that go into production and the financial position of the organisation makes employees understand the importance of “right the first time, every time.” With work-study officers involved in the process, all work measurement and method study evaluations were done with the team that shared ideas on methods and ergonomics. With the adoption of transparency in all activities employees understood their situation and that of the organisation [20].

6. RECOMMENDATION AND CONCLUSION

The nature of assembly line balance in the clothing industry is stochastic because of the existence of task time variations. The use of the fifteen minute break intervals for a machinist enables the smoothing of tasks which is appropriate for measuring the effectiveness of assembly line balance in the clothing industry. Based on the real production data collected from a men’s shirt manufacturing factory, the example illustrated in this study clearly showed that the tasks can be categorised in 15 minute intervals. This implies that should an operator complete the required target, the operator needs to be moved to another workstation so that line balance could continue. The time variance should be taken into consideration for improving effectiveness of line balancing. Operators should be moved between workstations to improve the line balancing efficiency [21].

This paper contributes to the existing body of knowledge in clothing manufacture and competitive strategy in two ways. First, it presents the relationship between competitive priorities and its impact within the clothing industry by using work study methodology. Secondly, this study provides empirical evidence on some natural rules of how strategic elements can be applied through management commitment and draws attention to the fundamental strategic issues that are often overlooked in the development of an appropriate manufacturing system for clothing manufacture.

The benefits of the findings in this study lie in their potential applicability to extending the operations of clothing manufacture beyond the system boundaries. This is important nowadays to reaffirm that the clothing industry in South Africa has the potential to survive under a strenuous climate. Managers should understand the importance of implementing competitive priorities and positioning strategies with their manufacturing systems [1].

7. REFERENCES


The Effect Of Time Variations In Assembly Line Balancing
Ramdass, Kruger


[4] Bell, T. 2006. Clothing chains wall over quotas is hypocrisy, SA Media, University of the Free


AN OVERVIEW OF RECENT MACHINING IMPROVEMENTS IN TITANIUM

N.F. Treurnicht* and G.A. Oosthuizen2
1,2Department of Industrial Engineering
University of Stellenbosch, South Africa
nicotr@sun.ac.za
tiaan@sun.ac.za

ABSTRACT

Ti6Al4V is being used increasingly in new generation aircraft, creating a market for high value components. It is argued that knowledge development is the key factor for South African machining suppliers to penetrate the global aerospace supply chains. Several new developments are discussed. Cooling of the machining process is covered from the high pressure through spindle perspective, gaseous application and liquid nitrogen concepts, all benchmarked against conventional flood cooling. The role of tool geometry, coatings and environmental concerns are also discussed from a perspective of the potential benefit a SA machining company can reap from a knowledge investment.

OPSOMMING

Ti6Al4V word toenemend gebruik in die nuwe geslag vliegtuie, wat ‘n mark skep vir hoë waarde komponente. Dit word gestel dat kennis ontwikkeling die deurslaggewende faktor is vir Suid-Afrikaanse masjineringsverskaffers om die globale lugvaart voorsieningsketttings te penetreer. Verskeie nuwe ontwikkelings word bespreek. Verkoeling van die masjineringsproses word gedek van hoëdruk deur spindel toepassing, tot gas verkoeling asook vloeibare stikstof konsepte, wat vergelyk word met konvensionele vloedverkoeling. Die rol van beitelgeometrie, bedekkings en omgewingsimpak word ook bespreek van die perspektief van watter voordeel ‘n SA masjineringsmaatskappy kan verdien van ‘n kennisbelegging.

*Corresponding author
1. INTRODUCTION

In order to discuss the significance of titanium machining, a brief reference is made to a few export initiatives from the local manufacturing industry. South Africa is the only country where a labour and knowledge pool with sufficient depth, exist in Sub-Saharan Africa, if not the whole of Africa, able to support advanced manufacturing. In the pre-1994 era sentiments that the country should build its own car for export or at least for import replacement were the political correctness of the day. Several attempts, dating back to the Dart and Flamingo sports cars, later including the Ranger, using General Motors technology, the Nomad, a Spartan 2WD sports utility vehicle, were made. Less serious endeavours such as the BMW 1800 and 2000SA were company takeovers where the taken over company's tooling was brought here to produce the “special, unique” South African model. None of these initiatives were sustainable. Recent initiatives include the electrical vehicle to be named the Joule.

In the late seventies and eighties technical knowledge spawning innovation in the arms field created the foundation for exports for many years after the change to a fully democratic dispensation in 1994. Exports included armoured vehicles, artillery, ammunition and relatively advanced missile defence systems. The country’s aircraft manufacturing capability was spectacularly displayed with the Rooivalk helicopter. The common denominator between these military equipment export successes and South Africa’s leading position in mining could be ascribed as a knowledge capability. The capturing of international markets in a sustainable way however remains elusive. This is partly because of a competitive landscape where access to large dispensable income home and regional markets, deep rooted technical education, low cost of capital and sophisticated political development support makes entry into the supply chains of the global brands from the developing world extremely difficult. On the contrary, component exports into the major automotive supply chains, fueled by import credit earnings, are a success story. The key seems to be the knowledge and business support the local subsidiary of automotive multinational gives the local manufacturer to develop the capacity. In a sense the country's buying power to buy imported cars is harnessed to empower local companies. A similar solution is perceived to exist when the buying power of the developing country is exploited to gain entry in the major global business chains in the form of counter trade clauses.

The post transformation government pioneered this concept with significant success. Several local companies have now become regular suppliers of Saab Aerospace following the acquisition of the fighter jets and the corresponding counter trade agreements. The prerequisite for this mechanism to succeed to stimulate manufacturing capacity development in the developing country is knowledge. Counter trade is no simple solution. In the case of the Airbus A400M counter trade contracts placed on South African manufacturers, are now clouded with uncertainty. South Africa cancelled its order of eight aircraft after major budget and lead time overruns. After major local investments, suppliers may lose manufacturing contracts such as the tail fin work package that was recently lost by the South African Denel Saab Aviation. This study is part of an initiative funded by the Department of Science and Technology to proactively contribute to a knowledge capability in titanium machining that will enable counter trade when opportunities arise. At the same time the vision that the country’s titanium reserves should be exported as high value components rather than low value ore, is supported. This paper discusses new developments in titanium machining against the background of their productive significance that may be used to attract the interest of global aerospace players in local suppliers, provided that a sufficient knowledge base can be developed.

2. IMPROVEMENTS IN TITANIUM MACHINING

2.1 Background

Aircraft manufacturing is in a growth phase. High fuel prices and environmental concerns further promote the use of titanium. Fuel is the largest cost for all airlines and a major priority for profitable airline operation. For some carriers, the fuel bill can account for up to 50% of direct operating costs. For most carriers fuel expense grew from 15% to more than
25% of total airline operating costs (2003 – 2006)[1]. By 2026, the fuel consumption of the average world fleet is expected to be at three litres per 100 passenger kilometres, similar to the benchmark that the A380 set recently [2]. Figure 1 illustrates the breakdown of different materials in aircraft construction. The specific thermal, corrosion and mechanical properties offered by titanium alloys, provide distinct advantages in engine applications. Their machinability is however poor, currently limiting their use.

A growing market niche for high value titanium machined components is perceived. It creates an opportunity for a supplier to gain a competitive advantage. These factors create an opportunity for a developing country like South Africa with proven technical and manufacturing expertise on an advanced enough level to be suitable for second or third tier supplier status in the global aerospace supply chains. The prospects for increased use of titanium are positive, creating niche areas for new entrants as specialist manufacturing suppliers. Given the high cost of titanium and the difficulty of machining, necessarily sets the stage for a high value part, for which the added manufacturing value could be significant as well. This is considered an export opportunity for South Africa. Current South African titanium practice includes relatively low speeds of around 20 m/min, for high speed steel and 50 m/min for cemented carbide tools. Two or four tooth insert carriers similar to those used for other light metals are frequently used. The tool manufacturer remains a key source of cutting condition knowledge. Competitive demonstrations or “shootouts” to determine whose tools should be bought is favoured above in company knowledge development. The improvements that are candidates for knowledge development in SA are cooling, specifically high pressure cooling, gaseous cooling, new tool materials and coatings, cutter configurations, as well as environmentally conscious machining.

Titanium’s status of being a light metal, but simultaneously falling in the category of difficult to cut heat resistant alloys such as Inconel is because of two major factors. The first is the phenomenon of heat concentration in the insert, close to the cutting edge[3]. The second is the element’s behaviour at elevated temperatures, especially above 500˚C. Generally the material possesses high temperature strength, making machining difficult. As the temperature increases this already problematic state of affairs develop into an even more difficult situation, namely severe chemical reactivity with the tool material. It promotes the formation of a built-up edge as well as accelerated diffusion wear of the tool material [4]. Finding cutting conditions resulting in a temperature at the work to cutting material interface that is within the hot hardness capability of the tool material and that improves productivity simultaneously is a major challenge. This is approached from two perspectives in this paper. The first is to address cooling effectiveness and the second, tool durability by means of investigating wear mechanisms and innovative cutting materials.
2.2 High pressure cooling
The concept of forced cooling to improve surface finish and tool life in machining dates from as early as the 1950’s. Initial work focused on turning, but later the concept expanded to include milling. Wertheim reports on tests done at Iscar in 1992 improving tool life from 3 min to 14 min with 16 bar pressure when turning Inconel 718 [5]. Coolant feed was through holes in the tool. Kovacec et al compared through tool coolant application from an external source as shown in Figure 2 [6]. The following results relate to titanium milling (alloy and heat treatment unknown) specifically. They observe a general reduction of cutting forces in the region of 40%. As cutting speed increases, surface finish generally deteriorates. High pressure cooling counteracts this effect with a 50% reduction in Ra with no significant difference for different pressures and cutting speeds up to 150 m/min. Both 69 and 200 MPa coolant pressures yields promising results. The relatively inferior results for an intermediate pressure of 138 MPa may indicate that coolant flow to achieve an improvement is a complex issue.

Using flood cooling, the conventional cutting fluid boils at 350°C forming an insulating boundary layer or “high temperature blanket” according to Ezugwu et al [7] limiting the cooling effect. High pressure coolant delivery can deliver the coolant closer to the critical point on the secondary shear zone. At lower speeds the benefit is limited but at higher speeds the benefit has been proven. Machado and Wallbank [8] report a 300% enhanced tool life over flood cooling.

In later work Ezugwu [9] reports benefits at lower speeds as well as higher speeds. His further work focused extensively on Inconel 718 where a “seven-fold improvement” was achieved at 50 m/min with coated carbide tools. The benefits of high pressure cooling was pronounced when machining (turning) Ti6Al4V. Under these conditions there were negligible differences in tool life between coated and uncoated carbides. The surface finish was uncompromised. Tool life and coolant pressure is generally positively correlated. At lower speeds (100m/min) even lower coolant pressures (70 bar) already has a significant effect as shown in Figure 2, increasing tool life from below 20 min for flood cooling to more than 70 minutes as shown in Figure 3.
An overview of recent machining improvements in Titanium
Treurnicht, Oosthuizen

Figure 3 - Recorded tool life when machining titanium alloy with uncoated carbide inserts at various coolant pressures, conventional coolant flow and in argon enriched environment[9]

Cutting speeds ranging from 175 m/min to 250 m/min under high pressure cooling also shows significant benefits for finish machining of Ti6Al4V with PCD inserts as shown in Figure 4. Again at the lower end of the speed scale (175m/min) the lower pressure of 70 bar yields good results at a tool life increase of more than 9 times. It is interesting to note that at the intermediate coolant pressure of 110 bars and 175 m/min cutting speed the tool life is less than for the low pressure of 70 bars.

Figure 4 - Tool life cutting titanium alloy with PCD with high pressure vs flood cooling[9]
In tests performed at Stellenbosch, a Mitsubishi coated carbide performed relatively well under high pressure through spindle cooling, yielding a 100% increase in tool life compared to flood cooling as shown in Figure 5. The material was Ti6Al4V heat treated and age hardened to 1080 MPa yield strength. Cutting condition were $V_c=100$ m/min and $f_z=0.05$ mm/z\[10\].

Figure 5 - Performance of Mitsubishi VP15TF coated carbide and PCD inserts with 60 bar coolant pressure vs flood cooling. $V_c=100$ m/min and $F_z=0.05$ mm/z \[10\].

However, in rather extensive tests at Stellenbosch mixed results that were not in agreement with the results in Fig 4 were found. It is clear from Figure 5 that at 100 m/min and in similar tests at 200 m/min, $f_z=0.05$, high pressure through spindle coolant at 40 and 80 bar pressure (TSL40, TSL80) application results in a reduction of tool life relative to flood cooling. Additional tool life tests performed at $V_c=100$ m/min and $f_z=0.05$ yielded similar results yielding an approximate 25% reduction in tool life for 80 bar high pressure cooling compared to flood cooling\[10\]. As part of the testing of new PCD grades at Stellenbosch, the question whether the benefits of high pressure through spindle cooling applies to PCD as well, was investigated. Ezugwu\[9\] reports the opposite, namely a consistent deterioration of tool life using high pressure coolant relative to flood cooling for ceramic tooling. The tests comprised of turning Inconel 718 with ceramic cutting tools with speeds ranging from 200 to 300 m/min, feeds from 0.15 to 0.25 and coolant pressures ranging from 110 bar to 203 bar. The benchmark for these tests are similar tests done on the same alloy with coated carbide tools at speeds ranging from 20 to 50 m/min, feeds from 0.25 to 0.3 mm and the same coolant pressures as for the ceramic tools. In these tests the benefit of high pressure cooling ranged from 8% at the low speed/low feed combinations to more than 700% at the high feed/high speed end of the experimental range. The experimental results are shown in Table 1 and Table 2.

Table 1 - Percentage improvement in tool life with high pressure cooling relative to conventional coolant supply after machining Inconel with coated carbide tools \[9\]

<table>
<thead>
<tr>
<th>Speed (m/min)</th>
<th>Feed rate (mm/rev)</th>
<th>110 bar</th>
<th>150 bar</th>
<th>203 bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.25</td>
<td>8</td>
<td>9.8</td>
<td>-33.8</td>
</tr>
<tr>
<td>30</td>
<td>0.25</td>
<td>87.7</td>
<td>50.6</td>
<td>64.1</td>
</tr>
<tr>
<td>50</td>
<td>0.25</td>
<td>335</td>
<td>411.1</td>
<td>462.8</td>
</tr>
<tr>
<td>20</td>
<td>0.3</td>
<td>8.6</td>
<td>11.5</td>
<td>-43.9</td>
</tr>
<tr>
<td>30</td>
<td>0.3</td>
<td>27.1</td>
<td>95.2</td>
<td>104.5</td>
</tr>
<tr>
<td>50</td>
<td>0.3</td>
<td>517.6</td>
<td>647.2</td>
<td>739.8</td>
</tr>
</tbody>
</table>
An overview of recent machining improvements in Titanium
Treurnicht, Oosthuizen

Table 2 - Percentage deterioration in tool life with high pressure cooling relative to conventional coolant supply after machining Inconel with SiC whisker reinforced alumina ceramic tools [9]

<table>
<thead>
<tr>
<th>Speed (m/min)</th>
<th>Feed rate (mm/rev)</th>
<th>110 bar</th>
<th>150 bar</th>
<th>203 bar</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.15</td>
<td>-48.6</td>
<td>-51.6</td>
<td>-49.6</td>
</tr>
<tr>
<td>250</td>
<td>0.15</td>
<td>-45.7</td>
<td>-38.8</td>
<td>-69.7</td>
</tr>
<tr>
<td>300</td>
<td>0.15</td>
<td>-44.7</td>
<td>-34.1</td>
<td>-63.1</td>
</tr>
<tr>
<td>200</td>
<td>0.25</td>
<td>-4.5</td>
<td>-35</td>
<td>-35</td>
</tr>
<tr>
<td>250</td>
<td>0.25</td>
<td>-47.9</td>
<td>-63</td>
<td>-61.7</td>
</tr>
<tr>
<td>300</td>
<td>0.25</td>
<td>-68.8</td>
<td>-67.5</td>
<td>-65.6</td>
</tr>
</tbody>
</table>

The question was raised whether PCD will benefit from high pressure cooling like tungsten carbide or follow the trend of ceramics. The results for PCD tool life at Stellenbosch as depicted in Figure 5 showed correspondence with the results for ceramics in Table 1 and 2. The result was a reduction of tool life for PCD using high pressure cooling, compared to flood cooling.

Conclusions that are of significance for a competitive advantage are listed below:

Jet design in high pressure cooling is important. In South Africa, where we mostly use imported tooling, the performance of the through spindle cooling of different tool bodies should be evaluated. In the case of turning, jet design can be evaluated against well established norms. Generally the chip needs to be lifted to minimise the length of chip contact. In the case of milling, evaluation is more complex, placing a larger emphasis on carefully designed scientific experimentation. In spite of ISO standardisation for insert geometry, cooling nozzle placement and spray angles are not standardised. Two different tool holders were used in the experiment for the PCD and the carbide inserts. The machinability of titanium under different conditions is reported extensively in literature, however mostly without reference to the specific alloy and heat treatment condition. Commercially pure titanium with a yield strength of 500 MPa is predominantly used for medical applications. Ti6Al4V is the alloy used extensively in machined aerospace components. In the solution treated and aged condition, yielding at more than 1000 MPa, it is a much more difficult to machine material than commercially pure titanium. The lack of reference to the alloy and heat treatment condition is therefore a serious deficiency in many published articles. Flood cooling systems have undergone significant development. Some machines are designed focusing on easy to cool materials, where a pressure of 1 bar and only one nozzle is sufficient. In the experiments reported in this paper a modern system designed for difficult to cut material was used. Five nozzles and 3 bars pressure are used to achieve good penetration. PCD has a higher heat transfer coefficient than carbide resulting in a better utilisation of the higher, better penetrating flow of the modern flood cooling system.
2.3 Gaseous cooling technologies
When titanium is machined at elevated speeds, the work material is cyclically heated to temperatures beyond 500°C. Kitagawa et al[11] reported a tool entry temperature of 820°C.

![Temperature fluctuation in milling operations of titanium](image)

Figure 6 - Temperature fluctuation in milling operations of titanium[11].

If the tool surface temperature is higher than the liquid to gas phase change temperature the water based coolant will be present in two phases, namely liquid and gas. Because the tool is mostly above 100°C, the water fraction of the coolant will be present in both phases. The gas or steam in this case forms an insulating barrier between the tool surface and the coolant. This phenomenon is referred to as delayed surface wetting or coolant jet impingement[12]. A further effect of higher cutting speeds is the large temperature difference between the tool surface during cutting and the coolant, promoting thermal shock. In order to avoid the reduced cooling effect of the coolant at higher temperatures and address thermal shock, experimental work is being performed with gas cooling. A mixture of compressed air with a small addition of emulsion coolant is used. Milling is simulated in a lathe by means of interrupted cutting. The cutting edge is gas cooled or flood cooled in the control. As shown in Fig 7 even under severe interrupted cutting conditions at the normal cutting speed of 50 m/min, the normal practice of flood cooling yields acceptable results without significant difference from the gas cooled cutting edge. In Fig 8 however the flood cooled specimen that was tested at 80 m/min is significantly more damaged than the gas cooled insert.

![Microscopic comparison of the cutting edge of a flood and a gas cooled insert](image)

Figure 7 - Microscopic comparison of the cutting edge of a flood and a gas cooled insert at a cutting speed of 50 m/min (200x magnification) (Koen)
An overview of recent machining improvements in Titanium
Treurnicht, Oosthuizen

Figure 8 - Microscopic comparison of the cutting edge of a flood and a gas cooled insert at a cutting speed of 80 m/min (200x magnification)[13]

Figure 9 - Tool life as material removed for LN2, compared to conventional cooling[13]

The potential is a 180% increase in tool life for finishing ($V_c = 120$ m/min) and 25% higher speed finishing at a cutting speed of 150 m/min. This work is being continued and expanded to address the environmental concerns regarding the disposal of used machining coolant, in collaboration with a Swedish university and industrial research consortium, Innovatum, in Trolhättan, Sweden.

2.4 Performance enhancement of milling strategies

Titanium alloys are used for high value components, not only components used in an aircraft’s frame and engine, but also in the biomedical field. Workshops able of sustained growth will migrate toward higher-end work, resulting in a growing percentage of machining shops that will encounter titanium alloys. Therefore, the investment in knowledge to mill titanium is worth while in order to achieve higher productivity, when raising the cutting speed is not an option. Results [14] from a research collaboration study details the effect of tool geometry on tool life. As illustrated in Figure 10 the side clearance angle of the tool is the most influential.
2.5 Tool coatings

Another solution to protect tool materials from the thermal load is the use of a coating. The perfect coating should have a high temperature resistance, a high toughness and limited thickness in the sub-10 micron range. Figure 11 shows results of different experimental coatings milling Ti6Al4V. None of these coatings improved tool life relative to the uncoated tool. It should however be noted that coatings such as TiSiN was not tested in this study.

Explanations might be that the coated surface is rougher, which leads to the sticking of the chips to the rake of the tool [15]. The cutting edge of the tool is also not necessarily as sharp as an uncoated tool and during the preparation phase (before applying the coating) the tool material can become brittle [16]. Based on this study, it can be concluded that none of the experimental coatings are suited for high speed machining of Ti6Al4V. This work is being continued with the support of Boeing Research and Technology.
4. CONCLUSION

South African companies that either already have achieved supplier status in the international aerospace supply chains or those aspiring to such a level, will find opportunities in knowledge development in titanium machining. The following opportunities have been identified in this review:

- High pressure cooling offers a potential of a 150% improvement in tool life. The technology and knowledge is in development at SA research institutions.
- Gaseous cooling has shown the potential to yield similar wear conditions at a cutting speed that is 60% higher than industrial practice. Practical implementation requires further work.
- Liquid nitrogen cooling can extend tool life by approximately 180% with much more potential. It can enable finishing at 25% higher cutting speed. It requires a more complex infrastructure in terms of liquid nitrogen storage and safety measures than conventional practice, but these services are available in South Africa and has been established for the laser cutting industry in the manufacturing environment. The knowledge is in the early stages of development in SA.
- Tool geometry is a knowledge area that a SA manufacturer can use by means of selection to improve tool life in the order of 30%.
- Coated tools are recommended by tool manufacturers on a large scale without convincing economic benefit according to the bulk of the literature. The work is being continued with strong external interest and support.
- Environmentally friendly machining may be enhanced using liquid nitrogen based cooling.

5. ACKNOWLEDGEMENTS

This research is funded by the Advanced Manufacturing Technology Strategy (AMTS) programme, an initiative of the South African Department of Science and Technology (DST). The support of Mr Ryan Hanks and Dr Kevin Slattery of Boeing Research and Technology, St Louis, Missouri, USA is gratefully acknowledged.

6. REFERENCES

An overview of recent machining improvements in Titanium
Treurnicht, Oosthuizen


AN EFFECTIVE LABOUR PERFORMANCE MEASUREMENT AIMED AT OPTIMISING PRODUCTIVITY FOR A SOUTH AFRICAN COMPANY

D.L. van Blommestein\textsuperscript{1*}, S. Matope\textsuperscript{2}, A.D. Swart\textsuperscript{3} and G.A. Ruthven\textsuperscript{4}

\textsuperscript{1}Department of Industrial Engineering  
University of Stellenbosch, South Africa  
14857790@sun.ac.za

\textsuperscript{2}Department of Industrial Engineering  
University of Stellenbosch, South Africa  
smatope@sun.ac.za

\textsuperscript{3}Department of Mechanical Engineering  
University of Stellenbosch, South Africa  
15025616@sun.ac.za

\textsuperscript{4}Department of Industrial Engineering  
University of Stellenbosch, South Africa  
gar@sun.ac.za

ABSTRACT

A system has been created to automate the work sampling procedure. The automation consists of a number of USB cameras linked to central computer. The computer then uses a random function in C++ to determine when measurements are to be taken. The video camera footage is analysed with OpenCV. OpenCV is used to track the movement of an object at the workstations by using a colour filter to identify the object. This data is then written to a number of text files. The data is then exported to a spreadsheet application. The spreadsheet application generates a report of the labour utilisation.

\textsuperscript{1} The author was enrolled for a BEng (Industrial) degree in the Department of Industrial Engineering, University of Stellenbosch  
\textsuperscript{2} The author was enrolled for a PhD (Manufacturing Engineering) degree in the Department of Industrial Engineering, University of Stellenbosch  
\textsuperscript{3} The author was enrolled for a BEng (Mechanical) degree in the Department of Mechanical Engineering, University of Stellenbosch  
\textsuperscript{4} The author is a senior lecturer in the Department of Industrial Engineering, University of Stellenbosch  
*Corresponding author
1. INTRODUCTION: AUTOMATED WORK SAMPLING

A significant amount of time is being spent by individuals in industry to perform time studies or work sampling. The ever increasing availability of new technology and technological breakthroughs affords us an opportunity to automate these actions. This paper describes the creation and evaluation of a program developed to perform work sampling. Work sampling is based on the fundamental law of probability in which an operator can assume one of two states depending on the amount of movement that occurs during the work sampling period: either busy or not busy.

The system that has been created to automate the work sampling process consists of a number of cameras linked up to a central computer; the computer then uses a random function in C++ to determine when “measurements” are to be taken. The video camera footage is subsequently analysed with Open Source Computer Vision (OpenCV) (Levin[8]). The primary concern was how to isolate and track a moving object. Once it was determined that an object could be efficiently tracked using a colour filter to isolate a coloured glove or sticker in OpenCV, the data required to perform a work sampling analysis then became available (Levin[8]).

OpenCV provides us with real time information about the centre point of a filtered object’s location in terms of pixel values. This data (time and velocity) is then stored to create manageable sized text files in order to afford efficient retrieval and analysis. After the work sampling period has elapsed, the data is extracted into a Microsoft Excel (or any spreadsheet program) for analysis. A report of the labour utilisation can be generated in Microsoft Excel which is then sent to the analyst for review.

2. METHODOLOGY

2.1 Work sampling

Work sampling is a method for analysing the performance of either human labour or machines. It typically provides the same information as a time study, however it is performed faster and at a significantly lower cost.

The automation of work sampling is far less complex than that of a time study because an operation only needs to be classified as either one of two possible states. The automation process of work sampling only requires the tracking of an object that will move or be stationary (or near stationary) in the corresponding two states.

The following list provides the most common applications of work sampling (Freivalds [1]).

- Determining machine and operator utilisation
- Determining allowances
- Establishing time standards

Presently there are a number of work sampling software packages available. These software packages save up to 35 percent of the time required to perform work sampling (Freivalds [1]). However, these software packages only save time by supporting the analysis of the values determined by the analyst. This data therefore still needs to be measured as well as entered into the program, which is time consuming and tedious process. A further improvement would therefore be the automation of sampling and analysis to further reduce the time and cost associated with the work sampling activity.

Work sampling is based on the fundamental law of probability, in which an event, measured at a random interval is classified as either present or absent. By analysing the data in
An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven

Microsoft Excel, and comparing the determined values against preset values, the operator or machine is classified as busy or not busy. The data follows a binomial distribution represented by equation (1);

\[ P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \] ..........(1)

where  
- \( p \) = probability of an event
- \( q = 1-p \)
- \( n \) = number of observations

As \( n \) becomes large the binomial distribution approaches that of a normal distribution. In work sampling studies, the number of observations is typically large and the normal distribution is a satisfactory approximation of the binomial distribution.

The distribution then has a mean = \( \mu \), and a standard deviation of \( \sigma_p \), given by equation (2) below,

\[ \sigma_p = \sqrt{\frac{pq}{n}} \] ..........(2)

Based on the concept of a confidence interval, consider the term \( z_{\alpha/2}\sigma_p \) as the acceptable limit of error \( l \) at a \( (1 - \alpha)100 \) percent confidence error, equation (3) (Freivalds [1]).

\[ l = z_{\alpha/2}\sigma_p = \frac{\sqrt{pq}}{n} \] ..........(3)

The \( l \) value is one of the values that the analyst will specify. The \( l \) value will depend on the desired precision of the results. The difficulty is that \( l \) is dependent on the mean of the data, which is often initially unknown. To illustrate this issue, consider an \( l \) value of 2.5 percent. With a mean of 0.8, the precision of the results would be adequate. Suppose however that the \( p \) value is actually 0.6, then the 2.5 percent value may result in the data not being precise enough. It is therefore suggested that the analyst define the \( l \) in terms of an overall precision. This value will then be adjusted and corrected for when the \( p \) value changes.

The following example best demonstrated the use of an overall precision value. In order to ensure a \( \pm 5 \) percent precision on a task, then \( l = 0.05 \)\( \times \)mean.

From equation (3) above and solving for \( n \) we get equation (4). Equation (4) is the equation used to determine the number of samples required in the study (Freivalds [1]).

\[ n = \frac{z^2_{\alpha/2} pq}{l^2} \] ..........(4)

As \( n \) increases so does that accuracy of the measurements. Typically, with manually performed work sampling the accuracy is limited by the number of samples that can be taken per day. This value increases in the automated system for two reasons;

- The number of samples taken per day can be increased
- A number of operators/machines can be sampled simultaneously

Traditionally the number of observations to be taken has been a tradeoff between what is practical while still maintaining a reasonable accuracy. With the automated system the limiting factor is the maximum cell range of the spreadsheet application (65 536 in...
An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven

Microsoft Excel 2003 and 1,048,576 in Microsoft Excel 2007. The automated system therefore enables a far greater number of samples to be taken per day than would be possible if done manually. This in turn results in more accurate results.

It is important to note that the program has been limited to run fewer than 1,000 samples per day at a 10 second interval length, using a 15 fps (frames per second) camera. This was done to facilitate the handling and data processing as well as to accommodate older spreadsheet applications.

Limiting the number of samples to 1,000 samples per day will not significantly influence the accuracy of the study. Furthermore, the accuracy of the data is also determined by the period over which the measurements are taken. This leads to another requirement of work sampling: The observations need to be taken over as long a period of time as feasible, preferably over several days or weeks. This ensures that the data is more representative of the true performance of the operator.

2.1.1 Calculated values

As was mentioned earlier due to the availability of the raw data, more calculations can be made than could typically be done in a (manual) work sampling study. In addition, the calculation of the following measures is possible:

- Operator utilisation
- Allowances
- Standard time

The variation in working speeds over a shift, between shifts and on different days of the week can also be calculated.

2.1.1.1 Operator utilisation

Operator utilisation is determined by dividing the number of observations classified as busy by the total number of observations of the study (Allan et al [7]).

\[ U = \frac{n_{busy}}{n_{Total}} \]  

2.1.1.2 Allowances

In order to determine a fair standard allowance time a time study is performed and the operator is classified as one of three states:

- In the screen and stationary
- Out of the screen
- Working

It can be assumed that avoidable delays are those which take place while the operator is seated at his station, but is not working. The values for allowances can then be determined using a similar equation to the one above.

\[ A = \frac{n_{Working}}{n_{Total}} \]  

The avoidable delay allowance is given by the following equation.
An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven

\[ D_{\text{avoidable}} = A \left( \frac{n_{\text{stationary}}}{n_{\text{Total}}} \right) \] ......(7)

### 2.1.1.3 Standard time

Observed time is calculated from the working time divided by the number of units produced during that time. The working time is calculated in the spreadsheet application, and the number of units produced is available from the job charts.

\[ OT = \left( \frac{T}{P} \right) \left( \frac{n}{N} \right) \] .................(8)

Normal time is determined by multiplying the observed time by the average performance rating.

\[ NT = OT \left( \frac{R}{100} \right) \] .................(9)

The standard time is then found by adding allowances to the normal time.

### 2.1.2 Planning

In planning for the implementation of the automated system, an estimate for the probability of a worker being classified as working need to be made. Typically, this estimation can be made from historical data. Alternatively, if no historical data is available, the estimate can be determined by using the program to sample for a few days and then using that data to generate the estimate.

It is important that the study be performed on critical activities; the utilisation at bottleneck stations is more important than that of a supporting activity (Sivasubramanian, Selladurai [4]).

### 2.1.3 The program

The program is written in C++ using NetBeans IDE 6.8. Netbeans was used because it creates a stand-alone application. The program therefore needs only OpenCV to be installed on the computer as well as the drivers for the webcam to operate. This makes the operation of the program easy to use as well as protecting the programming code.

The program consists of the following elements: random function, sorting function, OpenCV and writing to a text file.

#### 2.1.3.1 Random function

The program uses a random number generator to determine when measurements are to be taken. Currently the program is using the built-in random function in C++ (rand()). The rand() function is, however, not a good random number generator and a better Mersenne twister generator will be implemented in the next version of this program (Deitel, Deitel[5]).

#### 2.1.3.2 Sorting function

The random numbers generated need to be sorted into sequence and the bubble sort function is used together with the random number generator to ensure that the measurements do not overlap and that the recordings are processed in sequence. The bubble sort function is shown in Figure 1 (Deitel, Deitel[5]).
2.1.3.3 OpenCV

OpenCV is used to analyse the images captured. The image captured by the webcam is filtered, only allowing the desired identifying colour to be displayed. A function was then created to find the centre-point of the pixels. This point’s data is then analysed in real time and a moving average speed is determined. The filter and the center location are represented in Figure 1. In the case when the identifying object is moved out of the screen, as shown in Figure 2, the location returned to C++ is a large negative value. This number is then converted to zero. This in turn means that when the identifying object moves out of the screen a 0.00 units velocity is returned (Joines, Roberts [6]), (Yu et al [9]), (Levin [8]), (Chentao, Feng [3]).
2.1.3.4 Writing to file
The data is written to a text file so that the analysis of the data can be completed on any spreadsheet program. This in turn makes the program flexible and allows the analyst to perform the desired analysis on the data (Deitel, Deitel[5]), (Chandy, Kesselman[10]).

3. RESULTS

3.1 The user interface

The screen presented in Figure 3 is the initialisation screen for the program. Firstly, the user is required to enter the desired values for the confidence interval, the program then returns the z-value corresponding to the user's inputted confidence interval. The range of the confidence intervals is limited in this program. This is done in order to simplify the program, and to ensure that the range available is adequate. Secondly the user is asked to enter values for:

- The probability of an event, entered as a decimal value
- The control limits, entered as a decimal value

The value for the probability is typically taken from historical data. However, in the case that no historical data is available, the user can just enter an estimate for the probability and allow the program to run for a number of days. This data should then be used to determine the actual value for \( p \). The control limits are selected by the analyst.

From these three values the program determines the total number of samples required for the work sampling study.

The final user input is the period of study. This period should be a trade-off between, the accuracy of the data and the time available to perform the study, while still keeping in mind the spreadsheet application’s cell range. For ease of use it is suggested that the number of samples per day be limited to less than a 1000 samples per day. This value can, however, be increased if the data is analysed with Microsoft Excel 2007, which has a larger cell range.

Once the number of samples has been determined, the program runs its random function and determines when the samples are to be taken. The program then creates the text file to which the data will be written, it also opens the webcam and displays both the source

![Figure 3 - Initialization screen with analyst input](image-url)
An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven

and filtered images. These images are displayed for the entire period of the study to enable the analyst to ensure the camera location and filter settings are correct.

The program displays the time, the speed and the moving average when it is recording, and only the time when it is not. In Figure 4 it can be seen that the velocity is equal to 0 as is the moving average (the last column). This indicates that the identifying object was not in the screen. If the object was in the screen, but stationary, the moving average value would be a small number (typically less than 5), this is a result of the sensitivity of the program picking up a small amount of noise that is always present.

![Figure 4 - Application screen when the identifying object is out of screen.](image)

On the other hand, in Figure 5 it can be seen that the velocity and the moving average are not equal to zero. This indicates that the identifying object in the screen is moving and can therefore be classified as working. There are a number of samples in the shown data in which the velocity is equal to 0. This could result from one of the following two events: Firstly, it could be the result of the object moving out of the frame of the camera, or secondly, it could be that the operator is stationary. The second option is however unlikely because the program is highly sensitive to movement. The sensitivity is a result of the pixel size of the frame. Currently the frame size is set at 640 x 480. This value can be adjusted to higher or lower values depending on the performance capabilities of the camera and the computer on which the program is run. Empirical testing has indicated that a 640x480 resolution is adequate and sensitive enough to differentiate between the different possible states, given that the sampling time is at least 10 seconds long.
Next, the recorded values are written to a text file. After every shift a new text file is created. This ensures that the data is stored in manageable sized files. This applies to both the text file and the spreadsheet. When opening a file containing over a million lines typically results in Microsoft WordPad failing. The data stored in the text files can be accessed at any time and exported to the spreadsheet application for analysis.

3.2 The calculations and values

The testing process was done by taking 300 random samples over a three hour period. In the first hour work was done, in the second the item was idle and in the final hour the identifying object was out of screen. In the first part of the experiment the item was moving constantly, it can be seen that in six out of 105 samples the individual was classified as not working as indicated in Table 1. Further investigation found that the values corresponding to these classifications where zero values, indicating that the identifying object was not in the screen for those measurements. It can be seen that when the item is out of screen, the average velocity is zero. When the item is in the screen but stationary, the slight adjustments result in small velocities of the center point.

Table 1 - Test data results

<table>
<thead>
<tr>
<th>Hour</th>
<th>Number of samples</th>
<th>Working</th>
<th>Not working</th>
<th>Average Pixel Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105</td>
<td>99</td>
<td>6</td>
<td>67.50</td>
</tr>
<tr>
<td>2</td>
<td>108</td>
<td>0</td>
<td>108</td>
<td>2.23</td>
</tr>
<tr>
<td>3</td>
<td>87</td>
<td>0</td>
<td>87</td>
<td>0.00</td>
</tr>
</tbody>
</table>
4. DISCUSSION

4.1 Limited detail

One drawback of the automated system is that only three states can be determined, whereas the manual work sampling studies can be more specific in recording the specific activity that the individual was busy with at the time of the sample.

4.2 Light

One of the difficulties with using a low quality camera and filter to analyse images is that if the light varies, the colour that is filtered may no longer be effectively filtered, another more serious issue is that the level of noise from the background may become appreciable if the lighting levels are changed. This noise can cause the tracking dot to jump around and make it appear as though work is being done even when the tracking object is not in the screen. It is therefore important that the lighting on the task being analysed remain constant over the entire observation period. Furthermore from empirical testing, it has become apparent that the camera should be set up in such a way that it automatically sets the exposure and gain levels. The white balance should also be set up to automatically adjust (Levin[8]).

4.3 Cost

The cost of performing work sampling using the software is limited to the price of the camera and the USB extension if required. The program only requires the installation of OpenCV version 2.0 or later on the computer. Most office computers already have a spreadsheet program installed. The cost of utilising the system is therefore low. This in turn makes it available to smaller firms which might not have had the economical standing to perform work sampling.

4.4 Hawthorne effect

In manual time studies and work sampling, an individual watches over the shoulder of an operator. The impact of this human presence can negatively influence the accuracy of the results, as the individual being monitored may either work harder, or work slower so that slack standards are set. This is known as the Hawthorne effect (Barnes [2]).

"The effect is named after experiments that were conducted at the Hawthorne Works of the Western Electric Company from 1927-1932 in which workers' productivity increased in response to both positive and negative changes in working conditions. The investigators concluded that the increased attention brought on by the experimental set-up motivated the workers to improve their performance regardless of working conditions."

With the automated system, there is no stopwatch and no-one looking over the operators shoulder, and although a camera will be present, its impact is less noticeable than that of an individual. The study will therefore be a more accurate representation of the work environment.

5. CONCLUSION AND RECOMMENDATIONS

From the tests done on the system, it is apparent that the system adequately classifies an operator as in one of the three possible states. The text files provide a good data source for analysis in a spreadsheet application. The spreadsheet application that was used (Microsoft Excel 2007) handled the quantity of data without problems.
An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven

The system can also be used to determine machine utilisation by placing a coloured sticker on a moving part of the machine. The program will detect movements, and classify the machine as either being busy or not busy.

There are many advantages of utilising an automated system to perform work sampling. These advantages are summarised below.

- Time savings
- Fewer resource requirements
- Higher accuracy
- More accurate representation of the work environment
- Lower cost
- Can become a permanent installation.
- Data on the activity velocity is available.

There are also a number of disadvantages and difficulties that need to be overcome:

- Setting up the camera filter
- Glove may influence the performance of the operator
- Less detail available

There are a number of improvements to the system that will be implemented in the next version of the program:

1. The use of a glove and a colour filter will be replaced by hard training in OpenCV. This will enable the program to track an item without the negative impact of:
   - Setting up the camera filter
   - Interference (noise) problems
   - No glove or identifying object will be required.
2. The implementation of a Mersenne twister random number generator. The new generator will produce better random numbers than the inbuilt C++ random number generator.
3. The Visual Basic program will be improved to enable selection of the data rather than the presentation of all the data to the analyst.

6. REFERENCES

An effective labour performance measurement aimed at optimising productivity for a South African company - van Blommestein, Matope, Swart, Ruthven


XAP OPEN SOURCE COMMUNICATION PROTOCOL FOR TELEMEDICINE

A.F. van der Merwe¹
¹Department of Industrial Engineering
University of Stellenbosch, South Africa
andrevdm@sun.ac.za

SUMMARY

Control communication between networked telemedicine devices necessitates the activation of incidents or events. One Internet protocol is User Datagram Protocol (UDP). UDP packages can include text data. These text packages can be readily structured and broadcast over the Internet. UDP activated systems/devices are able to react when a UDP package is received. Every device is programmed to recognise packages which contain its UID. The xAP protocol is a schema for basic machines/devices to recognize their own UID’s and to take action according to the instructions contained in the xAP message.

The application of xAP in the intelligent device and tele-medicine environment was observed. It was found that the protocol can control basic event triggers and transfer/transmit basic text data. However, more complex data transfers would require a different protocol. The xAP messaging schema can be utilised to warn a data base of a change in the original device. Such an intelligent data base is then able to request the extensive data en transmit it from that device. FTP (File Transfer Protocol) or HTTP can be utilised here.

OPSOMMING

Beheer kommunikasie tussen netwerk gekoppelde telemedisyne toestelle noodsaak die aktivering van gevalle. Een internet protokol is User Datagram Protocol (UDP). UDP pakkies kan teks data insluit. Teks pakkies kan struktuur word en oor die internet versend word. UDP geaktiveerde toestelle reageer wanneer ‘n UDP pakkie ontvang word. Elke masjien word geprogrammeer om pakkies te herken wat sy UID bevat. Die xAP protocol is ‘n skema vir eenvoudige masjiene om hul UID te herken en aksie te neem na mate die instruksie in die xAP boodskap.

Die toepassing van xAP in die intelligente toestel en telemedisyne omgewing word beskou. Dit is gevind dat die protocol eenvoudige gevalle kan activeer en basiese data kan oordra. In die geval van meer komplekse data sal ‘n ander protocol nodig wees. xAP boodskap kan tipies ‘n databasis waarsku van die verandering by die oorspronklike toestel. So ‘n intelligente databasis kan dan die uitgebreide data aanvra en oordra vanaf die toestel. FTP of HTTP kan hiervoor gebruik word.

¹ The author was employed as lecturer in the Department of Industrial Engineering, University of Stellenbosch, South Africa.
1 INTRODUCTION

In the communication era in which we live, there is no excuse for not transferring medical information timeously and accurately. In South Africa, as in the rest of the world, people’s lives are priceless. There is no compromise or no cost too high to save someone’s life. However, many manufacturers of medical devices continue holding their patented communication protocols close to their chests.

This is understandable if they have spent extensive development costs to reach a reliable platform to which they are able to transfer complex data measured by their devices. Often web-based client software can interface with such devices to extract information on the patient. However, a higher level communication strategy is required to track patients’ health histories.

Simple information gathered on a patient’s health condition often needs only a ‘flag’ that the patient has actually been tested. Such a basic information pack could include the patient’s identification, the date and time, the test type and the result in abbreviated format. More detailed analyses, prognoses and discussions could follow. The question is which network protocol would reliably transfer updated data even when the “computers are offline”.

The success of the Internet lies in its pull strategy. When you need information, you go onto a search engine, often find several related websites, and browse your way through to get to a specific answer. A centralized telemedicine database website-based server would work in the same way. One would be able to log on to the website, retrieve the patient’s history, and page through the details, prognoses and medications. However, when the Internet is down, or where out of Global System for Mobile Communications (GSM) reach, like in rural Africa, another strategy is required.

A push pull strategy could work under these circumstances. The broadcast (push) of a basic information pack could trigger whenever an event occurs. An event could entail any interaction of any patient with any medical service. More detailed data could then be retrieved using a file transmission protocol (ftp) pull system as necessary and when the Internet is once again activated.

In this paper the concept of data transmission in telemedicine is investigated, followed by an overview of the type of data that would be required from a patient. We briefly look at the file types that would be required. Various existing protocols are discussed, including Transmission Control Protocol (TCP), User Datagram Protocol (UDP) and Short Message Service (SMS). Finally, the xAP text based schema and implementation is concluded with consideration to patient confidentiality.

2 DATA TRANSMISSION IN TELEMEDICINE

Telemedicine is a developing application of clinical medicine where medical information is transferred through media for the purpose of consulting, often followed by remote medical procedures or examinations. Typically, only basic patient data needs to be shared between locations. Patients often move between clinics, which would then require their patient records to be transferred.

File transfer is not always compatible between the several electronic devices and software systems which are commercially available to record, capture and store patient conditions electronically. Furthermore, many communication software and web-based interfaces are available to retrieve information from these devices and systems, each containing datasets of a common patient.

Ferrer-Roca et al [1] and Farmer et al [2] described an SMS service to serve Diabetics. The patient would SMS his/her glucose level and body mass to a server-based expert system,
which would compare it to the patient information and it would then SMS instructions back to the patient. Vashney [5] suggested the use of wireless local area networks (WLAN), radio frequency identifications (RFID) and personal area networks (PAN) to communicate. He further suggested using global positioning system (GPS) in certain cases where the location of a patient-coupled device is known. Caffery et al [6] investigated the use of an intelligent message handling system to route emails to a specific counselling service. The turnaround time was drastically reduced due to the correct allocation of questions to experts.

Joo and Kim [7] showed that SMS can be used as a reminder service to a community in a medium term weight loss lifestyle programme. Similarly, Hüslner and Peters [4] showed that SMS reminders can be used to assist in Cape Town tuberculosis (TB) management programmes. Kaplan [3] found that the use of personal mobile phones in telemedicine in developing countries could be effective if the telephone is not shared. He still questions the cost effectiveness of this application. Mao et al. [20] used a mobile pharmacy service system (MPSS) in China to carry information to patients via SMS. Patient feedback showed high levels of satisfaction on the increased service levels which included reminders about medication, practical information about medicines and information about adverse drug reactions.

### 2.1 Data types

Initially we will look at the several existing data types required to store and transport data within the telemedicine system. These look at the most common pieces of data that would accompany a patient prognosis.

#### 2.1.1 Patient identification data

Patients are identified by their ID number. All other information is linked to this single ID number. Different parts of the telemedicine system require different forms of information on the patient. These include: name, contact details, financial details, next-of-kin, medication, prognosis, tests and operation details. Biometric data would categorise here.

#### 2.1.2 Test operator data

A test operator is often also a patient in the system. Identification of the operator should be by ID number similar to that of patients. Additional human resources information could be required, such as qualifications to operate apparatus and to conduct specific tests.

#### 2.1.3 Device identification

Any apparatus has a manufacturer name, model number and date of manufacture. Most have serial numbers which are often related to batches and design versions. Firmware and software have version numbers. Each device used in the assessment of a patient’s condition should be identified according to these parameters for traceability.

#### 2.1.4 Test identification

Tests often require the use of several apparatus. In this case the use of a single device is not possible and results are reported with reference to an agreed test procedure. The expert doctor would trust the integrity of the standardised test results and make recommendations accordingly. Such standardisation of tests is paramount in remote treatment of patients. The test procedure would not necessarily specify the apparatus manufacturer and model. However, it is necessary to specify the apparatus used in a specific test on a specific patient in the test report.

#### 2.1.5 Test device calibration

Some test apparatus such as a ruler and a digital wrist watch do not easily become uncalibrated. However, audio phonic equipment can measure different results after each move. Calibration before and after each test, against a known reference is important to maintain test data accuracy. The measurement apparatus as well as the reference source against which it is checked should be checked at regular intervals by a calibration
laboratory. The laboratory certificate number and date should be reported with every test carried out using the apparatus.

2.1.6 Test result summary
Not all test result details need to be communicated to and stored in the central telemedicine programme. However, each test carried out on each patient needs to be noted, with a summary of the test results. The purpose is to supply summarised medical history to locum doctors and when patients move between clinics. Each test summary that has more detail stored at the source should include a URL from where the detail could be retrieved.

2.1.7 Test result detail
Modern sophisticated devices and older patient files contain more detail than what the test summary on the patient’s telemedicine files hold. The patient’s file would typically contain an URL link to the device from which the details could be retrieved. Such test result details are to be stored in ftp servers.

2.2 File types
In the previous section we list a wide variety of data types that would accompany a patient through the treatment process. In this section we consider the file types that could contain the suggested data types.

Starting with basic personal information of a patient we can suggest storage in text file format. Each text character in the text file is limited to the 256 ASCII characters. This requires only 1 byte (8bits) of data per character. Some compression systems also suggest the removal of spaces to decrease file size. More complex biometric data, RFID information or bar codes have to be stored in more complex files. Typically, bitmap or its compressed version, jpeg, would be used.

If data security is not critical, but information flow is, the text file approach is suggested. One could pose the question: what would a dying man prefer – life saving correspondence reaching his paramedic over a slow but secure data link or fast, insecure data?

An electrocardiogram (ECG), such as that which Jun & Hong-hai [14] suggest, could be transmitted in ‘real’ time over a link which is suitable for voice transmissions. However, ‘real’ time monitoring of a heart rate is not always required. Intelligent monitoring routines of ECG conditions exist in software that can be run on handheld devices which are able to send a trigger in the case of an alarm. This means that the device could report a simple heartbeat to a remote expert system. But when an alarm condition is detected, a full ECG is transmitted in an audio file format.

Prognoses like assessing a skin condition could require high resolution colour imagery which is possible by multimedia service (MMS). Capturing several pictures, at various levels of optical zoom, reduces the need for transmitting large sized multimedia files. In the simplest black and white format each pixel is represented by a one or zero requiring 1 bit per pixel. Grayscale is possible at 16 to 256 bits per pixel, or multi bit colour. In most systems, 16 bit colour is possible, giving 65000 colours, but up to 32million colours can be possible. At this level a 10megapixel picture would require 10million times 25bits of data, amounting to 250million bits or 32MegaByte of data in a single picture. The transmission time of such a single file would take a considerable amount of time even at a healthy 3G transmission rate.

JPEG Compression is not always desired as it creates artefacts that could distort the image should the doctor digitally zoom into the received data image. The use of optical zoom at the camera is therefore required to minimize the size of the file to be transmitted. The next section deals with the data carriers and layers required to transport data from the sender to the receiver devices.
2.3 Protocols & Layers

![Figure 1 - Layers in networked communication](Source: technet.microsoft.com/en-us/library/cc958821.aspx)

In Figure 1 it is shown how the physical wires that we know as the computer network hosts several layers to make the network function. The Internet is mostly used and transferred to our desktop computers. This is done on the Internet layer which carries the packages, TCP and UDP, as shown in the transport layer. Application layer protocols that we recognize include file transport protocol (FTP) and SMTP for emails.

2.3.1 WLAN and LAN

Wireless local area network (WLAN) is a wireless extension of a normal cabled Internet local area network (LAN). A WLAN access point is connected to the structured cabled network and set up to allow WLAN enabled devices to connect to the network. The LAN in turn connects to the Internet service provider via as ADSL or long distance WLAN network. Many cities already have WLAN networks which covers the city omitting the necessity for ADSL connection. High rate users may still require the speed and capacity of ADSL or similar.

2.3.1.1 TCP

TCP is a reliable, connection-oriented protocol that provides error checking and flow control through a virtual link that it establishes and finally terminates. TCP is responsible for the establishment of a TCP connection (TCP handshake), the sequencing and acknowledgment of packets sent, and the recovery of packets lost during transmission.

2.3.1.2 UDP

UDP is an unreliable, connectionless protocol that provides data transport with lower network traffic overheads than TCP. UDP is used when the amount of data to be transferred is small (such as the data that would fit into a single packet), or when the overhead of establishing a TCP connection is not desired or when the applications or upper layer protocols provide reliable delivery. UDP does not error check or offer any flow control; this is left to the application process. It can, however, still it can be used by protocols that provide reliable packet transmission like NFS.
Ohashi et al [9] claims that their experiments also showed that using TCP2 developed by TTT Co. protocol, even in the UDP protocol that does not have retransmission functions; it was possible to establish secure communications and to ensure the quality of ‘real’ time streaming.

2.3.2 GSM
In the context of this research, WLAN and GSM could be seen as performing the same function. Extensive research has been done to allow seamless transition between WLAN and Mobile networks. Several references exist, one of which is by Aust et al [15]. Talukder & Das [12], who proposes an algorithm that enables SMS as a transport bearer between mobile networks with a focus on unstable networks in developing countries.

2.3.2.1 GPRS
Where GSM is the normal audio mobile telephone link, very similar to a normal landline-type telephone system (POTS), GPRS is an enhancement which allows the transfer of voice and data packages at a more economical rate. GPRS would be able to transfer data only on request, and be connected, but idle when there is no data to be transferred. In the case of a data link over the GSM system, the connection needs to be up for the full duration of the connection. Older mobile phones don’t support GPRS.

GPRS, is in many cases, used as a substitute when WLAN connection is not available. Often symbian and other personal devices like mobile phones would connect to the Internet via WLAN when available and when not, connect via GPRS. If it is set up correctly, the inexperienced user should not notice the difference, except that the GPRS is often more costly than the WLAN connection.

2.3.2.2 SMS
SMS, from a technological point of view, refers to sending text that is limited to 160 characters, including spaces. SMS can be a carrier of basic xAP text messages.

2.3.2.3 MMS
Terry [9] mentions the more advanced SMS called Enhanced Messaging Service (EMS). This allows mobile phones to send and receive messages with special text formatting like: italics, bold, animations, pictures, icons, sound effects and special ring tones.

2.3.3 Other Radio frequency (RF)
Bluetooth seems to be the new trend in device interfacing. This is a protocol mainly used for linking between devices and controlling a small radius electronic environment. Istepanian [11] suggested that in a proprietary MedLAN system and due to its small data rates, Bluetooth can easily be used for transferring still images from medical equipment to a mobile cart. Their system is based on an intelligent medical cart that uses short range wired and wireless communication to connect various local devices, and long range WLAN / GSM links to a main system.

Interference between devices using different RF technologies is a further topic of ongoing research.

2.4 Text schema
Having now defined the carrier layers and protocols for the transport of text data, we can look at the format of the text to be transported. The xAP schema specification was developed by several specialists mentioned on www.xapautomation.org. The specification is widely used to carry automation triggers over UDP/IP between UDP enabled devices. The xAP protocol is solely text based, which makes it simple to program and debug. However, text only data has its limitations. With reference to figure 1, it is an application layer protocol. This research attempts to apply xAP as far as text data would suit the distribution of patient information.

XAP is an open source protocol for which a specification was agreed on by several users. The message format consists of a ‘header’ and ‘message’ area. The message area must contain at least one message block, but may contain many. Each message block has a
message section. The maximum message size over the Internet is 1500 characters (bytes) which is due to the UDP limit. Message fragmentation is not supported. However, sequential numbering of messages can be used to attempt transmission of sequential parts of larger text messages. Application layer error correction would have to be applied.

Several applications exist for Windows, Linux, TiVo, infrastructure and soft and hardcoded hardware devices. A UDP broadcast reaches all networked devices at almost the same instant. All devices are therefore triggered almost simultaneously which could reduce waiting time typically known as http and ftp server congestion.

The ‘header’ distinguishes the source of the message. All receiving devices that are interested in the state of a source device parameter would register that source and react on UDP packages from that source. Each UDP message is filtered as a result of having been in the xAP format first. If the incoming message is of importance/relevance, the receiving device will run a routine to acquire the information by parsing the text. The ‘header’ is parsed to assess the source and the UID. If these match the record held by the receiving device, the message area is parsed. The receiving device would then run further routines based on the information in the message block and message section.

Multiple routines could be triggered from a single xAP message. Normally the receiving device would store the last state of a source UID. This state could be boolean, analogue or a text string. Other state definitions can be created for an application. As defined by the schema, all source device states are broadcasted whenever a change in state (form) occurs. This is typically an “event” class message that carries only the information of that one device and the one parameter that changes state. This message is sent immediately when the state changes.

In addition to the event messages, there is a heartbeat which is a periodic message designated by the class “info” in the message area. This heartbeat would broadcast the states of all parameters of a device on a regular basis. To some degree this technique increases data transmission integrity as it repeats the information sent by the event class message. The heartbeat is also the mechanism to receive the condition of all other source devices on startup of the receiving device.

Due to the uncorrected (open loop control) nature of the UDP message which carries the xAP text, Intelligent error reduction should be implemented on the application layer. Sequential numbering of messages from a source device could enable the receiving device to request the resend of messages not received in the sequential numbering. This would require an error correction communication specification which is not currently covered by the xAP specification. The following is a typical example of a xAP message used to interface with SMS:

```
XAP-header
{
  v=13
  hop=1
  uid=FF.1234:0000
  class=sms.message
  source=acme.my.controller
}
outbound
{
  msg=This is a text message from the xAP network
  num=447788123123
}
```

Below is an example of a message from a device reporting several temperature parameters:
XAP-Open Source communication protocol for telemedicine
van der Merwe

XAP-header
{
  v=12
  hop=1
  uid=FF.0002:0000
  class=temp.reading
  source=Rocket.QTemp.Group1
}
reading
{
  Temp1=018.43
  Temp2=027.50
  Temp3=020.12
  Temp4=Unknown
}

The text used in the xAP message is simple and concise. Source device conformance to the schema is important as the receiving device only understands the schema as well as it was taught (programmed). Sending information outside of the bounds of the schema could cause the receiving device to interpret the message incorrectly. Malfunction or “deafness” is typically the result. Deafness occurs when the receiving device acknowledges receipt of a message, but does not react to it, even though it should.

The header and message areas are clearly separated and bound by the “{“ and “}” characters. Every source device should send a unique “source=” address. The UID is similarly unique for each device. There may be several parameters within each device which can be distinguished by the second part of the UID and source address. Note that the intended recipient is not mentioned in the message. The schema therefore does not send the xAP package from a source to a specific recipient. If required, such a restriction can be dealt with in the UDP transmission and port (port 3639 is used for xAP) settings on the Internet layer.

Whereas the UID= and the source= parameters are required to have unique values, the class parameter can only be from a limited list of pre-agreed values. Similarly, in the message area, parameters will only be recognized by the recipient if previously set. A message from a source having the correct UID and source values would be parsed by the receiving device. However, if the message area now contains unknown parameters or values out of range, the receiving device would not interpret such information.

If a source device is upgraded and is able to broadcast new parameters (previously heart rate only - now skin resistance measured together with heart rate) such a source device can simply add the new parameter and its value to the xAP message. The receiving device will still receive the heart rate and normal. However, there will be a new parameter in the set up of the receiving device that can be activated for monitoring should it be required to do so. The information is therefore pushed from the source to the recipient, but the recipient makes the choice to use the information or not. This means that a new patient’s “presence” will be broadcast on the network, but only when a receiving clinic is interested in a patient’s information will it activate that patient file for monitoring.

2.5 Implementation
Not all eventualities are able to be simulated in this experimental phase. However, an attempt is make to propose an architecture which would overcome the basic requirements of a push-pull system in telemedicine using the xAP schema. The transmission of xAP messages was successfully tested between two personal computers running Visual Basic in Excel. A typical telemedicine transfer of patient data was simulated. In device to database transmissions, reliable data transmissions were found, using twenty five Barionet (www.barix.com) programmable logic controllers, each reporting more than a hundred parameters into a single Excel spreadsheet. All devices were connected to a LAN.
Here we look at two examples of xAP message formats which we used to transfer data in the telemedicine network. The first is on the entry of a new patient’s information and the second on a simulated ECG device report:

Example 1: New patient database entry
Firstly each header of each message should be unique to each patient. A common identifier is therefore required. The patient’s national identification number is used. Where the person does not have a national ID, a sequential number is generated from a common list. As soon as a number is used and broadcast, all other systems on the network receive a message with the number which blocks the number from use for new patients. The xAP message header and UID also require an identifier to distinguish it from other UDP messages that could enter through the xAP port 3639. If the purpose of the message is to introduce a new patient to the network, the event would be classed as such to change the contact information of the patient. The header and message area would then look like this:

```plaintext
XAP-header
{
  v=13
  hop=1
  uid=TM.0002:0000
  class=xAPBSC.event
  source=clinic0002.7502185028080.infodb.Florence
}
database.name
{
  FirstName=John
  Surname=Smith
}
database.contact
{
  Cell=0821313382
}
database.history
{
  ActivationDate=20100601
}
```

Note that the UID firstly identifies telemedicine (TM). It then identifies the clinic where the patient was introduced as clinic 0002. Lastly, the system parameter is identified which, in this case, is the basic info parameter “0000”. The class tells us that it is an event of the XAPBSC type. The source again identifies clinic 0002 and the ID number of the patient, followed by the type of content that the message contains, namely, database info, as well as the operator who was logged into the system at the time of transmission, in this case, Florence. A receiving device that is not concerned with the “infodb” of a patient would ignore this message. However, every clinic would have at least one system updating the patient’s data which would then store this information in its files.

Due to the fact that nurse Florence entered (or changed) three types of database data, the message area consists of three message blocks. In this case, the first message block contains the name information (FirstName and Surname) and the identification of the person who made the last change. The same applies to blocks 2 and 3. It can therefore be seen that this system would broadcast only in the event of a change in information and send on only the information that was changed. The activation date would be sent to enable the receiving system to compare with that day’s date to determine whether this would be a new patient or not.

The enrollment of a patient with limited information can therefore be done by any clinic and the file, although containing only partial information, will be broadcast to all other clinics when online. There is no need for a central database for enrolment and for all clinics.
to be online all the time. This mechanism enables mobile clinics to venture into Africa and enroll new patients which will be automatically uploaded when an Internet link becomes available. In this case, the GPS coordinates of where the patient was last treated would be useful and could be included in the patient’s database history.

Example 2: ECG device xAP enabled
Where a device like an ECG apparatus is used to measure a patient’s condition, the full database information is not required by the device. The device would simply maintain a list of enrolled ID numbers of patients that it received (normally by class messages XAPBSC.event or xAPBSC.info). The operator would select the patient from the list and perform the ECG as normal. Upon completion of the tests, the device would then broadcast the abbreviated results of the tests in the following format:

```
xap-header
  {v=13
    hop=1
    uid=TM.0002:0001
    class=xAPBSC.event
    source=clinic0002.7502185028080.ECG0001.Mosadi
  }
ECG.info
  {
    MachineName=ECGmachine
    Manufacturer=Siemens
    Model=ECG05
    LastCalibrated=20100101
  }
ECG.test
  {
    TestType=Standard Cardiogram
    TestDuration(min)=5
    TestOperator=Mosadi
    TestDate=20100301
    TestTimeStart=1543
    PatientAge=45
    PatientBMI=5
    SamplePeriod(millesec)=25
  }
ECG.result
  {
    AverageHeartRate=85
    MinimumHeartrate=67
    MaximumHeartrate=165
    OperatorsComment=patient was anxious and coughed often during the test
    Graph=______
  }
```

In this second message the UID second identifier changed to 0001 which is the first device, i.e., the ECG apparatus. The class of message is an event that occurred. The source is again clinic 2 where the tests were done, and the patients ID number, followed by the device and the identification of the person who operated it i.e. nurse Mosadi. The message area now has three message blocks: ECG.info, ECG.test and ECG.result, respectively containing the machine information and calibration, then the test type and parameters, and lastly the test results. The test results also include a graph which could be created by the source device software in a byte or word format. A graph sent by text would be based on the 255 ASCII
characters set sequentially in relation to the graph shape. The information that can be transferred is very basic, but could be deemed life saving in certain cases. Graph to text encoder and decoder subroutines would be required.

A URL for further information could be included which would appear as a hyperlink to an ftp or http server on the recipient’s side. Details of the device test could then be retrieved when online. The location of this server, whether centrally or at the clinic, would be based on the Internet stability at the clinic. Often centralized server location and management is considered better, but when a clinic is offline, regular patients cannot be treated due to information held offsite. Servers based at the clinics and remote-managed is recommended. The retrieval rate is increased and no offline conditions would result in the non treatment of patients. However, some patients move between clinics, and offsite data backup is considered important in all organisations.

Here xAP automation’s natural mechanism of broadcasting any changes as they occur is helpful as the central robots would seek new data only from the list of xAP sourced systems. When a clinic comes online, the central robots would receive a burst of xAP messages that it would then use to update the central systems. Where the xAP message includes a URL, the central robot would automatically retrieve the file and store it on one or more mirror sites.

The combination of the push of events as they occur by xAP and the pull of more information by remote doctors for prognoses or by the central robots for data assurance naturally reduces data traffic. A reduction in data traffic is what is required by telemedicine which has to operate into areas where the Internet is intermittent and often offline. The proposed xAP mechanisms will function equally efficiently over dialup connections, whether they are GSM or landline connections.

The xAP schema also makes provision for query messages. Whenever a device or system has been offline and comes back online to serve a certain patient, the patient’s data held by other devices can be requested. All devices recognising the query will respond with all information held on the particular patient.

3 CONCLUSION

The effective transfer of telemedicine data via an Internet link that is intermittently off, requires a basic protocol that would broadcast new data as it becomes available. The pull only mechanism of a central server system is not feasible as the receiver can only function when the client application computers are online.

A push-pull system is recommended, where all updated information is broadcast to all systems on the telemedicine network. Each system is set up to listen only for the data that it requires. Certain central systems will be set up to record and store all relevant network traffic. Such systems would rebroadcast updates as required. A query would also be able to be used to manually trigger the rebroadcast function.

Many publications confirm the use of UDP over GPRS and support the use of SMS as a carrier for application layer protocols like xAP. Several simulations of actual data transport between Visual Basic enabled software programs on a LAN. Simultaneous broadcasts from multiple devices on the same network proved flawless.

Ongoing research is being done on the implementation of UDP transmission from WLAN to GPRS in order to transport xAP messages to mobile phones. xAP already provide standard applications for transport of WLAN xAP messages to SMS messages.

It is recommended that all telemedicine devices adopt the simplicity of a xAP protocol for high level communication on a common base.
4 REFERENCES


TELEMEDICINE - LEVERAGING COMPETITIVE ADVANTAGE THROUGH THE USE OF ICT INVESTMENT

L. van Dyk
Department of Industrial Engineering
Stellenbosch University, South Africa
lvd@sun.ac.za

ABSTRACT

The purpose of the paper is to find trends and themes with respect to Telemedicine to contribute to the leveraging of the R300 million ICT infrastructure investment, spurred by the 2010 Soccer World Cup in South Africa, towards a healthier, and consequently, more competitive country. A statistical topic modeling analysis is performed on a total of 1771 peer reviewed articles from the two international journals on Telemedicine. A significant body of research is available on the use of Telemedicine to address diabetes and heart diseases, but not with respect to HIV/AIDS and Tuberculoses. South Africa can benefit from research outputs related to ICT infrastructure and device development as well as systems integration and management or change.

OPSOMMING

Die doel van hierdie artikel is om temas en tendense te identifiseer met betrekking tot Telegeneeskunde, ten einde die meeste te maak van die R300 miljoen ICT infrastruktuurbelegging, wat deur die FIFA Sokker Wereldbeker aangevuur is. ’n Statistiese onderwerpmodelanalise is uitgevoer op ’n total van 1771 eweknie geevalueerde artikels uit twee international Telegeneeskunde journale. ’n Betekenisvolle hoeveelheid navorsing is beskikbaar oor die gebruik van Telegeneeskunde vir die behandeling van diabetes en hartverwante siektes, maar nie met betrekking tot HIV/AIDS en Tuberkulose nie. Suid Afrika kan voordeel trek uit navorsingsuitsette rakende IKT infrastruktuurontwikkeling asook stelselintegratie en die bestuur van verandering.
1. INTRODUCTION

The South African Government recognised the quality of Information and Communications Technologies (ICTs) as a critical “invisible” success factor to the 2010 World Cup. R300 million was invested in telecommunications and broadcasts in preparation for the this event, including the integration and expansion of existing ICT infrastructure and services. In the Global Information Technology Report 2008-2009 [2] the question is asked: “Will the Soccer World Cup of 2014 Bridge the Social Gap through the promotion of ICT … in Brazil?”. Accordingly, a pledge is placed for (amongst others) for a health-care network “creating an environment conducive to training, fostering medical consultations, making appointments online, obtaining second opinions, and practicing telemedicine.”

The purpose of this paper is to find trends and themes with respect to Telemedicine that can be assist in leveraging of the ICT infrastructure investment, spurred by the 2010 Soccer World Cup in South Africa, towards a healthier and, consequently, more competitive country.

1.1 Telemedicine defined

Sood et al [3] systematically reviewed of 104 peer-reviewed articles and position papers as well as definitions from organizations such as the World Health Organization (WHO), the International Telecommunications Union (ITU), the European Commission, the National Health Service (NHS) and professional bodies like the American Telemedicine Association to conclude with definition for telemedicine:

“Telemedicine is a subset of telehealth, uses communications networks for delivery of healthcare services and medical education from one geographical location to another, primarily to address challenges like uneven distribution and shortage of infrastructural and human resources.”

Telemedicine’s dependence on telecommunications technologies is a major reason for its continuous evolution [3]. Hence, the current Telemedicine trends and themes should also link closely with the current ICT infrastructure trends.

1.2 Challenges and opportunities specific to South Africa

The South African government is committed to providing basic healthcare to all South African citizens, of which about 50% is situated in rural areas [4]. In line with the definition above [3], the government identified telemedicine as a strategic tool to improve delivery of equitable healthcare and educational services [5]. The uneven distribution and shortage of human resources in the public health Sector of South Africa, are confirmed by the data in Table 1 and Table 2. [6]
Telemedicine - Leveraging Competitive Advantage
Van Dyk

Table 1 - Medical specialists per 100 000 population in the (Public Sector, South Africa)

<table>
<thead>
<tr>
<th>Year</th>
<th>EC</th>
<th>FS</th>
<th>GP</th>
<th>KZN</th>
<th>LP</th>
<th>MP</th>
<th>NC</th>
<th>NW</th>
<th>WC</th>
<th>ZA</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>2.6</td>
<td>10.9</td>
<td>32.4</td>
<td>7.4</td>
<td>1.0</td>
<td>1.2</td>
<td>2.0</td>
<td>1.5</td>
<td>42.7</td>
<td>11.2</td>
</tr>
<tr>
<td>2001</td>
<td>2.6</td>
<td>9.3</td>
<td>24.6</td>
<td>6.5</td>
<td>0.8</td>
<td>0.6</td>
<td>2.4</td>
<td>1.7</td>
<td>44.5</td>
<td>10.3</td>
</tr>
<tr>
<td>2002</td>
<td>2.3</td>
<td>9.2</td>
<td>25.0</td>
<td>6.3</td>
<td>0.7</td>
<td>0.7</td>
<td>2.2</td>
<td>1.5</td>
<td>39.3</td>
<td>9.8</td>
</tr>
<tr>
<td>2003</td>
<td>2.3</td>
<td>9.2</td>
<td>19.7</td>
<td>6.0</td>
<td>1.0</td>
<td>0.7</td>
<td>2.7</td>
<td>1.5</td>
<td>32.7</td>
<td>8.9</td>
</tr>
<tr>
<td>2004</td>
<td>2.3</td>
<td>13.8</td>
<td>18.3</td>
<td>6.2</td>
<td>1.5</td>
<td>0.5</td>
<td>2.5</td>
<td>1.9</td>
<td>29.1</td>
<td>8.8</td>
</tr>
<tr>
<td>2005</td>
<td>2.2</td>
<td>13.3</td>
<td>20.8</td>
<td>6.2</td>
<td>1.4</td>
<td>1.3</td>
<td>3.1</td>
<td>1.9</td>
<td>31.3</td>
<td>9.2</td>
</tr>
<tr>
<td>2006</td>
<td>2.4</td>
<td>14.2</td>
<td>21.9</td>
<td>6.3</td>
<td>1.6</td>
<td>1.8</td>
<td>2.7</td>
<td>1.7</td>
<td>33.2</td>
<td>9.8</td>
</tr>
<tr>
<td>2007</td>
<td>2.5</td>
<td>14.7</td>
<td>22.3</td>
<td>6.2</td>
<td>1.7</td>
<td>1.6</td>
<td>2.9</td>
<td>1.3</td>
<td>31.9</td>
<td>9.8</td>
</tr>
</tbody>
</table>


Table 2 - Medical practitioners per 100 000 population (Public Sector, South Africa)

<table>
<thead>
<tr>
<th>Year</th>
<th>EC</th>
<th>FS</th>
<th>GP</th>
<th>KZN</th>
<th>LP</th>
<th>MP</th>
<th>NC</th>
<th>NW</th>
<th>WC</th>
<th>ZA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2000</td>
<td>12.3</td>
<td>24.3</td>
<td>36.6</td>
<td>24.0</td>
<td>12.5</td>
<td>16.4</td>
<td>28.9</td>
<td>11.9</td>
<td>39.7</td>
<td>21.9</td>
</tr>
<tr>
<td>2001</td>
<td>12.2</td>
<td>22.2</td>
<td>28.7</td>
<td>22.3</td>
<td>12.2</td>
<td>16.4</td>
<td>26.3</td>
<td>12.2</td>
<td>32.5</td>
<td>19.8</td>
</tr>
<tr>
<td>2002</td>
<td>11.3</td>
<td>23.4</td>
<td>29.1</td>
<td>22.4</td>
<td>9.1</td>
<td>16.6</td>
<td>24.2</td>
<td>11.8</td>
<td>33.1</td>
<td>19.3</td>
</tr>
<tr>
<td>2003</td>
<td>12.7</td>
<td>23.2</td>
<td>25.4</td>
<td>21.3</td>
<td>14.3</td>
<td>17.9</td>
<td>28.4</td>
<td>11.5</td>
<td>31.9</td>
<td>19.7</td>
</tr>
<tr>
<td>2004</td>
<td>15.8</td>
<td>20.7</td>
<td>26.0</td>
<td>24.2</td>
<td>14.3</td>
<td>21.5</td>
<td>36.5</td>
<td>13.6</td>
<td>36.1</td>
<td>21.9</td>
</tr>
<tr>
<td>2005</td>
<td>16.1</td>
<td>21.4</td>
<td>29.7</td>
<td>27.5</td>
<td>14.8</td>
<td>22.0</td>
<td>34.7</td>
<td>14.9</td>
<td>38.8</td>
<td>23.7</td>
</tr>
<tr>
<td>2006</td>
<td>17.0</td>
<td>22.8</td>
<td>29.6</td>
<td>30.1</td>
<td>17.4</td>
<td>20.9</td>
<td>41.2</td>
<td>14.9</td>
<td>33.8</td>
<td>24.4</td>
</tr>
<tr>
<td>2007</td>
<td>17.9</td>
<td>23.2</td>
<td>32.0</td>
<td>34.7</td>
<td>17.4</td>
<td>18.3</td>
<td>35.7</td>
<td>14.1</td>
<td>26.0</td>
<td>19.8</td>
</tr>
</tbody>
</table>


According to the Human Science Research Council, cardiovascular diseases and strokes are responsible for 37% of deaths in South Africa. HIV/Aids (30%) [7] are next, followed other infectious diseases, such as tuberculosis, cancer and diabetes. These five diseases are often referred to as “The other Big 5" [8]. Telemedicine holds promise as a solution of bringing health care to rural areas, especially if a reliable and efficient ICT infrastructure is available. This article seeks areas for further development and research, with the goal of leveraging the ICT infrastructure investment to address the healthcare challenges, specifically with respect to the public health sector.

Internet access charges in South Africa are relatively high when compared with those in industrialized countries and mobile phone coverage is good and uptake is high [4]. According to the International Development Research Centre’s [9] increasing access to mobile telephony in developing countries will have a tremendous impact on the delivery of health services by ICT in rural areas.
2. METHODOLOGY

International peer reviewed journals articles are analyzed by means of statistical topic modeling to identify the current trends and themes with respect to Telemedicine, relevant to the public health sector of South Africa.

2.2 Peer reviewed international articles on Telemedicine

Being a multidisciplinary field, articles on Telemedicine are found in all types of journals, ranging from highly technical journals, such as *IEEE Computer Science*, to the other side of the spectrum, journals such as *The Critical Care Nurse*. Two journals are specifically focused on Telemedicine and were selected as such for purposes of this investigation. Information concerning these journals and the articles included in this study are shown in Table 3:

Table 3: Journals and articles included in this study

<table>
<thead>
<tr>
<th>Journal</th>
<th>Journal of Telemedicine and Telecare</th>
<th>Telemedicine and e-Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Developments in telemedicine and telecare. How different countries and health systems are using new technology in health care.</td>
<td>All aspects of clinical telemedicine practice, technical advances, medical connectivity, enabling technologies, education, health policy and regulation and biomedical and health services research dealing with clinical effectiveness, efficacy and safety of telemedicine and its effects on quality, cost and accessibility of care, medical records and transmission of same.</td>
</tr>
<tr>
<td>Telecare/ eHealth</td>
<td>Telecare is a term given to offering remote care of vulnerable people, providing the care and reassurance needed to allow them to remain living in their own homes.</td>
<td>Telemedicine is generally seen as a subset of e-Health, focusing specifically on solutions where distance is an issue.</td>
</tr>
<tr>
<td>ISSN</td>
<td>ISSN 1758-1109</td>
<td>ISSN: 1556-3669</td>
</tr>
<tr>
<td>Publisher</td>
<td>Swets Blackwell Ltd</td>
<td>Mary Ann Liebert</td>
</tr>
<tr>
<td>Most recent impact factor</td>
<td>0.890</td>
<td>1.389</td>
</tr>
<tr>
<td>Articles included</td>
<td>All of the articles (excluding some special editions) published between September 2003 and June 2010</td>
<td>All articles published between September 2007 to April 2010</td>
</tr>
<tr>
<td>Number of articles</td>
<td>651</td>
<td>1121</td>
</tr>
<tr>
<td>available for this study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>629335</td>
<td>374834</td>
</tr>
<tr>
<td>Unique word count</td>
<td>1643</td>
<td>1216</td>
</tr>
</tbody>
</table>

2.3 Statistical Topic Modeling

In a recent article of the South African Journal of Industrial Engineering, Uys et al. [10] used statistical topic models as mechanism for characterizing themes in the 283 articles published in this journal by Industrial Engineering Researchers over the past 22 years.
“In essence, a topic model is a machine learning technique that can be applied to count data to find the underlying (or latent) structures for a given input dataset ... Statistical methods are applied to the generative model to make the underlying structures explicit.”

The same topic modeling technique and software as used by Uys et al. [10] are used for purposes of this study. The technique used is the Latent Dirichlet Allocation (LDA) topic modelling technique [11]. The software programmed used for both studies, is called CAT, which is an acronym for Content Analysis Toolkit. It was developed by Indutech and a temporary license was granted for purposes of this project.

The different aspects of a topic model are shown in Figure 1 [10]: With the use of topic modeling software, such as CAT, all words are extracted and the word frequency recorded. The words with the highest frequencies are included in the vocabulary list. Then a word-to-document co-occurrence matrix is compiled, where each row represents a word, each column represents a document, and the entries indicate the frequency with which the specific word occurs in the specific document. Then, statistical inference algorithms are used to develop topics, consisting of list of associated words as well as an indication of the strength of this association.

Figure 1 - Different aspects of a topic model [10]

Blei & Lafferty [11] as well as Uys et al. [10] conducted a single analysis in which all documents available to them were included. Crawford et al. [12] performed two separate analyses and then compared the trends uncovered from the International Journal of Project Management and the Project Management Journal respectively. For this study, three analyses were run, following both approaches. Two analyses was firstly done on the respective corpuses of the two journals, followed by an analysis in which all articles of both journals were included (refer to Table 4).
3. RESULTS AND DISCUSSION

The output of these analyses is a set of hyper texted pages that can assist in information discovery. For purposes of the paper the vocabulary-list (words) was used in the first place the get an idea of scope of telemedicine specializations. Secondly, certain topics are focussed on, given the research purpose.

3.1 Current trends and themes in terms of telemedicine specializations

Every statistical topic model produces, amongst others, a vocabulary list in which all words that appear in the topic lists are indexed and ranked. Table 5 shows all of the words from the vocabulary list, respectively for the *Journal of Telemedicine and e-Health* and *Journal for Telemedicine and Telecare*.

Table 5 - Word count ranking: words starting with “tele-“

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>telemedicine</td>
<td>1</td>
<td>1</td>
<td>telephone calls</td>
<td>624</td>
<td></td>
</tr>
<tr>
<td>telehealth</td>
<td>15</td>
<td>15</td>
<td>telesurgery</td>
<td>840</td>
<td></td>
</tr>
<tr>
<td>telemed</td>
<td>106</td>
<td>35</td>
<td>tele-ultrasound</td>
<td>864</td>
<td></td>
</tr>
<tr>
<td>teledermatology</td>
<td>88</td>
<td>82</td>
<td>teleophthalmology</td>
<td>965</td>
<td>776</td>
</tr>
<tr>
<td>telecare</td>
<td>192</td>
<td>3</td>
<td>telementoring</td>
<td>907</td>
<td></td>
</tr>
<tr>
<td>e-health</td>
<td>21</td>
<td>79</td>
<td>teleconsultations</td>
<td>1618</td>
<td>254</td>
</tr>
<tr>
<td>telephone</td>
<td>228</td>
<td>14</td>
<td>telemedicine</td>
<td>979</td>
<td></td>
</tr>
<tr>
<td>telepsychiatry</td>
<td>324</td>
<td>72</td>
<td>applications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>telerehabilitation</td>
<td>217</td>
<td>250</td>
<td>teleoncology</td>
<td>1157</td>
<td></td>
</tr>
<tr>
<td>telemonitoring</td>
<td>237</td>
<td>247</td>
<td>telemedicine</td>
<td>1263</td>
<td></td>
</tr>
<tr>
<td>teledemed telecare</td>
<td>417</td>
<td>71</td>
<td>telestroke</td>
<td>1266</td>
<td></td>
</tr>
<tr>
<td>telepathology</td>
<td>278</td>
<td>358</td>
<td>telecommunication</td>
<td>1458</td>
<td>1203</td>
</tr>
<tr>
<td>telemedicine system</td>
<td>514</td>
<td>371</td>
<td>television</td>
<td>1336</td>
<td></td>
</tr>
<tr>
<td>telemedical</td>
<td>539</td>
<td></td>
<td>telewound group</td>
<td>1346</td>
<td></td>
</tr>
<tr>
<td>telepharmacy</td>
<td>558</td>
<td></td>
<td>tele-audiometric</td>
<td>1364</td>
<td></td>
</tr>
<tr>
<td>telediagnosis</td>
<td>957</td>
<td>266</td>
<td>telepresence</td>
<td>1440</td>
<td></td>
</tr>
<tr>
<td>teleradiology</td>
<td>134</td>
<td></td>
<td>telehomecare</td>
<td>1557</td>
<td></td>
</tr>
<tr>
<td>radiology</td>
<td>619</td>
<td>230</td>
<td>telerobotic</td>
<td>1920</td>
<td></td>
</tr>
</tbody>
</table>
It is not surprising that the word telemedicine is indexed first, with telehealth following shortly thereafter. The index rank of E-Health and Telecare correlates with the titles of the respective journals. Most of the other words are formed through the combination of tele- and a term relating to health service delivery and health education e.g. -health, -dermatology, -cardiology, -radiology, -ophthalmology, -ultrasound, -consultations, -stroke, -surgery, -wound, -pharmacy, -wound, -mentoring, -monitoring, homecare. This list gives us an idea of the extent of telemedicine specializations. Interestingly, teleradiology, does not appear in this vocabulary list of Telemedicine and E-Health.

Non-medical terminologies are -phone (calls), -communication, -vision, -presence, and -robotic. These non-medical terminologies appear more frequently in the Journal of Telemedicine and eHealth than in the Journal of Telemedicine and Telecare.

3.2 Topics relevant to leveraging ICT infrastructure in South Africa

The purpose of this paper is to search for the current trends and themes with respect to Telemedicine that can be assist the the leveraging of ICT infrastructure investment in South Africa, towards a healthier and, consequently, more competitive country. Keywords were extracted from the CAT-analysis vocabulary list, that pertains to the discussion in paragraph 1.2 (challenges and opportunities specific to South Africa). The ten topics (out of 40) mostly associated with these keywords are listed in Table 6.
### Table 6 - Ten topics (out of 40) relevant to South Africa

<table>
<thead>
<tr>
<th>Topic</th>
<th>Keywords in sequence of association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bringing healthcare to rural areas</td>
<td>costs, cost, travel, time, analysis, economic, patient, rural, savings, equipment, clinic, system, direct, days, visits, remote, value, avoided, estimated, hospital, pay, expenses, case, benefits, alternative</td>
</tr>
<tr>
<td>Diabetes: Preventative treatment</td>
<td>diabetic, eye, screening, retinopathy, digital, images, glaucoma, examination, fundus, optic, teleophthalmology, nerve, vision, disease, hospital, standard, software, detection, diagnosis, stereoscopic, camera, referral, risk</td>
</tr>
<tr>
<td>Heart diseases: Diagnoses and monitoring</td>
<td>ecg, heart, telecardiology, patient, disease, fetal, transmitted, diagnosis, risk, myocardial, hospital, acute, coronary, follow-up, trial, symptoms, fibrillation, recorded, cardiologist, standard, quality, recording,</td>
</tr>
<tr>
<td>Telepsychiatry in rural areas</td>
<td>service, telepsychiatry, rural, mental, sites, clients, consultations, clinical, telecare, access, providers, videoconferencing, provider, programmes, community, face-to-face, appointments, remote, data, adolescent, utilization, referral, standards, urban, administrative, education</td>
</tr>
<tr>
<td>Healthcare education, through ICT</td>
<td>education, students, training, learning, knowledge, professional, skills, continuing, distance, practice, sites, nurses, rural, college, e-learning, interactive, quality, participants, web-based, medicine, materials, cme, faculty, rounds, video</td>
</tr>
<tr>
<td>National ICT infrastructure</td>
<td>information, systems, data, electronic, standards, patient, records, Security, national, services, process, quality, management, guidelines, privacy, clinical, software, integration, legal, requirements, access, support, consent, messaging, implementation, decision, technology, providers, model, components, interoperability, policy</td>
</tr>
<tr>
<td>mHealth</td>
<td>mobile, system, phone, information, data, application, mobile phone, messages, patient, service, users, server, sms, management, database, text, wireless, communication, china, applications, software, hospital, interface, platform, layer, request, local, access, web, query, doctors, computer, client, pc, trial, control, automatically, time, file, technology, architecture, multimedia, picture, project</td>
</tr>
<tr>
<td>Image transmission</td>
<td>images, camera, digital, quality, compression, imaging, resolution, display clinical, size, transmission, video, pixels, color, light, capture, screen, computer, dicom, digital camera, file, original, important, view, room, telecare, settings, pictures, compressed, diagnostic, ct, conditions, pc, evaluation, radiology, stereoscopic, software</td>
</tr>
<tr>
<td>Video transmission</td>
<td>video, network, system, transmission, data, audio, wireless, communication, bandwidth, networks, remote, quality, internet, software, connection, devices, emergency, device, connected, communications, application, broadband, computer, packet, real-time, digital, mobile, screen, voice, isdn, ieee, architecture, access, protocol, satellite, evaluation</td>
</tr>
<tr>
<td>Change management</td>
<td>Technology, system, staff, project, communication, organizational, Implementation, nurses, information, attitudes, acceptance, important, training, change, support, providers, professionals, adoption, managers, changes, success, research, development, quality, attitude, problems</td>
</tr>
</tbody>
</table>
The cloud diagram of an eleventh topic is shown in Figure 2. Instead of topic-vocabulary being listed in order of association, the level of association is shown graphically on this diagram. This figure gives a global view in Telemedicine. Of particular interest is the relatively strong association between India and Africa. Much more is published on Telemedicine in India, than (South) Africa. South Africa has much in common with India, since in both countries have both a developing and a developed aspect.

3.3 Time-trend analysis

The example of Uys et al. [10] was followed for purposes of this study, in that the calculated topic model was extended manually, including the issue and year of publication, with the purpose of establishing a time trend of each topic. This is also one of the main reasons why the corpus of each journal was analyzed individually. However, in contrast with Uys et al. [10] the time-trend analysis results for this study were disappointing and not worth publishing. The successful time-trend analysis from the previous study [10] may be attributed to the fact their corpus included articles from 1987-2009, whilst the oldest article in this Telemedicine study dated from 2003. A follow-up study, including all articles, starting from 1996, may deem more satisfying time-trend information.

The only recognizable trends were the themes of the special journal editions. This outcome was disappointing on the one hand, but the positive association between themes of special editions and certain topics, provided unplanned validation of the statistical topic model.

4. CONCLUSION

Thanks to, amongst others, the soccer world cup, considerable ICT infrastructure investments are made in South Africa. This conference considers ways in which investments like these can be exploited towards the attainment of global competitiveness. A country’s competitiveness relies, amongst others, on the health of its population. Telemedicine uses ICT deliver healthcare services where distance is an issue, primarily to address challenges.
Telemedicine - Leveraging Competitive Advantage
Van Dyk

like uneven distribution and shortage of infrastructural and human resources. This study searched for the current trends and themes with respect to Telemedicine, that a post-soccer South Africa can leverage.

The use of Telemedicine for the prevention and treatment of three of the so-called Big-5 priority diseases, namely diabetes, heart diseases and mental illness, receives substantial attention throughout the globe. At least one topic is devoted to each of these (refer to Table 6). Most of these technologies and techniques are applied or can be applied to South Africa. However, little information surfaces concerning the use of ICT to prevent and treat HIV-AIDS and tuberculoses, the other two of the so-called Big-5. Neither of these words, nor variants thereof, appears in the vocabulary list of 2559 words. This can either be interpreted as that Telemedicine is not applicable to the treatment of these illnesses or that it is applicable, but not satisfactory developed, researched and published in these journals. It is also evident that most of the published research applies to first world circumstances, which may explain why so little is published on HIV-AIDS and tuberculoses. The conclusion is made that Telemedicine for the treatment and prevention of HIV-AIDS and Tuberculoses is a specific research and development challenge for South Africa.

This statistical topic analysis also uncovered the fact that not much is published concerning Telemedicine in developing countries: Typically, the names of developing countries have a much lower vocabulary-index-rank that the names of developed countries.

Of interest to the Industrial Engineer and Technology Manager, topics related to ICT infrastructure and device development, integration and management of related change have a strong presence in the corpus of these journals (refer to Table 6). Words like standards and quality can be found throughout these topics. The topics on mHealth, Video Transmission and Image Transmission indicates the opportunities that can be leveraged from the current ICT infrastructure investment. The advantage of bringing health services to rural areas are well documented and closely associated with cost and economic benefits, which will drill down to the competitiveness of the country.

As final concluding remark, the usefulness of statistical topic modelling to explore a multidisciplinary research field is commended. Only the tip of the iceberg is shared in this article. However, through the use of operations research techniques, such as Expectation Propagation, Expectation-Maximisation, Markov Chain Monte Carlo and Variational Bayes [10] a tool for the exploration of other Telemedicine trends and themes was created effectively and efficiently.

5. REFERENCES


SUPPLY RESPONSE ANALYSIS OF THE NAMIBIAN MUTTON INDUSTRY: A METHODOLOGY OVERVIEW

D.N. van Wyk¹;
Department of Industrial Engineering
Stellenbosch University, South Africa
danievw@sun.ac.za

N.F. Treurnicht²;
Department of Industrial Engineering
Stellenbosch University, South Africa
nicotr@sun.ac.za

ABSTRACT

In terms of contribution to the agricultural economic activity in Namibia, the small stock industry is the second most important sector being smaller than only the beef industry. This sector makes a significant contribution to the agricultural sector in Namibia due to the sector’s provision of employment, use of natural resources, contribution to GDP, exports and the contribution to consumer spending as well as food security. Agricultural activities in Namibia contributed 5.7% to Namibia's GDP while 70% of the population relies on agriculture for employment and day to day living.

Livestock in Namibia is free ranging on natural pastures and therefore produces high quality meat that is in high demand in national and international markets. Small stock production in Namibia is unstable due to high variability weather patterns, changes in economic and social environments, unpredictable droughts as well as political and structural changes.

This study investigates the relationship between the various factors contributing to supply dynamics. Previous work in this field that is discussed in this paper includes econometric modelling of various agricultural commodities in USA, Australia, and New Zealand. Similar research on supply response and supply chain dynamics in Southern African countries, such as Zimbabwe, Botswana, Namibia and South Africa is included as background to the investigation.

¹ The author is enrolled for a MScEng (Engineering Management) degree in the Department of Industrial Engineering, Stellenbosch University
² The author lectures at the Department of Industrial Engineering, Stellenbosch University
1. INTRODUCTION

Meat consumption plays an important role in the daily food consumption patterns of the consumer and is therefore seen as a primary product of the Namibian agriculture economy. Because of the demand for red meat and the geography in Namibia, livestock industry is the major contributor to Namibia’s Gross Domestic Product. Namibia is located in the South Western part of Africa with arid and semi-arid bio-resource regions and a population of 2.2 million people. Agriculture contributes 5.7% to the Namibian GDP while 70% of the population is dependent on agriculture to sustain a living. Because of Namibia’s variable and harsh climate agricultural production is dominated by livestock farming, in contrast to a small area of approximate 1% that is suited available for horticulture and crop production [1].

Livestock production consists of cattle, sheep, goats and pigs. Rainfall and vegetation determines the distribution of these livestock types through the various geographical regions of Namibia. Figure 2 shows the relative proportion of livestock marketed in Namibia.

![Spread of Livestock Marketed (2008)](source)

The production sector consists of commercial and communal farmers where the latter predominantly practice subsistence farming. In figure 2 the small stock population is summarized as provided by the Meat Board of Namibia.
Supply response analysis of the Namibian mutton industry
van Wyk, Treurnicht

Figure 2: Sheep Population from 1990-2006 (Source: Meat Board of Namibia Monthly Stats, 2008)

Namibia exports 85% of the annual produce to South Africa. Deboned high quality value added cuts are also exported to Norway [1]. Since 2003 the Namibian government introduced the Small Stock Marketing Scheme that established export restrictions for live animals. Since the Small Stock Marketing Scheme was introduced, live exports decreased while the export of slaughtered carcasses of sheep increased (refer Figure 3).

Figure 3: Total Sheep Marketed 1990-2009 (Source: Meat Board of Namibia Monthly Stats, 2009)

Sheep production is a biological process. This introduces cyclical movements in the number of stock produced and marketed per annum. The fact that stock numbers vary from cycle to cycle suggests that there are many economic, biological and physical (weather) factors affecting the cycles. The time lag between the producer who decides to expand its flock and production resulting from that decision contributes to the cyclic nature of production. Economic and weather conditions as well as other random effects can influence a given cycle both in terms of timing and amplitude [2]. Management and expectations are also factors contributing to the variability of instability of production cycles. Due to variability in
Supply response analysis of the Namibian mutton industry
van Wyk, Treurnicht

Supply, the purpose of conducting a supply response study is to analyse the factors that influence the total agricultural output supply, in this case marketed sheep.

The remainder of the paper is constructed as follows: section two gives a brief background on supply response where section three describes the approaches followed for these studies, whilst section four discusses the previous work done in this field and a model for supply response is presented in section five. The relevance of this study to stakeholders is examined in the last section of this paper.

2. BACKGROUND TO SUPPLY RESPONSE

The supply curve of a product such as mutton is a schedule indicating the quantities producers are willing to supply at a given price, time and locality. The slope of the supply curve is determined by the time it takes producers to adjust to changes. According to the law of supply, if the price of a product increases the supply should also increase. Supply response studies are performed in order to reveal the relationship between inputs and outputs. However as any model generating outputs from inputs, limitations and problems exist, typical problems that occur in supply response studies are:

- Uncertainty in expectations
- Flexibility of fixed factors
- Problem of technological changes
- Measurement of the influence of weather

Proper model design is important for supply response analysis. It is of paramount importance that it should be based on accurate data. The development of these models provides good support to human judgement (management) in marketing, planning, control and policy. According to Sartoruis Von Bach [2], supply response studies for agricultural commodities and products can be differentiated into the following categories:

1) Studies of the supply of individual commodities based on time series data
2) Studies based on budgeting techniques or linear programming models using typical farms or regions as units of analysis
3) Studies of aggregate supply including both the development of theoretical concepts and the estimation of the response of total farm output to changes in product and factor prices

Most studies of supply response, aggregate or individual crop, are based on times series data and either use the Nerlove [3] partial adjustment model for single commodities or the method developed by Griliches [4] for aggregate supply response [5]. For the purpose of this paper the first study on supply of a single commodity is considered for the analysis of supply response for the Namibian sheep industry.

3. APPROACHES TO MEASURE SUPPLY RESPONSE ANALYSIS

To measure agricultural output supply responses to price and other contributing factors broadly two approaches can be followed; 1) Programming 2) Econometrics [6]. The econometric approach can further be subdivided into 2 categories. As there has been considerable scepticism regarding the generated results from econometric analysis a greater understanding of the methodology used to derive the results will provide some indication of the confidence that can be placed in econometric estimates.
3.1 Programming:

Programming Models, usually *linear programming*, involves the creation of a linear production model that represents the typical production system of a specific product or various products. An objective function is usually specified related to profit maximization. Other objectives such as risk minimization can also be defined.

By solving the model with various sets of data, assuming the profit is maximised, the supply-price relationship can be established for a specific product. The advantages of this approach are that linear programming is capable of handling complex multi-relationships on farm level in a production system. The complex multi-relationships involve recognition of all the effects of supply of product prices, input prices, technological and physical restrictions. However, the data requirements are extensive and the collection of data at farm level is costly and the development of such models require some extensive time to be developed [7]. The assumption that farmers maximise their profits may lead to the overestimation of supply that is not always true in practice. Due to the restricted data availability and resources, this approach is not widely used among researchers when supply response studies are conducted.

3.2 Econometrics:

Econometrics was developed out of a desire to further understanding of economic phenomena and was used as being the nature of extension and expression of neoclassical economics. Statistical data proved to be valuable in establishing economic regularities, in presenting economic arguments and performing measurement of economic variables [8]. At the start disproportionate trust was placed in the quantitative techniques, however after initial optimism of the quantitative revolution some disillusionment and scepticism became common. According to Townsend [8] scepticism pivots around thus methodology adopted in formulating these econometric models. Although various econometric approaches exist for analysis of supply response functions, only two of most widely used approaches are discussed in this paper. These approaches include Directly Estimated Partial Models and Error Correction Models, and are briefly discussed below.

3.2.1 Directly Estimated Partial Supply Models

The first approach is *directly estimated partial supply models* that involve the direct estimation of single commodity supply functions from time series data. Production in agriculture is not instantaneous and is dependent on post investment decisions and expectations, meaning the production in any period or season are affected by past decisions. The supply level of supply is a function of current economic conditions, at the time decisions were made as well as the expectation about future conditions [6]. The majority of supply response studies fall in this category.

The partial adjustment model, used by Nerlove [9], is an earlier version of an econometric approach used in measuring agricultural supply response for a single commodity. Nerlove’s partial adjustment model is used to capture agricultural supply response to price incentives. The general static supply function can be mathematically presented as;

$$ Y_t = c + \beta P_{t-1} + \gamma T + \delta_t $$

Where $Y_t$ → expected long-run equilibrium output level at time $t$,
$c$ → constant term
$\beta$ → long-run supply response
$P_{t-1}$ → output price at time $t-1$
$\gamma$ → coefficient of linear deterministic time trend $T$
$\delta_t$ → independently normally distributed error
Supply response analysis of the Namibian mutton industry
van Wyk, Treurnicht

The dynamic adjustment supply response equation is based on Nerlove’s hypothesis that “each year farmers revise the output level they expect to prevail in the coming year in proportion to the error they made in predicting the output level of this period”.

This is presented as;

\[ Y_t^* - Y_{t-1}^* = \lambda(Y_t - Y_{t-1}^*) \quad \text{where } 0 > \lambda > 1 \]  (2)

Where \( Y_t^* \rightarrow \) expected output level at time \( t \),

\( \lambda \rightarrow \) coefficient of expectation about price or elasticity if
variables are in logarithm

By substituting equation (1) in equation (2);

\[ Y_t^* = \lambda c + \lambda \beta P_{t-1} + (1 - \lambda)Y_{t-1}^* + \lambda yT + \lambda \theta_t \]  (3)

where \( \lambda \beta \) captures the short-run price elasticity of supply. Nerlove states that if producers have static expectations and if supply depends on expected normal prices or prices of the preceding year, than the coefficient of expectation, \( \lambda \) is equal to one in equation (2). In such a case producers do not immediately adjust their production decisions to changes in prices observed in period \( t \), such that;

\[ Y_t = Y_t^* = c + \beta P_{t-1} \]  (4)

If \( \lambda \) is less than 1, the fluctuation in expected output level is less than the fluctuation in the observed output level such that the actual change in output level in the periods \( t \) or \( t-1 \) is only a small part of the change required to achieve the expected output level. In this case the only condition for observing significant differences between short and long-run elasticity’s is the initiation non-static assumptions [10], [11].

According to Abou-Talb et al. [10] the previous studies that are based on the Nerlovian partial adjustment model have found low values or even zero long-run price elasticity of agricultural supply. This method assumes that difference between current and long-run planned outputs is eliminated i.e. it assumes that farmers are not forward looking in their production decisions.

This method also lacks the capacity to measure the effect of non-price factors such as rural infrastructure and credit.

According to Abou-Talb et al. [10] and Alemu et al. [12] the Nerlovian partial adjustment model is considered weak for the following reasons:

1. The inability to distinguish between short-run and long-run elasticities
2. The model use integrated series which poses the danger of spurious regression
3. The assumption that production adjusts to a fixed target of supply, towards which actual supply adjusts, is considered unrealistic under dynamic conditions
4. The empirical evidence that Error Correction Models (refer Section 3.2.2) describe the dynamics of supply better that the Partial Adjustment Models

In can be concluded that the partial adjustment model was used as a framework by many previous studies on supply response analysis but due to its limitations and the improvement
of methods the partial adjustment model is less appropriate for the study of supply response on agricultural output.

3.2.2 Error Correction Models

Empirical dynamics of supply can also be described by Error Correction Models (ECM). The ECM form of dynamic specification has been used by various authors in macroeconomic modelling since its appearance in the Davidson, Hendry, Srba and Yeo (DHSY) consumption function of 1978 [13]. The ECM offers a means of reincorporating levels of variables alongside their differences and hence of modelling long-run and short-run relationships between integrated series. In addition to this, economic time series data contain trends overtime, although regression analysis shows significant results with high $R^2$, the results may be spurious. ECM and co-integration analysis is used to overcome the problem of spurious regression [14].

Due to the limitations of the partial adjustment model, the ECM is favoured above the Nerlove method. The ECM overcome the restrictive dynamic specification and captures the forward-looking behaviour of producers optimizing their production in dynamic situations [10]. The ECM approach is used to analyse non-stationary time series data that are know to be co-integrated. This method also assumes co-movement of the variables in the long-run. The general form of the ECM method is;

$$\Delta Y_t = c + \sum_k a_k \Delta Y_{t-k} - \lambda (Y_{t-1} - \sum_j \beta_j X_{j,t-n}) + \gamma T + \vartheta_t$$

(5)

Where,
- $\Delta$ → deference operator such that $\Delta Y_t = Y_t - Y_{t-1}$
- $a_j$ → short run supply elasticity
- $\beta_j$ → long run supply elasticity

$Y_t$ are assumed to be co-integrated time series variables (including previous supply levels of $Y_{t-n}$ and other explanatory variables $X_{t-n}$).

Co-integration analysis can be carried out with the Johansen or Engle-Granger test approach.

**Engle-Granger approach to co-integration analysis**

The Engle-Granger test for co-integration involves estimating a static Ordinary Least Squares (OLS) model where all variables enter at levels. The OLS model usually shows a high $R^2$ with a low Durbin-Watson statistic that indicates significant evidence of serial correlation in the residual. To test for the co-integration of the series, the residual is expected to be stationary i.e. having a unit root. In the spurious OLS regression the t-statistic can not be used to test hypothesis, because the variables are not stationary in levels.

The Augmented Dicky-Fuller (ADF) regression of the form;

$$\Delta \vartheta_t = \delta_0 \vartheta_t + \sum_{i=1}^k \delta_i \Delta \vartheta_{t-k}$$

(6)

is fitted to test for stationarity of the residual $\vartheta_t$ in equation (1). If there is an evidence of a unit root the residual is stationary. I addition to this there is proof of co-integration of the time series in equation (5). K in equation (6) is arbitrarily chosen lagged period until the residual is found stationary when $\delta_i$ is significantly different from zero. Then $\beta$ in equation (5), that is the parameter estimate of long-run supply elasticity, may be interpreted as the co-integration parameter from the linear combination of the series. However, error
Correction representation may result from this co-integration. The error is subsequently corrected by using the residual $\vartheta_t$ to estimate an ECM of the form;

$$\Delta Y_t = c + \alpha \Delta X_{t-n} - \lambda \vartheta_t$$

(7)

However, some researchers criticise the Engle-Granger approach because the estimate of the co-integrating parameter from the statistic regression equation (1) could be subjected to small sample bias. They argue that this can be overcome by estimating the ECM in dynamic form by replacing residual $\vartheta_t$ with the lagged variables $(Y_{t-1} - X_{t-1})$ and $X_{t-1}$ resulting in the following unrestricted regression;

$$\Delta Y_t = c + a_t \Delta Y_{t-n} + \lambda_1 (Y_{t-n} - \beta X_{t-n}) + \lambda_2 X_{t-n}$$

(8)

where, $\alpha$ is the short run elasticity of supply

$\lambda_2$ is dynamic adjustment of supply

$\lambda_1$ is the coefficient of the error/equilibrium correction term of the co-integration equation, that presents the period to adjust to the long-run equilibrium. $\lambda_1$ must be negative and significantly different from zero. Being negative implies that if there is a deviation from the current and long-run levels, there would be adjustment back to the long-run equilibrium in subsequent periods to eliminate the disequilibrium.

If $\beta$ is significant, the ECM captures the speculative behaviour of producers, otherwise the ECM reduces to the Nerlove partial adjustment model. It should be noted that while static co-integration regression predicts the level of supply, the error correction predicts the changes so that the variation in the supply is necessarily higher. The new, corrected co-integration parameter estimate (long-run elasticity), $\beta^*$, is computed as;

$$\beta^* = 1 - \frac{2n}{4n} \frac{\lambda_1}{\lambda_2}$$

if the fraction is negative

(9)

The Engle-Granger approach to co-integration is suited to bivariate relationship [11].

**Johansen Approach to Co-integration Analysis**

The Johansen test of co-integration involves estimating Vector Error Correction Models of the form;

$$\Delta Y_t = c + \sum_j a_j \Delta Y_{t-1} + \delta D_t + \gamma T + \lambda \epsilon_{t-1} + \vartheta_t$$

(10)

where, $\epsilon_{t-1} = ln Y_{t-1} - \sum_j \beta_j Y_{t-1}$ (error/equilibrium correction term)

$D_t$ → vector of stationary exogenous variables

$\delta$ → vector of parameters of exogenous variables

$\lambda$ → coefficient of error correction term $\epsilon_{t-1}$

The Johansen method provides two likelihood ratio tests, namely the Trace and the Maximum Eigen value statistic test, which are used to determine the number of co-integration equations given by the co-integration rank $r$. A co-integration equation is the long-run equation of co-integrated series. The Trace statistic tests the null hypothesis of $r$ co-integrating relations against the alternative of $k$ co-integrating relations, where $k$ is the number of endogenous variables for $r = 0, 1, \ldots, k - 1$. The Maximum Eigen Value statistic tests the null hypothesis of $r$ co-integrating vectors against the alternative of $r + 1$ co-integrating vectors.

When the co-integrating rank $r$ is equal to 1, the Johansen single equation dynamic modelling and the Engle-Granger approach are both valid. When $r$ equals 1, the normalisation restriction for the parameters produces a unique estimate of what the economic theory suggests. However, when there is more than one co-integrating equation
the Johansen approach to co-integrating analysis is preferred to the Engle-Granger approach [11].

4. PREVIOUS WORK DONE ON SUPPLY RESPONSE ANALYSIS

Several research studies undertaken worldwide have been done in the field of supply response. Recent studies increasingly focused on developing countries in Africa such as Ghana, Zimbabwe, Botswana, Namibia, South Africa and Egypt. Earlier studies on supply response predominantly focus on one commodity, where price responsiveness is the major factor of influencing supply. More recent studies used improved quantitative methods as explained in the previous section.

Whipple et al. [15] developed a dynamic model for the U.S. sheep industry to examine the supply response of sheep products, including lamb, mutton and wool. The approach followed was based on dynamic modeling where the producer’s objective is to maximise profits. Reynolds et al. [16] analysed the Australian sheep industry to construct a model of supply that is believed to be faced by the producers in Australia. Relationships of quantities produced between the output of lamb, mutton and wool and changes in inventory levels were hypothesised as describing the producer’s response to prices and seasonal changes. Court [17] investigated the New Zealand sheep industry in order to obtain numerical estimates for economic influences on New Zealand wool, mutton and lamb supply.

Sartoruis von Bach et al [18], analysed the influence of prices and access to markets on the cattle population in certain regions of Namibia, where communal farming dominates. The approach to their study was based on the fundamentals of the work done by Nerlove [9]. Abbott and Ahmed [19] analysed the effect of supply response of wool production towards price changes in South Africa.

Over the last decade various studies have been conducted in Africa on several commodities with new approaches towards supply response. These studies include the analysis conducted by Schimmelpfennig et al [20], Mckay et al. [5], Ocran et al. [21]. Akinboade [22], investigated the influence of producer price and capital/labour ratio on the production of livestock and sorghum in Botswana.

After reviewing most of these studies in the field of supply response analysis on various commodities a few conclusions can be drawn. Firstly most of the studies have used the same methodology by Nerlove in the original form or with modification. These studies mostly have reported low supply response elasticities. Previous studies have also shown that non-price factors are of the same order of importance than price factors in supply response.

5. EMPIRICAL METHODOLOGY FOR SUPPLY RESPONSE MODELLING

For a supply response analysis, investigating the relationship between inputs and outputs, a model is required that includes these inputs and outputs.

Agricultural supply depends on both output and input prices. The fundamental result from free market theory is that output price is the most important determinant of supply [23]. If the output prices increase the profit increase and that motivates producers to produce more. Similarly, an increase in input prices leads to increase in production costs that depress supply. However there are several other factors affecting agricultural production. These factors include; lack of infrastructure, human capital, technology and agroclimatic conditions [24]. Infrastructure includes accessibility of roads, market facilities, farmer access to credit, agro extension services, availability of good breeding stock, pesticides, communication and transport facilities have an affect on the agricultural output through
the effect on productivity and cost of production. Other factors including education, research and extension services have a positive effect on production by increasing production and reducing costs [14].

According to Muchapondwa [23], differences arise between theory and practice when constructing a supply response model. Theory assumes that there is an instantaneous response between inputs and outputs, and this is not experienced in agriculture in practice. Firstly, the agricultural sector is characterised by biological lags between input application and output production. Secondly, for an agricultural production unit, the technical rules implied by the production function are a variable that may change during the course of the production process. Thirdly, for an agricultural production unit, institutional factors may exist that prevent optimal decision making in the production process. Fourthly, perfect knowledge and foresight is not valid in the majority of agricultural firms during the production process. The last factor mentioned by Muchapondwa [23] is that in agriculture, risk and uncertainty is faced by agricultural production units at a higher level than other standard non-agricultural firms or production units. This concludes that the result from production behaviour in practice may differ from the logical theoretical behaviour.

A supply response model of Namibian sheep was hypothesised and constructed according to work completed by previous researchers. As mentioned earlier, in developing countries data availability may restrict the comprehensiveness of the model. The total supply response model presented in equation (11) is based on a marketed cattle production model constructed by Sartoruis von Bach [2];

\[
Z_t = f \left( X_t, X_{t-1}, M_{t-1}, M_{t-2}, E_{t-1}, E_{t-2}, CA_{t-1}, CA_{t-2}, NP_t, NP_{t-1}, N_{PR_{t-1}}, N_{I_{t-1}}, N_{RI_{t-1}}, T_t, T_{t-1}, RP_t, RP_{t-1} \right) \tag{11}
\]

- \(Z_t\) = Total head count of sheep marketed, including all sheep marketed by the producer, i.e. live sheep marketed to export markets, local butchers, export abattoirs etc.
- \(X_{t-1,t-2}\) = Total sheep stock, lagged one year, lagged two years. Sheep stock is influenced by external factors, rainfall, sheep numbers of the previous year etc.
- \(M_{t-1,t-2}\) = Total sheep marketed, lagged one year, lagged two years
- \(E_{t-1,t-2}\) = Total sheep exported, lagged one year, lagged two years
- \(CA_{t-1,t-2}\) = Total carcasses slaughtered, lagged one year, lagged two years
- \(NP_{t-1}\) = Namibian average sheep producer price, lagged one year
- \(NPR_{t-1}\) = Namibian real producer price, lagged one year
- \(CM_{t-1}\) = Mean carcass mass, lagged one year, lagged two years (this variable is used to calculate the Income \(NI\))
- \(N_{I_{t-1}}\) = Namibian Income per carcass, lagged one year, lagged two years
- \(N_{RI_{t-1}}\) = Real Namibian incomer per carcass, lagged one year, lagged two years
- \(T_{t-1}\) = Time, lagged one year,
- \(RP_{t-1}\) = RSA average sheep producer price, lagged one year, lagged two years
Supply response analysis of the Namibian mutton industry
van Wyk, Treurnicht

Time lags were used due to the biological lags between input application and output production, where the inclusion of the time trend, \( T \) is intended to identify the influence of non measurable monotonic time related factors on overall output, including advances in agro-technology, infrastructural facilities designed to increase the production and finally to incorporate fluctuations in natural elements [22].

Other factors influencing production output included by other researchers that was not included into this Sartoruis von Bach [2] study is capital to labour ratios [22], prices of substitutional farming enterprises with competing outputs like cattle, goats or game and feed prices (including the price of maize) [19].

The livestock industry in Namibia is an established commercialised agricultural sector with the necessary infrastructure, marketing systems and extension services in place for optimal production. The elimination of factors that influence the production of livestock in this regard is therefore ignored by Sartoruis von Bach.

Due to the fact that most producers in Namibia, market both sheep and cattle through the same marketing systems, and where the environmental conditions remain the same for both livestock enterprises, the cattle supply response model of Sartoruis von Bach [2] can be used as the fundamental framework for the sheep production supply model. Although some differences may exist in the marketing due to the Small Stock Marketing System, the marketing channels remain the same.

6. CONCLUSION: USEFULLNESS OF SUPPLY RESPONSE MODELLING TO INDUSTRY STAKEHOLDERS

The main reason for conducting research on supply response is to improve the understanding of the price mechanism. The knowledge of supply response greatly assists farmers in their decision making regarding the allocation of resources towards business goals. It can support planners and policy makers to allocate resources and achieve production targets in long term planning. Supply response equations can be used for forecasts based on current agricultural supply response parameters. Therefore, a thorough knowledge of the supply response of food commodities as well as implications of policies will be useful for planning food production against the background of a well balanced development strategy.

A research project on supply response analysis of the Namibian sheep industry will add value to the livestock industry in Namibia due to the fact that important relationships between the producers and other role players in the sheep value chain can be analysed. Stakeholders in the industry can take advantage of this analysis by incorporating these relationships into future decisions and policies.

It can be concluded that in a developing country such as Namibia, research on the field of supply response on a specific food commodity is relevant to a secure sustainable future of food security.
Supply response analysis of the Namibian mutton industry
van Wyk, Treurnicht

6. REFERENCES


Supply response analysis of the Namibian mutton industry

van Wyk, Treurnicht


