\[ F = \frac{1}{\Sigma W \sin \alpha} \sum \left[ \frac{c' B + (W - uB) \tan \phi'}{\cos \alpha + (\sin \alpha \tan \phi'/F)} \right] \]

\[ k_x \frac{\partial^2 h}{\partial x^2} + k_z \frac{\partial^2 h}{\partial z^2} = 0 \]
## CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>The creation and application of a national freight flow model for South Africa</td>
<td>J H Havenga, W J Pienaar</td>
</tr>
<tr>
<td>14</td>
<td>Challenges confronting road freight transport and the use of vehicle-pavement interaction analysis in addressing these challenges</td>
<td>W J vd M Steyn, C L Monismith, W A Nokes, J T Harvey, T J Holland, N Burmas</td>
</tr>
<tr>
<td>22</td>
<td>Air void characterisation of HMA gyratory laboratory-moulded samples and field cores using X-ray computed tomodraphy (X-ray CT)</td>
<td>L F Walubita, B Jamison, A E Alvarez, X Hu, C Mushota</td>
</tr>
<tr>
<td>32</td>
<td>Weak interlayers in flexible and semi-flexible road pavements: Part 1</td>
<td>F Netterberg, M de Beer</td>
</tr>
<tr>
<td>43</td>
<td>Mechanistic modelling of weak interlayers in flexible and semi-flexible road pavements: Part 2</td>
<td>M de Beer, J W Maina, F Netterberg</td>
</tr>
<tr>
<td>55</td>
<td>The effect of parameters on the end buffer impact force history of the crane</td>
<td>T N Haas, P Mainçon, P E Dunaiski</td>
</tr>
<tr>
<td>63</td>
<td>Estimation of the maximum end buffer impact force for a given level of reliability</td>
<td>T N Haas, P Mainçon, P E Dunaiski</td>
</tr>
<tr>
<td>69</td>
<td>Assessment of the behaviour factor for the seismic design of reinforced concrete structural walls according to SANS 10160 – Part 4</td>
<td>R C le Roux, J A Wium</td>
</tr>
<tr>
<td>81</td>
<td>A rational approach to predicting the buckling length of compression chords in prefabricated timber truss roof structures braced by means of diagonal bracing</td>
<td>W M G Burdzik, N W Dekker</td>
</tr>
</tbody>
</table>
The creation and application of a national freight flow model for South Africa

J H Havenga, W J Pienaar

South Africa suffered from a historical lack of freight-flow information that was detrimental to infrastructure planning, optimal network development and market structuring. This paper proposes a methodology that can fill this gap, be repeated annually and is, by its nature, not prone to the errors of market surveys. The methodology develops a comprehensive description of South Africa’s surface freight flow market space based on the definition of four definitive freight flow market segments. The results from the annual South African National Roads Agency (SANRAL) traffic counts are allocated to these segments to develop national road freight flows. For rail freight flows, the rail database is reclassified on a station-to-station (i.e. origin-destination) basis to match the freight flow market segments developed. Consequently, modal flows, market share and total flows for all freight flow market segments and the geographical groupings with the segments can be analysed and reported each year. The results confirm the deteriorating role of South Africa’s rail system amidst growing freight demand, as well as the concomitant over-cropping of the road network, and therefore enable the development of specific national freight transport policy recommendations.

INTRODUCTION AND BACKGROUND

Accurate market information about freight flow was not historically available in South Africa, which compromised planning for optimal network development (Conraddle 2007). Historically, freight transport in South Africa was railroad-driven and formal statistics were available for this sector. Following transport liberalisation and eventual deregulation, rail freight market share declined and flow statistics became increasingly elusive.

The only historical studies that attempted to measure freight flows – the studies of Verburgh (1958 and 1968), Smith (1973), Hamilton (1983 and 1986) and Pretorius (1991) – had limited success, and no permanent data set of freight flows (including modal market share and size) by mode was ever established. Furthermore, their work was not repeatable given the substantial effort required to carry out the research according to the survey methodology. Even the work of Statistics South Africa was discontinued in 2003 owing to issues with validity. This paper proposes a new approach based on a different methodology that can be repeated annually, and is less subject to human error and the ‘lie’ factor encountered in the survey method.

The objectives of the proposed national freight flow model (NFFM) are firstly to provide the first annually repeatable observations on modal market share and freight flows in South Africa for use in infrastructure policy decisions and planning, and secondly to provide lead and lag indicators of the performance of the freight transport network.

Despite the lack of detailed information the challenges experienced in the surface freight flow market space have been well documented by government, researchers and the media (DoT 1998 and 2005; CSIR, Imperial Logistics & University of Stellenbosch 2010; Van Eeden & Havenga 2010). The challenges caused by the modal imbalance (between road and rail) are not sustainable, but the development of more optimum solutions requires the availability of appropriate market intelligence. In the case of South Africa where, in the collective consciousness, rail has mostly been relegated to a provider of bulk transport services, this market intelligence includes a renewed understanding of the drivers of rail sustainability. This is expounded in the next section.

THE DRIVERS OF RAIL SUSTAINABILITY

In normal rail economics about 75% of costs would be fixed over the short term and 50% over the medium term. During deregulation traffic will shift from rail to road (especially time- and value-sensitive cargo), with a concomitant loss of income. But only a partial reduction in costs. This shift will be accelerated and pronounced in the absence...
of intermodal solutions. The remaining fixed costs would have to be cross-subsidised from ‘rail-captured’ freight in the short and medium term. In the absence of cross-subsidisation the remaining traffic would have to shoulder a significant tariff increase, which would result in more traffic shifting to road.

This sequence of events did take place in South Africa and the following hybrid, unsystematic strategy was followed to alleviate the situation:

1. Cross-subsidisation did occur to some extent between rail’s export lines and general freight, making South Africa’s ‘captured’ rail traffic slightly less competitive. This was partly offset by state-of-the-art world-class engineering, especially on the export lines.

2. Relative tariff growth on some higher-value, time-sensitive freight did occur at faster than PPI growth rates, but this exacerbated the problem as more traffic left the railway.

3. The loss of income, amidst high fixed costs, resulted in declining investment and expenditure on maintenance, inducing further freight losses and a self-reinforcing downward spiral.

Globally, this problem usually has three possible outcomes:

1. Rail decline on shorter haul is allowed to continue and the railways concentrate on bulk heavy haul over long distances (the American model).

2. Investment is made in intermodal solutions to get the best of both possible worlds (part of the European and American models).

3. Re-regulation (part of the European model).

Congestion and social factors in South Africa necessitate the adoption of the European model. South African railway management in the mid-1990s, however, planned for the American option, even though this would have been impossible to implement due to inter alia potential job losses and strain on road infrastructure. The impact of deregulation was therefore compounded by poor strategy and the fact that assets were allowed to deteriorate even further to the brink of collapse. No investment in intermodal solutions occurred, and these problems were exacerbated by an absence of guidance from the policy makers, the failure to develop a revised national transport policy following the De Villiers deregulation report, and a lack of implementation of Moving South Africa initiatives.

One of the major causes of this situation was a poor understanding of the freight flow market segments, the real trends in transport volumes for these segments and the limited or erroneous assumptions that followed this shortage of information. The belief was that rail decline was ‘normal’ (erroneous: the decline should not have been so pronounced for corridors and especially high-density, long-haul corridors); that it was caused by poor service (erroneous: poor service was not the only cause of decline, as the events described above highlight); and that it was not rapid (erroneous: this was masked by the rise of ‘rail-captured’ export machines and the concomitant absence of a network view).

The drivers of rail sustainability are best described by comparing or offsetting the positive contribution of a railway to an economy, together with its disadvantages. A railway will, under the right circumstances, save an economy money and provide systemic access for freight and passenger movements and environmentally sustainable transport solutions. These advantages should be offset or compared to rail’s major disadvantage, i.e. that it provides only one degree of freedom of movement. The only way in which the advantages of rail can be monetised in the face of its disadvantage is by its not competing with other modes directly, but by exploiting the intrinsic technologies of rail, i.e. its bearing, guiding and coupling technologies compared to the capacity that it can leverage.

Bearing indicates the weight of axle load that can be maintained and, therefore, volumes. Guiding indicates the wheel on track differentials and, therefore, the speed of movement. Coupling refers to long trains with massive volumes. Combined, these technologies provide high-volume, long-distance solutions (Van der Meulen 2007).

Figure 1 depicts two drivers of these intrinsic rail technologies—the speed of guiding and the axle load of bearing.

The four areas of competitiveness in this depiction of drivers indicate that:

- position A is suitable for heavy-haul traffic
- position B is suitable for heavy intermodal traffic
- position C is suitable for fast, intercity, high-value traffic or passengers, and
- position D is suitable for general freight solutions in a regulated environment.

Van der Meulen (2007) maintains that all railways in the D group (where South Africa’s rail system is located) will gradually become redundant and that the problem can only be solved at the state level, where it was created (i.e. by redesign).

South Africa’s rail system was designed without A, B and C in mind; it was highly regulated for a long time (and was able, therefore, to survive for a long time), but was destined to fail with deregulation. This failure is, therefore, a combination of the incorrect application of rail economics caused by de-densification of loads (as a result of deregulation), which in turn was caused by the absence of intermodal solutions and rail network design errors. All of these errors could have been avoided by macroeconomic research on actual freight transport demand, coupled with design options to meet that demand in a sustainable package of solutions and design.

Macroeconomic research of this kind requires time-series information, and to understand this it is necessary to see the detailed picture that underlies the freight flow market segments and the trend in traffic movement across the segments. The
This established base makes correlation a descriptive model fitted on historic data. It is however, not possible to isolate the actual routes to and from the border posts. It is, the presence of counting stations on all the same set of assumptions should be applied. when the model application is repeated, the however, always be borne in mind and, infrastructure of) more than one segment.

data and research are used for planning technically correct, especially where the nature of the input information for road data (described in the section on road freight flows), the model produces aggregate tonnages as well as corridor flow estimates, but not at the level of detailed origin-destination flows. Certain trucks that were counted may have been traversing more than one of the three freight flow market segments, i.e. corridor, rural or metropolitan, and even more than one geographical grouping within the freight flow market segments without the specific distinctions being determinable. This is technically correct, especially where the data and research are used for planning purposes, since this freight impacts on (the infrastructure of) more than one segment. The specifics of this observation should, however, always be borne in mind and, when the model application is repeated, the same set of assumptions should be applied.

The impact of cross-border traffic on South Africa’s network is included, due to the presence of counting stations on all the routes to and from the border posts. It is, however, not possible to isolate the actual cross-border traffic for impact analysis on border posts and neighbouring countries. This is being addressed through engineering access to border post information from the Treasury, as well as initiatives to conduct input-output modelling for neighbouring countries.

The model as presented in this paper is a descriptive model fitted on historic data. This established base makes correlation analysis with GDP and other macroeconomic variables possible, which renders the model useful for forecasting in future. Subsequent to the development of the NFFM, a detailed origin-destination model was, however, developed for South Africa, based on the input-output model, with a 30-year forecast (Havenga, Simpson & Fourie 2011). The relative ease of updating of the NFFM (compare to the onerous input-output model) allows the two models to act as invaluable mutual validation tools, and the validity of both models have been established and is being tracked over time (CSIR, Imperial Logistics & University of Stellenbosch 2010).

### Definition of freight flow market segments

The first step is to develop a definitive definition of the surface freight transport driven by appropriate freight flow market segments. Given that the modelling objective was to understand freight flows to facilitate infrastructure planning, segment definition was driven by a geographical approach and an iterative process was followed – the segments were defined at the outset to facilitate data classification, but once the data had been analysed, outliers were interrogated to ensure that the market description was appropriate. This resulted in the following overarching segments: corridor, metropolitan, rural and rail export machines. The latter are South Africa’s world-class, dedicated, bulk mining export flows – a captive market for rail and therefore showed separately. The salient characteristics of each freight market segment are compared in Table 1.

Subsequently, the results from the South African National Roads Agency Ltd’s (SANRAL) Traffic Counting Yearbooks, compiled by Mikros Traffic Monitoring (2005 and 2006), as well as actual rail freight flow data obtained from Transnet Freight Rail, are classified according to these segments to develop total freight traffic flows in South Africa.

### Rail freight flows

The rail freight data obtained by previous researchers was one-dimensional, i.e. it provided only volumes shipped and not rail traffic flows between origins and destinations. Transnet Freight Rail (TFR) made detailed origin–destination rail flows available, which are adequate for use in this model, as TFR is virtually the only rail freight operator in South Africa. Rail freight flows therefore do not have to be modelled, but still have to be allocated to the freight transport market segments to allow market share comparisons between road and rail. Transnet Freight Rail has subsequently added this classification system to its core database, which enables, and fast-tracks, the use of rail’s historical data and the annual repeatability of the model.

### Road freight flows

Comprehensive traffic observations started in South Africa in 1984 with a pilot study on the 600 km N3 route between Johannesburg and Durban. As a result of the success of this study the National Transport Commission (now SANRAL) decided in June 1985 to expand the Traffic Counting Network to traffic counting stations (Mikros 2006, p iv). SANRAL now repeats this work annually and a reasonable degree of stability in the work process has been achieved since the 1990s. The main objective of SANRAL’s efforts is not to develop freight flows, but to understand congestion points and enable planning in terms of all vehicular road usage in South Africa (of which heavy vehicles are only a subset). No attempt has ever been made to use the information for freight flow estimation purposes, and a reasonable amount of modelling of the available data is therefore necessary.

The SANRAL Traffic Counting Yearbook is a compendium of traffic information obtained at traffic counting stations on primary roads, which highlights the latest available traffic characteristics. Both permanent and secondary stations are utilised. A permanent station makes continuous traffic observations; a secondary station is one where traffic observations are made on a sampling basis for at least 168 consecutive hours per annum. The 2006 SANRAL Traffic Counting Yearbook contains information on 398 permanent stations and 430 secondary stations. However, the position and number of the secondary stations change from year to year – an issue that the modelling process should consider. The stations are placed on

---

**Table 1 Definition of transport typologies (Havenga 2007: 146)**

<table>
<thead>
<tr>
<th>Traffic type</th>
<th>Corridor</th>
<th>Rural</th>
<th>Metropolitan</th>
<th>Rail export machines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mostly manufacturing</td>
<td>Mostly agriculture</td>
<td>Mostly final delivery</td>
<td>Bulk, low-value (export coal and iron ore)</td>
<td></td>
</tr>
<tr>
<td>Some agriculture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>Long and short</td>
<td>Medium and short</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td>Origin-Destination</td>
<td>Few long distance ODs</td>
<td>Many</td>
<td>Few, one-directional</td>
<td></td>
</tr>
<tr>
<td>pairs</td>
<td>Many ODs at end-points</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major challenge</td>
<td>Spatial organisation</td>
<td>Development</td>
<td>Congestion alleviation</td>
<td>Global competitiveness</td>
</tr>
<tr>
<td>Efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistics approach</td>
<td>Intermodal solutions</td>
<td>Effective road</td>
<td>Effective freight delivery</td>
<td>Ring-fenced, bulk rail systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>feeder system</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Methodology of the National Freight Flow Model**

The NFFM focuses on defining surface freight transport (i.e. road and rail) based on the two definitive measures of flows, namely tons and ton-kilometres. Due to the nature of the input information for road data (described in the section on road freight flows), the model produces aggregate tonnages as well as corridor flow estimates, but not at the level of detailed origin-destination flows. Certain trucks that were counted may have been traversing more than one of the three freight flow market segments, i.e. corridor, rural or metropolitan, and even more than one geographical grouping within the freight flow market segments without the specific distinctions being determinable. This is technically correct, especially where the data and research are used for planning purposes, since this freight impacts on (the infrastructure of) more than one segment. The specifics of this observation should, however, always be borne in mind and, when the model application is repeated, the same set of assumptions should be applied.

The impact of cross-border traffic on South Africa’s network is included, due to the presence of counting stations on all the routes to and from the border posts. It is, however, not possible to isolate the actual cross-border traffic for impact analysis on border posts and neighbouring countries. This is being addressed through engineering access to border post information from the Treasury, as well as initiatives to conduct input-output modelling for neighbouring countries.

The model as presented in this paper is a descriptive model fitted on historic data. This established base makes correlation analysis with GDP and other macroeconomic variables possible, which renders the model useful for forecasting in future. Subsequent to the development of the NFFM, a detailed origin-destination model was, however, developed for South Africa, based on the input-output model, with a 30-year forecast (Havenga, Simpson & Fourie 2011). The relative ease of updating of the NFFM (compare to the onerous input-output model) allows the two models to act as invaluable mutual validation tools, and the validity of both models have been established and is being tracked over time (CSIR, Imperial Logistics & University of Stellenbosch 2010).

### Definition of freight flow market segments

The first step is to develop a definitive definition of the surface freight transport driven by appropriate freight flow market segments. Given that the modelling objective was to understand freight flows to facilitate infrastructure planning, segment definition was driven by a geographical approach and an iterative process was followed – the segments were defined at the outset to facilitate data classification, but once the data had been analysed, outliers were interrogated to ensure that the market description was appropriate. This resulted in the following overarching segments: corridor, metropolitan, rural and rail export machines. The latter are South Africa’s world-class, dedicated, bulk mining export flows – a captive market for rail and therefore showed separately. The salient characteristics of each freight market segment are compared in Table 1.

Subsequently, the results from the South African National Roads Agency Ltd’s (SANRAL) Traffic Counting Yearbooks, compiled by Mikros Traffic Monitoring (2005 and 2006), as well as actual rail freight flow data obtained from Transnet Freight Rail, are classified according to these segments to develop total freight traffic flows in South Africa.

### Rail freight flows

The rail freight data obtained by previous researchers was one-dimensional, i.e. it provided only volumes shipped and not rail traffic flows between origins and destinations. Transnet Freight Rail (TFR) made detailed origin–destination rail flows available, which are adequate for use in this model, as TFR is virtually the only rail freight operator in South Africa. Rail freight flows therefore do not have to be modelled, but still have to be allocated to the freight transport market segments to allow market share comparisons between road and rail. Transnet Freight Rail has subsequently added this classification system to its core database, which enables, and fast-tracks, the use of rail’s historical data and the annual repeatability of the model.

### Road freight flows

Comprehensive traffic observations started in South Africa in 1984 with a pilot study on the 600 km N3 route between Johannesburg and Durban. As a result of the success of this study the National Transport Commission (now SANRAL) decided in June 1985 to expand the Traffic Counting Network to traffic counting stations (Mikros 2006, p iv). SANRAL now repeats this work annually and a reasonable degree of stability in the work process has been achieved since the 1990s. The main objective of SANRAL’s efforts is not to develop freight flows, but to understand congestion points and enable planning in terms of all vehicular road usage in South Africa (of which heavy vehicles are only a subset). No attempt has ever been made to use the information for freight flow estimation purposes, and a reasonable amount of modelling of the available data is therefore necessary.

The SANRAL Traffic Counting Yearbook is a compendium of traffic information obtained at traffic counting stations on primary roads, which highlights the latest available traffic characteristics. Both permanent and secondary stations are utilised. A permanent station makes continuous traffic observations; a secondary station is one where traffic observations are made on a sampling basis for at least 168 consecutive hours per annum. The 2006 SANRAL Traffic Counting Yearbook contains information on 398 permanent stations and 430 secondary stations. However, the position and number of the secondary stations change from year to year – an issue that the modelling process should consider. The stations are placed on

---

**Table 1 Definition of transport typologies (Havenga 2007: 146)**

<table>
<thead>
<tr>
<th>Traffic type</th>
<th>Corridor</th>
<th>Rural</th>
<th>Metropolitan</th>
<th>Rail export machines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mostly manufacturing</td>
<td>Mostly agriculture</td>
<td>Mostly final delivery</td>
<td>Bulk, low-value (export coal and iron ore)</td>
<td></td>
</tr>
<tr>
<td>Some agriculture</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>Long and short</td>
<td>Medium and short</td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td>Origin-Destination</td>
<td>Few long distance ODs</td>
<td>Many</td>
<td>Few, one-directional</td>
<td></td>
</tr>
<tr>
<td>pairs</td>
<td>Many ODs at end-points</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major challenge</td>
<td>Spatial organisation</td>
<td>Development</td>
<td>Congestion alleviation</td>
<td>Global competitiveness</td>
</tr>
<tr>
<td>Efficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistics approach</td>
<td>Intermodal solutions</td>
<td>Effective road</td>
<td>Effective freight delivery</td>
<td>Ring-fenced, bulk rail systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>feeder system</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The average daily truck traffic (ADTT) SANRAL allocates counting stations to traffic counts: used to estimate road freight flows from per annum are shown in Table 2. The number of permanent counting stations Calculation of road tonnages The number of permanent counting stations per annum are shown in Table 2. The following modelling approach is used to estimate road freight flows from traffic counts: SANRAL allocates counting stations to specific routes across the country, e.g. N1, R30, etc in a geographical order. For example, the N1 stations start in Cape Town and end in Beibridge. The average daily truck traffic (ADTT) (i.e. the number of trucks) and the percentage split of these trucks between short, medium and long trucks (SMLT) are captured at each counting station per route from SANRAL’s data. The SANRAL counting methodology enables the counting monitors at the stations to distinguish between trucks and passenger vehicles by the length of the vehicle. The trucks can then be categorised further in terms of axles. This provides a good approximation of vehicle types. ADTT is the total number of trucks observed in each direction during the actual period monitored, divided by the total number of hours monitored, multiplied by 24. SMLT percentage split means the percentage of trucks in each direction which fall into each of the following categories: A short truck is typically a rigid-chassis, two-axle vehicle designed for the transport of goods, or a bus with at least one of its axles bearing four wheels. A medium truck is typically a truck-tractor, plus semi-trailer combination. A long truck is typically a combination of a truck-tractor plus a semi-trailer and a full trailer. The indicated split is established from the combination of measurements of vehicle length and chassis height, as follows: A vehicle shorter than 4.6 m is always regarded as a light vehicle, not a truck. A vehicle between 4.6 m and 11 m long is classified as a short truck if the counting signal indicating the chassis height is ‘high’. (If the signal indicates a ‘medium’ or ‘low’ chassis height, then the vehicle is considered to be a long, light vehicle, e.g. a car towing a caravan.) A vehicle between 11 m and 16.8 m is classified as a medium truck, irrespective of the chassis height. A vehicle longer than 16.8 m is classified as a long truck, irrespective of the chassis height (Mikros 2005, p.10). The number of trucks had to be translated into the actual weight of the freight. For this purpose two figures were calculated – the gross weight (i.e. truck + freight) and the truck tare (i.e. truck weight only). The difference between the two figures is the weight of the freight itself. The average total ton per SMLT was calculated by multiplying the truck mass by the number of trucks. The tare for SMLT was then calculated based on the average tare per vehicle type as published by the RFA (see Table 3). For the national routes, counting stations were depicted graphically to determine the split between metropolitan peaks, rural traffic and long-distance (corridor) traffic. The assumption is that corridor traffic is indicated when a levelling of traffic counts occurs, i.e. when metropolitan peaks taper off and traffic counts remain consistent over a number of consecutive counting stations, while other stations are either metropolitan or rural, depending on their count size and location: Corridor traffic: the stations where levelling occurred were allocated to national routes. The average of the annual weight for all the counting stations per corridor was calculated to reflect the tonnage per corridor. Metropolitan traffic: the key metropolitan areas were identified (through sharp peaks in traffic counts). Different routes lead into these metropolitan areas. For each route the annual average was calculated. The annual totals per route were aggregated to obtain the metropolitan traffic total. The remaining stations were allocated to rural traffic. In the same way as metropolitan traffic (where a group of different routes combine to form a metropolitan freight market segment), each province has a number of rural routes that also combine to form a rural freight market segment. For each route, the annual average was calculated. The annual totals per route were aggregated to obtain the rural traffic total. Extreme outliers were discussed with SANRAL who indicated problems at these counting stations – these outliers were therefore excluded from the analysis (i.e. counting stations which gave figures that were significantly higher or lower than other stations on the same route, and often significantly different from a series of adjacent stations, also clearly visible in the graphs). For illustrative purposes, Figures 2 and 3 graphically depict the Cape Town–Johannesburg and Durban–Johannesburg routes for 2006. (All the SANRAL routes are depicted in this fashion to enable allocation of counting stations).

### Table 2 Number of counting stations analysed per year

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>346</td>
</tr>
<tr>
<td>1993</td>
<td>368</td>
</tr>
<tr>
<td>1997</td>
<td>239</td>
</tr>
<tr>
<td>2003</td>
<td>626</td>
</tr>
<tr>
<td>2004</td>
<td>583</td>
</tr>
<tr>
<td>2005</td>
<td>521</td>
</tr>
<tr>
<td>2006</td>
<td>398</td>
</tr>
</tbody>
</table>

### Table 3 Average gross weight and tare per truck type (calculated from Road Freight Association data)

<table>
<thead>
<tr>
<th>Truck type</th>
<th>Average gross weight (tons)</th>
<th>Tare (tons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short</td>
<td>18.8</td>
<td>5.1</td>
</tr>
<tr>
<td>Medium</td>
<td>42.4</td>
<td>14.1</td>
</tr>
<tr>
<td>Long</td>
<td>51.6</td>
<td>18.2</td>
</tr>
</tbody>
</table>

### Notes
- **ADTT** is the total number of trucks observed in each direction during the actual period monitored, divided by the length of the vehicle. The trucks can then be categorised further in terms of axles. This provides a good approximation of vehicle types.
- **SANRAL** allocates counting stations to specific routes across the country, e.g. N1, R30, etc in a geographical order. For example, the N1 stations start in Cape Town and end in Beibridge.
- The number of permanent counting stations per annum are shown in Table 2.
- The following modelling approach is used to estimate road freight flows from traffic counts:
  - SANRAL allocates counting stations to specific routes across the country, e.g. N1, R30, etc in a geographical order. For example, the N1 stations start in Cape Town and end in Beibridge.
  - The average daily truck traffic (ADTT) (i.e. the number of trucks) and the percentage split of these trucks between short, medium and long trucks (SMLT) are captured at each counting station per route from SANRAL’s data. The SANRAL counting methodology enables the counting monitors at the stations to distinguish between trucks and passenger vehicles by the length of the vehicle. The trucks can then be categorised further in terms of axles. This provides a good approximation of vehicle types.
  - ADTT is the total number of trucks observed in each direction during the actual period monitored, divided by the total number of hours monitored, multiplied by 24.
  - SMLT percentage split means the percentage of trucks in each direction which fall into each of the following categories:
    - A short truck is typically a rigid-chassis, two-axle vehicle designed for the transport of goods, or a bus with at least one of its axles bearing four wheels.
    - A medium truck is typically a truck-tractor, plus semi-trailer combination.
    - A long truck is typically a combination of a truck-tractor plus a semi-trailer and a full trailer.
  - The indicated split is established from the combination of measurements of vehicle length and chassis height, as follows:
    - A vehicle shorter than 4.6 m is always regarded as a light vehicle, not a truck.
    - A vehicle between 4.6 m and 11 m long is classified as a short truck if the counting signal indicating the chassis height is ‘high’. (If the signal indicates a ‘medium’ or ‘low’ chassis height, then the vehicle is considered to be a long, light vehicle, e.g. a car towing a caravan.)
    - A vehicle between 11 m and 16.8 m is classified as a medium truck, irrespective of the chassis height.
    - A vehicle longer than 16.8 m is classified as a long truck, irrespective of the chassis height (Mikros 2005, p.10).
  - The number of trucks had to be translated into the actual weight of the freight. For this purpose two figures were calculated – the gross weight (i.e. truck + freight) and the truck tare (i.e. truck weight only). The difference between the two figures is the weight of the freight itself.
  - The average total ton per SMLT was calculated from Road Freight Association (RFA) data. The total weight per SMLT was calculated by multiplying the truck mass by the number of trucks. The tare for SMLT was then calculated based on the average tare per vehicle type as published by the RFA (see Table 3).
Geographical aggregation of corridors, metropolitan and rural areas

The defined freight market sub-segments were grouped geographically to facilitate analysis, reporting and the development of recommendations. The two main corridors, i.e. Gauteng–Durban and Gauteng–Cape Town, deemed to carry the heaviest traffic, were however kept separate. The corridor grouping is shown in Table 4, while the metropolitan and rural groupings are shown in Table 5.

National freight flows

The results from the NFFM were tabulated to provide modal market share per corridor, rural area and metropolitan area in South Africa. The methodology was applied to counting data from counting stations and actual Transnet Freight Rail data for 1993, 1997, 2003, 2004, 2005 and 2006. This application would, therefore, indicate (i) whether a comparable link with the previous sporadic surveys of Verburgh, Smith, Hamilton and Pretorius, which ended in 1990, could be established; (ii) whether the data from the late 1990s (obtained using this model) up to 2006 has a reasonable correlation with GDP; and (iii) whether the 2003 to 2006 data (when the model was applied annually) seems stable and useful enough to serve as the basis for an annual model. The results from the modelling exercise are reported in the next section.
Total freight flows in South Africa amounted to 1 493 million tons in 2006. The trends for road and rail between 1993 and 2006 are depicted in Figure 4.

Irrespective of the flaws of previous surveys, the definite downward trend in rail market share is clear. Unfortunately, because of the flaws of previous surveys and the absence of flow data, the specifics of this decline have never really been understood.

Growth rates for constant GDP at 2000 prices, physical production in the economy and tons transported can now be compared between the NFFM and the previous studies analysed. This is shown in Table 6.
The survey methods followed by Verburgh, Hamilton, Smith and Pretorius are essentially the same and could, therefore, serve as a good base for comparison.

The Verburgh-Smith-Hamilton time series performs satisfactorily in this comparison, as the growth in tons transported correlates well with the growth in GDP. The faster growth in the physical volume of production in the economy (compared to GDP and transport growth), however, could not be explained by the commissioning of South Africa’s two export machines in the 1970s, namely the Richards Bay coal line and the Sishen to Saldanha iron ore line, because in both these cases the results should also have translated into higher transport volumes. It is also hypothesised that more double handling of goods in a more mature economy (caused by specialisation) should occur, which means that tons transported should probably grow somewhat faster than the physical volume of production in the economy.

The Pretorius time series did not do well in these comparisons. Constant GDP growth in South Africa was slow in the years just before the legitimisation of the ANC and the release of Nelson Mandela, and the physical volume of production slowed to a compound annual growth rate of just 0.61%, but a negative correlation of more than 3% of tons transported is highly unlikely.

The NFFM correlation performs well in all comparisons. Constant GDP grows faster than the physical volume of production as the economy matures and the tertiary sector expands. The expected increase in double handling results in a faster transport growth rate than that of physical production, but is more in line with GDP growth.

A more precise measure of correlation would be to measure the correlation coefficient for the various data sets. For this calculation, for tons transported, the four annual data points from Verburgh, Smith and Hamilton’s two surveys were used, the six data points from Pretorius’s surveys between 1985 and 1990, and four data points from the NFFM between 2003 and 2006. This was correlated with GDP and physical production for the same time periods. The calculation is shown in Table 7. The results from the NFFM consistently perform significantly better than those of previous surveys.

The data has been assessed at workshops by peers at the Council for Scientific and Industrial Research, Transnet, the Department of Trade and Industry, the Department of Agriculture and Transnet Freight Rail. Current indications are that the high correlation will continue in future applications of the model. As the work continues, the academic soundness of the research should be continually tested and reported on to determine if the positive correlation is valid for longer time series.

**Freight flow market segments and the current distribution of freight**

In South Africa there are 329 billion ton-kilometres of freight movement per annum (National Freight Flow Model estimates), compared to approximately 15 000 billion ton-kilometres worldwide (Rodrigue 2007). South Africa has just less than 1% of the world’s population, produces less than 0.4% of the world’s GDP, and yet it requires more

---

**Table 7 Correlation coefficient comparison for all surveys**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP to physical production</td>
<td>0.992</td>
<td>0.935</td>
<td>0.986</td>
</tr>
<tr>
<td>GDP to tons transported</td>
<td>0.627</td>
<td>-0.811</td>
<td>0.996</td>
</tr>
<tr>
<td>Physical production to tons transported</td>
<td>0.653</td>
<td>-0.720</td>
<td>0.985</td>
</tr>
<tr>
<td>GDP to tons transported: known errors in Hamilton and Pretorius surveys removed</td>
<td>0.933</td>
<td>-0.868</td>
<td>0.996</td>
</tr>
<tr>
<td>Physical production to tons transported: known errors in Hamilton and Pretorius surveys removed</td>
<td>0.874</td>
<td>-0.794</td>
<td>0.985</td>
</tr>
</tbody>
</table>
less than 2% of global freight transport in terms of ton-kilometres. Therefore, South Africa’s transport demand is excessive in the light of these indicators.

This situation has arisen historically from the country’s economic development, which has resulted in:

- centres of production and population density far from the coastal areas
- a relatively open mineral-export economy, and
- a beneficiated product- and energy-import economy.

These factors consequently necessitate long export and import corridors.

This situation places, *ipso facto*, a massive burden on South Africa’s transport infrastructure, and, because of the country’s poor productivity, a specific need for excessively cheap transport. The South African economy is still relatively primary, especially when compared to developed economies, and a better understanding of the specifics of the freight flow market segments is required in order to formulate an appropriate freight transport strategy.

The 329 billion ton-kilometres of freight is divided into the four freight flow market segments as depicted in Figure 6.

The tonnage growth across the various freight flow market segments is depicted in Figure 7.

The stagnation in rail is clear. All growth over dense corridors occurred within the road mode, which expanded by more than 70% over 13 years. This growth would be understandable if the corridors in question were short or the density per corridor low. In these instances the economy will have to absorb this growth in the road mode. Cheaper options are, however, available in intermodality, if the density per corridor can be calculated as sufficiently high. In South Africa’s case, as observed by the national freight flow model, the spatial efficiency objective of the corridor freight market segment is not achieved. If this density is sufficient to entertain an intermodal solution, future investment should be considered in such solutions. This could release funds for the development of the secondary economy in rural areas, which is a major objective of the country at the moment.

Rural road traffic has grown more slowly than corridor traffic (less than 60% growth), which supports the hypothesis that South Africa is not succeeding in the stimulation of rural economies as desired. A major cause of this failure is declining road infrastructure, the impact of which can now be measured for the first time. Metropolitan growth is also slower than corridor growth.

The detailed results per freight market segment will point towards specific issues and possible solutions. These results are discussed in the following sections.

**Trends in movement over the various corridors**

The various corridors have performed differently in terms of growth over the past 13 years. The growth is depicted in Figure 8.

Growth on the densest road freight corridor – the route between Gauteng and Durban – was relatively slow (26%), but this is also the corridor that is probably the most overstretched in the country because...
Figure 8 Corridor growth from 1993 to 2006

Figure 9 Metropolitan freight transport growth between 1993 and 2006
of current density. Alternative routes (even as far asfield as the Energy–Demoina route, which is the alternative, much longer route between Gauteng and Durban over Piet Retief) are often used because of, inter alia, lack of policing, lower density and the presence of fewer toll roads. From this high base, if the road/rail market-share position should continue, major long-haul congestion problems will arise in the future.

The eastern corridors describe the roads to Witbank, Nelspruit and Maputo from Gauteng. There has been a major initiative over the past few years to develop this corridor, but freight has grown by only 46% over the time period (the second lowest – and slower than economic growth). Furthermore, the expected growth in rail traffic has not yet been realised.

The Gauteng to Cape Town corridor achieved 135% growth over the time period – almost all on road. This is perhaps the greatest error in South Africa’s infrastructure planning framework, as this is also the longest corridor and should, by any standard, be more rail-bound than the rest.

**Trends in movement in the various metropolitan areas**

Almost all known metropolitan areas (or areas usually classified as metropolitan) experienced growth of between 50 and 60% over the period, as depicted in Figure 9.

This pattern of rail freight decline is normal for the shorter distances involved and may seem acceptable to planners, but it should be remembered that dense metropolitan areas are growing from an already high base (given the current available infrastructure that is installed and planned). In addition, 86% of all metropolitan freight in South Africa moves within three metropolitan areas, and 41% of all freight shipped in South Africa originates from just three metropolitan areas. In very specific cases rail solutions are possible to alleviate pressure on metropolitan infrastructure. An example includes the bale-by-rail solution that was developed for Vissershok in Cape Town, whereby a ring-fenced fleet delivers baled waste over defined and limited origin-destination pairs to a landfill site outside Cape Town.

**Developing transport performance measures**

The application of the NFFM as lead indicator has now been adequately defined, given the specific planning information for the freight flow market segments that was generated. Possibilities also exist to use the data to develop lag indicators that can measure the performance of the economy in terms of transport consumed. GDP data, physical production data and sectoral GDP are known and a calculation is, therefore, possible (see Figure 11).

The time frame over which the performance of the NFFM is measured is short (with only five data points), and the previous measurements erratic and unreliable. The increasingly poor performance of the economy in terms of transport productivity is, however, clearly visible where the economy generated less than R800 for each ton transported in 2005. This measure over time could inform the spatial performance of the economy (about which only hypotheses have existed up to now) and contribute towards a better understanding of the spatial dilemma. As the tertiary sector in a mature economy grows, the measure should actually improve and, furthermore, improve with increased productivity in transport. This is, however, not the case in South Africa.

---

**Figure 10 Rural freight transport growth between 1993 and 2006**

<table>
<thead>
<tr>
<th>Year</th>
<th>1993</th>
<th>1997</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Cape</td>
<td>9%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>Northern Cape</td>
<td>21%</td>
<td>21%</td>
<td>21%</td>
<td>21%</td>
<td>21%</td>
<td>21%</td>
</tr>
<tr>
<td>Eastern Cape</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
<td>14%</td>
</tr>
<tr>
<td>North West</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Limpopo</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>

**(a) Road**

*Figure 10* Rural freight transport growth between 1993 and 2006

<table>
<thead>
<tr>
<th>Year</th>
<th>1993</th>
<th>1997</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gauteng</td>
<td>24%</td>
<td>24%</td>
<td>24%</td>
<td>24%</td>
<td>24%</td>
<td>24%</td>
</tr>
<tr>
<td>Free State</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>KwaZulu-Natal</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Northern Cape</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Limpopo</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>Mpumalanga</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
</tr>
</tbody>
</table>

**(b) Rail**

*Figure 10* Rural freight transport growth between 1993 and 2006
The measure can be extended to indicate measures for the transportable economy (see Figure 12). This measure also indicates a declining trend and the poor performance of spatial reorganisation, spatial requirements and transport productivity in South Africa.

**CONCLUSIONS**

This was the first successful attempt to develop annually repeatable flow data for South Africa, and the results provide the most complete picture of surface freight flows yet in the country, including modal market share. The data was compared with other economic data and with the Verburgh, Smith, Hamilton and Pretorius studies and performed well on all accounts.

The results have also been applied to initiate the process of generating economic performance data for transport. The lead indicators (namely, modal market share in total and correlation of freight transport with GDP per freight market segment), which could inform transport and logistics infrastructure planning, and the lag indicators (namely, performance of GDP and transportable GDP per ton), which measure past performance of transport in the economy, have now been defined.

The results indicate that long-distance corridor growth is being captured by road transport. The spatial efficiency objective of the corridor freight market segment is therefore not being achieved and the development of intermodal solutions should be paramount. This could release funds for the development of transport infrastructure in rural areas to support the development of the second economy. In addition, more sustainable rail solutions for specific ring-fenced metropolitan flows should be investigated.

The NFFM also provides a basis for forecasting. The existence of a volumetric measurement means that associated logistics costs can be developed, which will inform the affordability of transport infrastructure, as well as determine the degree to which economic infrastructure supports or dampens South Africa's global competitiveness. South Africa is also in the process of developing frameworks for freight transport regulation, and volumetric data on freight flows will be invaluable in the process of establishing a freight regulation framework.

**NOTES**

1. This means, for example, that if a specific load originated in Caledon in the Western Cape and was sent to Johannesburg, it will be observed by the model as a rural Western Cape load and a load that used the Cape Town to Gauteng corridor. In another example, if a load originated in Durban and was sent to Beitbridge, it will be observed as using both the Durban to Gauteng and Gauteng to Beitbridge corridors.

2. The current counting technology does not allow distinguishing between light delivery vehicles (colloquially ‘vans’ or ‘bakkies’) and passenger vehicles. The model results therefore exclude freight transport that is conducted in vehicles less than small truck size.

3. The current counting technology only allows the distinguishing of bus counts at toll stations, not at other counting points. Bus counts at toll stations account for approximately 1% of traffic counts, and are aggregated with heavy vehicle counts in the SANDRAL data. Currently no adjustment to the NFFM is made for this, as bus counts on other routes are probably significantly less than 1%.


**REFERENCES**


Hamilton, C C 1986. The estimated market share of public and ancillary road carriers in a specifically defined South African goods transport market.
Pretoria: CSIR National Institute for Transport and Road Research.
Challenges confronting road freight transport and the use of vehicle-pavement interaction analysis in addressing these challenges

W J vd M Steyn, C L Monismith, W A Nokes, J T Harvey, T J Holland, N Burmas

Traditional arguments for maintaining riding quality of pavement are expanded in this paper to examine the effects of deteriorating riding quality on vehicle operating costs, freight damage and logistics. The objectives of this paper are to analyse the effects of different levels of riding quality on a truck and its freight, and to discuss potential applications of the analysis in terms of effectiveness of the freight transport system. The paper discusses needs and drivers influencing freight transport costs, vehicle-pavement interaction concepts, and the potential physical effects and costs from roads with deteriorating riding quality. A case study is presented analysing vehicle-pavement interaction for selected roadways in California. It is concluded that investments in pavement and freight transport industry improvements can be investigated by applying vehicle-pavement interaction analysis to evaluate damage to pavement, vehicle and freight that would result from alternative levels of pavement riding quality. The paper recommends that existing concepts, tools and resources such as dedicated truck lanes and vehicle-pavement interaction analysis can help to improve the freight transport system. A framework is proposed to better understand the scale of potential impacts of riding quality from localised effects to larger-scale influences, including costs to customers and global competitiveness.
INTRODUCTION
Highways are the primary means of transporting people and freight in the United States (US) (Lambert 2003). A reliable land transportation system is essential to a robust economy (Transportation Research Board (TRB) 1994). Issues and concerns about ways to improve the transportation system are broad, complex, and controversial (Burks et al 2010). This paper examines the effects on freight transport vehicles, freight damage and logistics, and pavement infrastructure based on quantitative measures of Vehicle-Pavement Interaction (V-PI). Traditional arguments for maintaining riding quality of pavement are expanded in this paper to examine the effects of deteriorating riding quality on Vehicle Operating Costs (VoCs), freight damage and logistics, which have downstream impacts on costs of goods and ultimately on a nation’s economy.

The objectives of this paper are to analyse the effects of different levels of riding quality on a standard US truck (from dynamic loads applied to the pavement) and its freight, and to discuss potential applications of the analysis in terms of effectiveness of the freight transport system and implications.

The paper starts with an evaluation of the needs and drivers that influence freight transport costs, and reviews transport infrastructure and freight statistics in the US and California. V-PI concepts are introduced and the potential physical effects of deteriorating riding quality are discussed. This is followed by an evaluation of the costs of roads with low riding quality and a case study V-PI analysis for two roads in California. A framework is proposed to better understand the scale of potential impacts of riding quality from localised effects to the broader economy and global competitiveness. Recommendations are made to expand the scope of the process to more countries and vehicle types.

NEEDS AND DRIVERS OF FREIGHT TRANSPORT COSTS
The major needs and drivers influencing freight transport costs on a road network are shown in Table 1.

Infrastructure
Freight transportation is the backbone of US commerce with imports and exports typically doubling every ten years (CAGTC 2008). The US transport system serves seven million businesses that depend on it to move goods to markets, mainly using trucks (Bureau of Transportation Statistics (BTS) 2007). Transport costs represented 10.1% of the Gross Domestic Product (GDP) of the US in 2007 (Council of Supply Chain Management Professionals (CSCMP) 2008).

The US has approximately 6.4 million km of highways, representing 38% of the transportation capital stock in 2005 (BTS 2007; US Climate Change Science Program (USCCSP) 2007).

The overall condition of US interstate highways generally improved between 1995 and 2005, although rural and urban collectors and urban minor arterials showed a higher percentage of roads in poor or mediocre condition (BTS 2007). Highway system condition has fallen behind escalating use, creating traffic gridlock and delayed business deliveries. Over 256 000 km of the National Highway System need resurfacing or reconstruction (CAGTC 2008).

California’s State Highway System (SHS) consists of around 24 200 centre-line-km (479 000 lane-km) with 79% classified as freeways and expressways (California Highways 2011). In 2007 the California Department of Transportation (Caltrans) classified 25% of the SHS as being in Total Distressed Pavement condition (low riding quality and/or structural distress), 32% as requiring Corrective or Preventive Maintenance, and 41% as being in Excellent condition (Caltrans 2008).

California had six of the ten roughest urban road networks in the US in 2004, causing road users in California’s biggest cities to pay as much as $700 annually for repairs and early vehicle replacement costs, compared to the national average of $400 (Lawson 2006). It is estimated that keeping roads from getting worse would require a 32% increase in annual funding, while a 62% increase is required to improve them.

Without timely repairs the road network will require major rehabilitation and reconstruction that will increase costs by a factor of four to five over the current maintenance costs. About half of motorists’ rough road tax is spent on tyre replacement, wasted fuel and repairs to suspension systems, while the other half is a portion of the costs for having to replace vehicles more frequently because of the condition of roads (Lawson 2006).

Freight
International trade increased steadily in recent decades rising about 12% (as percentage of GDP) in the 1970s to about 25% in the

Table 1 Summary of needs and drivers influencing freight transport cost on a road network

<table>
<thead>
<tr>
<th>Needs</th>
<th>Drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>Inefficient traffic</td>
</tr>
<tr>
<td>Vehicles</td>
<td>Higher fuel and repairs to suspension systems</td>
</tr>
<tr>
<td>Freight</td>
<td>Economic movement without damage</td>
</tr>
<tr>
<td>Congestion</td>
<td>Demand for various types of freight</td>
</tr>
<tr>
<td>Logistics</td>
<td>Lowest total costs</td>
</tr>
<tr>
<td>Energy</td>
<td>Energy cost</td>
</tr>
<tr>
<td>Environment</td>
<td>Environmental impacts</td>
</tr>
</tbody>
</table>

Pavement riding quality is used as one of the primary indicators of the condition of the pavement (Sayers et al 1996). Deteriorating riding quality translates into negative impacts on road users (motorists and truckers) and agencies that own/operate the roads.

Vehicles
Various studies show that pavement riding quality affects the vibrations and responses in vehicles (Singh et al 1991; Slaughter et al 1998; Steyn & Visser 2001). The main conclusion from all these studies is that lower riding quality is a major cause of increased vibrations and subsequent structural damage to vehicles. Studies indicate that more heavily loaded trucks showed higher lateral and longitudinal levels of vibration than lightly loaded ones. Trailer vibration levels typically decrease with reduced speed over rough surfaces.

In a continuing study that forms part of the South African State of the Logistics effort (Steyn et al 2011), a database of VoCs for an operational fleet of 577 trucks indicated that trucks travelling on roads with poor riding conditions experienced increased suspension replacement costs of between 68% and 1 560%.

Short-term highway maintenance cost savings obtained by postponing pavement rehabilitation are realised by the transportation agency and do not include higher vehicle costs from wear-and-tear caused by pavements with lower riding quality. Life-time VoCs are typically four times the initial construction costs of a highway, while highway maintenance is only 1 to 2% of the total road transport costs. Neglect of highway maintenance can cause an increase of up to 15% in VoCs, while further neglect of maintenance will cause a paved road to start disintegrating with further increases of up to 50% in VoCs (Robinson 1987).

Freight
International trade increased steadily in recent decades rising about 12% (as percentage of GDP) in the 1970s to about 25% in the
1990s. Explosive growth, improvements in manufacturing processes and new technology are continuing this trend, increasing the strain on trade gateways (CAGTC 2008). More than 40% of containers moving through the US are transported through California (McKim 2011), with an average annual growth rate of 9.4% since 2001 (BTS 2007). The trucking industry is a crucial part of the modal mix, particularly with trucks typically acting as the “last mile” service provider, bringing goods directly to consumers (USCCSP 2007).

The agricultural sector accounts for nearly 30% of all freight transportation services in the US, making it the largest single user of the US and California freight transportation networks (FTA 2008).

Studies on the effects of freight vibration of fruit caused by low riding quality found that differences between the levels of vibration-induced damage during transport may be attributed to mechanical differences between trailers, as well as road condition. Fruit damage was found to be greatest in the uppermost container for every combination of road, truck type and travelling speed, which also corresponded to the highest vibration levels measured in the container (Slaughter et al. 1998; Bundit et al. 2005). The pattern of adverse effects on freight caused by pavements with low riding quality is not expected to be restricted to the agricultural sector and could have wide-ranging cost impacts, e.g. vibration effects on sensitive electronic components in freight.

Testimony to the Senate Committee on Agriculture, Nutrition and Forestry in 2005 observed that, “This U.S. transportation system is turning from a strength into a potential weakness. Because of higher fuel and energy costs, congestion on railroads and highways, a lack of investment in modernisation and maintenance of the inland waterway system, the costs of moving agricultural products to markets is escalating sharply.” (CANF 2005). Also, the American Association of State Highway and Transportation Officials (AASHTO) cited freight movement in its ten top priorities for 2011 (AASHTO 2010). Failure to invest in freight transportation will lead to increases in congestion, slower movement of goods, diminished capabilities to access critical international markets and increases in consumer prices.

**Congestion**

Road congestion costs the US economy around $63.1 billion in 2001. Current road congestion levels, budget limitations, effects of road condition on transport costs and forecasts of dramatic traffic volume increases all suggest that the highway system will fail to meet demand in the near future (CEC 2011). The entire US vehicle fleet logged nearly 4.4 trillion km in 2000 representing a 22.4% increase in total annual Vehicle Kms Travelled (VKT) since 1992 (FHWA 2003). The steady annual increase leads to an expected 6.48 trillion VKT in 2020 (47% increase relative to 2000).

The California Department of Finance projects a 31% population increase (CDOF 2002), which, when combined with Federal Highway Administration (FHWA) data (FHWA 2003), suggests a possible traffic volume increase in California of about 83% between 2000 and 2020. The growth of truck traffic volume has outpaced that of passenger vehicles (Lambert 2003), which is significant, because service lifespan of pavements is related to truck traffic (Samuel et al. 2002).

**Logistics**

Road transport logistics consists of planning and implementing the acquisition and use of resources for freight movement on roads (Cass Logistics 2000). More than 40% of the difference in economic performance of low- and high-growth countries can be explained by differences in the effective use of infrastructure (Rulistia 2008). This has significant importance for the economic development of a country such as South Africa. Deteriorating roads lead to delays, higher costs for road users and taxpayers, vehicle and freight damage and greater risks of accidents.

Logistics costs in the US changed from 12.3% of GDP in 1985 to an all-time low of 8.6% of GDP in 2003, compared to 13.5% of GDP in South Africa in 2009 (SOL, i.e. State of Logistics) (CSIR 2011). Total US logistics costs have risen by 32.3% over the last five years. Transportation costs rose 5.9% during 2007 and now account for 6.2% of nominal US GDP, similar to levels experienced 20 years ago (Schulz 2008). A major challenge to supply-chain logistics in the US is highway congestion and port delays combined with the need for a long-term vision in highway infrastructure improvement and sustainability.

**Energy**

Optimal energy use requires the evaluation of traffic flow characteristics, since constant speeds enhance optimal energy use and generally reduce travel times. Exclusive truck-only facilities may assist in improved productivity and lower energy use (Roorda et al. 2010). The transportation sector used 19% more energy in 2006 than in 1995, consuming 67.9% of US petroleum usage in 2006 (BTS 2007).

Long-term pavement studies at the WesTrack test facility found increasing pavement roughness increased fuel consumption of trucks by 4.5% (under otherwise identical conditions). Higher roughness also increased the frequency of fatigue failures of truck and trailer components (Sime et al. 2000).

US transportation energy use is expected to increase by 46% from 2003 to 2025, mainly due to economic growth. The most widely proposed fuel options are increased vehicle fuel economy, increased prices for carbon-based fuels, and developing alternative and renewable fuel sources (USCCSP 2007).

**Environment**

The environment links directly to energy and logistics. Improved logistics generally lead to lower energy demands and environmental impacts. The US freight transportation network is the second-largest source of GreenHouse Gas (GHG) emissions in the US, after electricity generation. US road transport contributes approximately 18% of total global CO₂ emissions and around 14% of GHG emissions, making it one of the three major contributors to these emissions (BTS 2007).

While emissions from light-duty vehicles are expected to drop 12% by 2030, freight truck emissions are projected to increase 20%. Truck GHG mitigation strategies include recommendations such as system optimisation and operational efficiency improvements through decentralised supply chains, less excess packaging, reduced shipment frequency, congestion mitigation measures and advanced vehicle technology (CEC 2011).

Reductions are possible by improving road network performance. A Norwegian study concluded that better road alignment, coupled with sufficient width and infrastructure capacity, led to decreases of between 11 and 61% in CO₂ emissions (IRF 2007).

**V-PI AND RIDING QUALITY**

V-PI describes the vehicle components, pavement components, and the ways in which these influence one another. Pavement riding quality influences the vertical dynamics of the moving vehicle. The resulting dynamic tyre loads cause pavement responses which lead to pavement distress. Under repeated loading of heavy vehicles the profile changes further, leading to further changes in dynamic tyre loads. The pavement structural strength reduces over time, due mainly to cumulative effects on pavement responses. The net effect on pavement distress will depend on the spatial relationship between dynamic tyre loads, pavement profile...
changes and pavement structural strength variations (DIVINE 1997).

Pavement roughness indices provide quantified riding quality values and trends. The International Roughness Index (IRI) is widely accepted as the index of choice for reporting pavement roughness. IRI is most appropriate when a roughness measure is desired that relates to overall VoC, overall riding quality and overall surface condition (Sayers and Karamihas 1998). It is noteworthy that IRI was intended to reflect pavement roughness attributes that affect riding quality of passenger vehicles, and not intended to describe pavement roughness characteristics affecting heavy trucks, because it does not show sensitivity to excitation frequencies observed under heavy vehicle traffic (Papagiannakis and Gujarathi 1995).

Vehicle response to the pavement profile can be modelled in the frequency domain as a response function. Mathematically, the vehicle frequency response function acts as a multiplier to the input road profile Power Spectral Density (PSD) to give the PSD of the vehicle response. This combined frequency characterisation of road profiles and frequency domain analysis of vehicle responses to the profile has resulted in the road profile being categorised into eight classes (A to H, from smooth to rough), as set by the International Organization for Standardization (ISO) based on PSD (ISO 1995).

In studying the interaction between vehicle and pavement, the objective is to determine how the vehicle and pavement components affect one another in order to determine the resultant effects on the pavement, vehicle components and freight. There are three standard components in the majority of the vehicle-pavement interaction evaluations (DIVINE 1997; Collop & Cebon 1995). They are (1) the cause of the problem (pavement profile), (2) the load history generator (the vehicle and freight) and (3) the component on which the forces are exerted (the pavement structure). This paper focuses on pavement performance and maintenance, vehicle damage and operating costs, and freight damage and logistics.

Pavement performance and maintenance

The performance of pavements over time depends on construction quality, traffic loading and environmental conditions. A rough pavement surface increases moving dynamic loads, resulting in higher than average pavement deterioration, shortening pavement service life. The simplest way to reduce pavement deterioration caused by V-PI is high quality, smooth pavement construction and timely maintenance that ensure high riding quality.

Riding quality directly influences the quality of the driving experience for vehicle occupants and freight. Steyn et al (2011) identified the potential effects of deteriorating riding quality on the broader economy (Figure 1). Decreasing pavement riding quality, and the increased dynamic vehicle loads it induces, cause direct increases in pavement maintenance costs, shortened pavement service life, and ultimately higher costs for a given pavement structure and transported freight.

Driving on a pavement with low riding quality affects the speeds at which the vehicle can travel safely, immediately affecting the logistics of delivering goods at optimum times to consumers. It also raises fuel consumption and increases carbon emissions for the same amount of freight delivered. In an evaluation of road maintenance requirements in sub-Saharan Africa, Heggie (1995) calculated that for each $1 spent on road maintenance, between $3 and $22 VoC is saved. Values in other studies may differ, but the pattern of return on investments in road maintenance is typical.

Vehicle damage, operating costs and environment

Potential damage to vehicles travelling on rough roads can be addressed through improved design of each vehicle, which multiplies the costs to all travelling vehicles.
Low riding quality thus has an increased cost effect on vehicle design, manufacturing and maintenance. All these costs are routinely incorporated into the costs that the vehicle owner charges for transporting freight, the logistical costs and ultimately the costs of goods paid by the consumer.

Lower riding quality affects the environment and environmental costs through increased VoCs, increased emissions due to slower speeds and longer durations of transport, increased use of natural resources, and increased costs as more materials are required to maintain pavements.

**Freight damage and logistics**

Increased vibrations result in more vibrations translated to the freight and resultant freight damage. Various methods exist for establishing the expected damage to freight exposed to ranges of vibrations and shock loads. Potential solutions to this problem include improvements in freight packaging and in the design of the freight itself, but both of these solutions add further costs. The costs of packaging of fresh produce are in the order of 10% of the final costs of the delivered product. Road-induced freight damage prevents transportation of some produce that are prone to damage, such as ripened soft fruits. This adds a cost as these fruits typically are transported at an earlier stage and then stored in warehouses until sold. The high volumes of produce (fruit, as well as nuts) transported in California makes this an issue of state-wide significance.

Increased logistics costs due to decreasing riding quality are experienced by most transportation companies in a country. These costs have cumulative effects and will result in massive increases in logistics costs of the country. These increased logistics costs will translate to higher costs of products in the marketplace.

Freight movement, with its linkages to infrastructure and vehicles as well as implicit role of logistics, is a national priority in the US as Congress develops a new land transportation bill. In California, goods movement underpins this eighth largest world economy and the 14% of the US GDP generated in the state. In US Senate testimony in early 2011, Caltrans Director McKim (2011) affirmed that the economic health of states like California requires establishment of freight movement as a local and national economic priority.

**V-PI CASE STUDY**

In coordination with the University of California Pavement Research Center (UCPRC), V-PI analyses were conducted for a 35-2 (articulated) truck on sections of the I710 and US101 located in Southern California. Route I710 is a heavily trafficked route that carries substantial freight traffic to and from the ports of Long Beach and Los Angeles (Average Daily Traffic (ADT) of 155 000 during weekdays, with 13% trucks). The centreline length of the I710 section is approximately 4.4 km with three lanes in each direction (Monismith et al 2009). The US101 between Santa Barbara and Ventura carries rural traffic and is approximately 13.3 km in length. Road surface profiles were measured on both highways for use in V-PI analyses. A total of 660 V-PI simulations were conducted using TruckSIM™, simulating four speeds and three load levels (10 km/h to 100 km/h; empty to full). Riding quality data was analysed using Provat™. Frequency analysis compared the displacement PSD with the ISO PSD classifications to indicate pavement frequency profiles. The riding quality classifications for both roads ranged between 0.8 and 7.3 m/km with the mean IRI being 2.1 m/km and the 90th percentile value being 3.8 m/km. An IRI-value below 1.5 m/km is viewed as good while a value above 2.7 m/km is viewed as unacceptable (AASHTO 2009). More than 70% of the profiles in this study had at least acceptable riding quality values. Based on the ISO classification of PSD pavement profile data, the majority of the pavement profiles in this study were classified as A (46 segments), with a limited number as B (10 segments) and only three segments as C.

The V-PI output data can be divided into tyre loads and vertical movement data. Eighteen individual tyre loads were generated for the 35-2 truck tyres at frequencies depending on the speed used in the simulation. Changes in individual tyre load are evaluated in terms of standard deviation (STDev) and Coefficient of Variation (CoV) of load data.

Analysis of STDev values showed a skew distribution, with the higher values being related to lower load levels and higher pavement roughness. Analysis of axle group load data indicated that the mean tyre loads were related to the static load levels and were not affected by the pavement roughness or speed. STDev values all increased with increasing pavement roughness, decreasing loads and increasing speeds. CoV values were related to the pavement roughness, although the combination of a fully loaded vehicle on a rough pavement (relatively high STDev) can provide a lower CoV than for an empty vehicle on a smoother pavement (relatively low STDev). When analysing specific load cases a similar pattern appears where lower load levels and higher speeds and pavement roughness lead to higher CoV values. These findings are consistent with previous studies (Steyn & Visser 2001).

Vertical acceleration and displacement movement of the Centre of Gravity (CoG) were evaluated to determine potential freight and vehicle damage. The STDev of the CoG's vertical acceleration increased directly with increasing pavement roughness for all speeds and loads, for both the truck-tractor and semi-trailer. The STDev of the CoG's vertical accelerations were highest for the empty semi-trailer simulations. The half- and fully-loaded simulations showed much lower STDev values at all levels of speed and pavement roughness. STDev values show a dramatic difference in vertical acceleration of the truck-tractor and the semi-trailer, with STDev values for the semi-trailer nearly twice as high as the truck-tractor in most cases.

The STDev of the vertical movement of the CoG of the truck-tractor and the semi-trailer generally increases with increasing pavement roughness, with a marked increase at riding quality values above 4.2 m/km.

Findings from this case study are consistent with the literature. Lower riding quality causes increased STDev of tyre loads, with higher speeds and lower masses causing further increases. The vertical acceleration and movement of the vehicle's CoG are affected in a similar way.

**DISCUSSION**

This paper examines linkages between diverse aspects of the roadway transportation system with regard to freight movement. Having identified the needs and drivers that affect freight transport costs, the paper discusses damage and its costs to pavement infrastructure, vehicles, and freight. Damage to all three results in higher costs of products to consumers. Investing in improvements to one or more of these aspects ultimately should reduce costs to consumers. Providing lower cost goods also improves a state's and nation's competitiveness in today's global marketplace.

This paper focuses on improving the pavement infrastructure through an improved understanding and quantification of V-PI effects. V-PI analyses enable better understanding of pavement roughness impacts on vehicles and freight. The roles and interdependence of pavement, vehicles and freight are complex. On-going freight mobility and related planning efforts in California reflect technical and organisational complexities, which are influenced by integrating different transport modes, addressing land use issues and working with many diverse (sometimes competing)
goals, stakeholders, and decision-makers (Scherzinger 2011).

A context for discussion of V-PI and freight issues is presented in a three-level framework in Figure 2. The context ranges from localised level (at left in Figure 2) up to state-level and ultimately up to a national/international level. Tasks, aims, and responsibilities differ at each level.

At the local level, vehicle and road owners are responsible for improving V-PI components under their control in order to reduce transportation costs and damage during freight transport. This approach focuses on improvement in the infrastructure for functional as well as economic purposes. Road condition affects all the vehicles running on the road, and therefore investment at this level has a major multiplier effect – much more than investments in improved vehicle components or goods packaging and design. Lower logistics costs also reduce subsequent costs in the economic stream.

The second level focuses on a sustainable and effective system by local, regional and state agencies to ensure less congestion and smoother operations in logistics and transport. Examples include dedicated trucking lanes, life cycle analyses, and evaluation of environmental issues. It also may include more freight-focused legislation and regulations to enable more cost-effective freight movement.

The third and broadest level focuses on federal, national and cross-border responsibilities for evaluating issues of total logistics costs, road vs rail comparisons, and cross-border flow of goods. A more efficient system should increase a nation’s GDP and global competitiveness. An ongoing dialogue on the efficiency and sustainability of the freight transportation system should be maintained, with the focus on developing a long-term vision of sustainable freight transportation for the US. Collaboration should be facilitated on a continental basis among freight industry, transportation experts and stakeholders.

This three-level approach is based on the understanding that the responsible entities at each of the three levels focus on improving their own efficiencies in the tasks shown, thereby strengthening the whole system. The recently published performance measures for freight transport (NCFRP 2011) focused on the development of a scorecard to gauge the performance of the US freight transportation system to support investment, operations, and policy decisions by public and private stakeholders. It covers six major categories (freight demand, efficiency, freight system condition, environmental impacts, safety, and the adequacy of investment in the freight system), with the freight system condition category closest to the focus of this paper.

A high quality freight system condition requires the best possible roads in terms of riding quality that causes low vehicle and freight damage. One of the causes for deteriorating pavement conditions is low quality construction. Designs may focus on high riding quality with minimal deterioration and maintenance, but achieving these design goals relies on high quality construction. Broad implementation of effective bonus/penalty schemes has localised effects on better riding quality, less vehicle and freight damage and their associated costs. These localised improvements benefit other freight performance measures, such as efficiency, safety, and environmental impacts at all three levels within the framework.

Use of dedicated truck lanes (DTLs) is one approach to improving freight system condition and efficiency. Studies focusing on DTLs in specific corridors (including I-10 coast-to-coast, I-81 in Virginia, corridors across Texas, Georgia, Chicago and Los Angeles) addressed opportunities that could arise from use of new vehicle control technologies to enhance the effectiveness of DTLs. Substantial savings in rehabilitation and maintenance costs could be achieved if trucks drove only on lanes designed exclusively for them. Further improvement would result if trucks consistently followed the same path within a lane through automatic lateral control. The pavement structure could be optimised in ways that are impossible for wandering traffic (e.g. thicker pavement on truck lanes, etc.). Longer vehicles and electronically coupled “platooning” of trucks could also improve freight productivity of the highway system (Browand et al 2004), as well as improving safety and congestion (Roorda et al 2010; Samuel et al 2002).

DTLs, freight performance measures, V-PI analysis, maintaining road smoothness, and economic assessments typically are addressed by an array of specialists working in their own respective fields. The analysis, discussion, and framework in this paper illustrate the interconnectedness of these issues and suggests the need for better coordination and a more coherent, multi-specialty approach to improving freight movement.

Potential application of the concepts discussed in this paper in other environments (such as South Africa) is currently investigated through efforts such as the SOL study (CSIR 2011). The differences in infrastructure and vehicle population play a role in the detailed application of the concepts. However, the principles around well-engineered planning, construction, maintenance and management of the logistics chain and all its components are as relevant as in the US. Implementation of the concepts presented in Figure 2 for the South African context should likewise lead to improvements in the local logistics chain and GDP.
CONCLUSIONS AND RECOMMENDATIONS

- The freight transport industry in the US is a significant contributor to state and national economies, with diverse needs and drivers affecting freight transport costs.
- Properties and complex interactions of pavement, vehicles and freight determine the condition of the freight transport system and resultant overall costs.
- V-PI analysis tools are available for quantifying physical response and damage for pavement with known riding quality. Results from a case study in California are consistent with international literature showing increased damage to pavement, vehicles and freight as the pavement condition deteriorates, leading to higher operational and logistics costs. These are ultimately paid by customers, reducing economic competitiveness. These tools can help strategic planning, as well as near-term operational planning, for a wide range of priorities in both the public and private sectors.
- Improved pavement maintenance results in reduced long-term costs for road authorities and induces cost-savings for road users, including freight transport. Improved management of components of the freight transport industry can lead to lower variations in tyre loads and acceleration levels of vehicle components and freight. Investments in these improvements can be investigated by applying V-PI analysis to evaluate response and damage to pavement, vehicle and freight that would result from various alternative levels of pavement riding quality. Results could also provide insights about effects on other drivers of freight transport costs, such as energy costs and environmental impacts.
- The proposed framework for V-PI, freight damage and logistics provides a context of explicit tasks, aims, and responsibilities for the scale of various spheres of influence. Use of such a framework might help to foster better coordination and a more coherent, multi-specialty approach to improving freight movement and, ultimately, improve cost-efficiencies, reduce costs to customers and enhance economic competitiveness.
- The significance of this study should be expanded through additional studies incorporating a larger range of vehicle types, road conditions and economic areas.

ACKNOWLEDGEMENT

Financial sponsorship by Caltrans in the development of the concepts discussed in this paper is acknowledged.

REFERENCES


Air void characterisation of HMA gyratory laboratory-moulded samples and field cores using X-ray computed tomography (X-ray CT)

L F Walubita, B Jamison, A E Alvarez, X Hu, C Mushota

The research work presented in this paper deals with the characterisation of the internal structure of hot-mix asphalt (HMA), incorporating both gyratory compacted samples produced in the laboratory and field cores. The primary objective was to determine the optimum trim depth on either end of laboratory-moulded HMA cylindrical samples that would optimise the air void (AV) uniformity in the test specimens. The analysis was based on the X-ray Computed Tomography (X-ray CT) scanning tests and subsequent image analyses. Two Texas HMA mixes, namely a coarse-graded (Type B) and a fine-graded (Type D) mix, with gyratory samples compacted in the laboratory to two different heights (110 and 164 mm) were evaluated for their internal structure in terms of the distribution of both the AV content and AV size. Analysis of the results indicated that the coarse-graded HMA mix (Type B) and the taller (164 mm in height) gyratory-moulded samples would be more likely associated with a more heterogeneous distribution of the AV content and AV size, respectively. Supplemented with field cores, the X-ray CT results indicated significantly poor AV content distribution (i.e. higher AV content and weakest area) at the ends, particularly in the top and bottom 20 mm zone of the samples. Thus, for 150 mm diameter samples of height equal to or greater than 110 mm, trimming a minimum of 20 mm on either side of the gyratory compacted samples should be given due consideration without compromising the specimen aspect ratio and NMAS coverage requirements (NMAS – nominal maximum aggregate size). In general, test specimens should always be cut from the middle zone of the SGC moulded samples where the AV is less heterogeneously distributed.

Keywords: hot-mix asphalt (HMA), superpave gyratory compactor (SGC), air voids (AV), mix internal structure, X-ray computed tomography (X-ray CT)
INTRODUCTION
The internal structure of a hot-mix asphalt (HMA) can be analysed in terms of the air voids (AV) characteristics (e.g. distribution, size, and connectivity), and the aggregate orientation, contact, and distribution (Alvarez et al 2010a). This HMA characteristic (internal structure) is one of the key factors that greatly influences the performance of HMA, including variability in the test results of not only the gyratory laboratory-moulded samples, but also of field cores. This is particularly critical for HMA tensile loading tests such as the direct-tension (DT) test. Thus, it is important to characterise the mix internal structure, for example via the distribution of the total AV content (or AV content) as a function of HMA sample height (or depth). The primary objective was to determine the optimum trim depth on either end of laboratory-moulded HMA cylindrical samples that would optimise the AV uniformity in the test specimens. X-ray CT scanning tests were conducted on cylindrical HMA samples (150 mm diameter) that were gyratory laboratory-moulded to two different heights of 110 and 164 mm. Two Texas HMA mixes, namely a coarse-graded (Type B) and a fine-graded (Type D) mix, were used for the study. To relate to in situ field conditions, the results of these two laboratory mixes were supplemented with X-ray CT results of field cores from in-service perpetual pavement (PP) sections.

In the paper, the X-ray CT scanner and the concepts of image analysis are described first, followed by the experimental design plan and the laboratory test results including the effects of the AV distribution on the DT test failure mode. A summary of key findings and recommendations is then presented to wrap up the paper.

X-RAY COMPUTED TOMOGRAPHY (X-RAY CT)
The X-ray CT is a non-destructive test used to capture the internal structure of materials. Various applications of this method have been discussed by Masad (2004) and others (Braz et al 1999; Shashidhar 1999). In HMA mixes, X-ray CT has been used successfully for characterising the influence of compaction on HMA internal structure (Masad et al 2009), assessing internal structure of open-graded mixes (Alvarez et al 2010a; Muraya 2007), analysing water transport characteristics (Kassem et al 2008; Masad et al 2007), and stone-on-stone contact assessment (Alvarez et al 2010b; Watson et al 2004). The X-ray CT setup at Texas A&M University, used to perform the HMA scanning for this study, is shown pictorially in Figure 1.

The setup shown in Figure 1 includes two separate systems placed in the same shielding cabinet. The mini-focus system has a 350 kV X-ray source and a linear detector, whereas the micro-focus system has a 225 kV X-ray source and an area detector. The mini-focus source can penetrate thicker and denser samples than the micro-focus source. The micro-focus system, however, provides better resolution than the mini-focus system. All the experimental measurements in this study were conducted using the mini-focus 350 kV X-ray source system. This system has the necessary power and resolution to penetrate the HMA samples and provide good quality grey scale images for subsequent analysis of AV characteristics (Kassem et al 2008).

In the X-ray CT, a test specimen is placed between an X-ray source (Figure 1) and a detector. The intensity of X-rays change from $I_0$ before entering the specimen to $I$ after penetrating the specimen due to the absorption and scattering of radiation. The relationship between $I_0$ and $I$ is related to the linear attenuation coefficients of the materials that constitute the specimen, which are related to the densities of these materials. As such, determining the attenuation coefficients allows calculating the density distribution within a specimen section. These different densities are represented in a grey scale image of the section that consists of 256 grey intensity levels with a lower density material represented by a darker colour (e.g. AV are shown as black pixels) – see Figure 2.

In this study, images were captured every 1 mm in the vertical direction of the specimen and with a horizontal resolution equal to approximately 0.17 mm/pixel. Images were

Figure 1 Pictorial setup of the X-ray CT scanner

Figure 2 Example of a grey scale X-ray CT image (Kassem et al 2008)
analysed using macros that were developed (Masad et al 2007) in Image-Pro® Plus software (Media Cybernetics 1999). Using a suitable grey intensity threshold value, AV can be separated from other HMA mix constituents (aggregate and asphalt-binder). The threshold level represents a boundary value below which pixels in the analysed image are considered as part of the AV, whereas pixels that have intensity values above the threshold value are considered to belong to the remaining phases. The analysis is capable of quantifying the vertical and horizontal distributions of AV, size distribution of AV, and connectivity of the AV. Additional details on the image analysis and computation of these parameters are documented elsewhere (Alvarez et al 2009). Readers are referred to Masad (2004) and Masad et al (2009) for more details on the different X-ray CT configurations, operational features and other capabilities.

However, it must be stated herein that the accuracy of the image analysis is also considered to be a function of the maximum scanning resolution of a specific X-ray CT device that is used. To attain more accurate results, some researchers proposed to apply a correction scheme such as a Weibull distribution (Luo & Lytton 2011). In this study, the mini-focus 350 kV X-ray source that was used was considered to have sufficient resolution and to be accurate enough not to warrant the need for data correction adjustments.

**EXPERIMENTAL DESIGN**

Two Texas HMA mixes were evaluated in the laboratory: a coarse-graded Type B and a fine-graded Type D mix. The HMA mix-design characteristics are summarised in Table 1 and the aggregate gradations are shown in Figure 3. As shown in Table 1, two moulded cylindrical sample heights of 110 and 164 mm respectively were investigated, all with a diameter of 150 mm. These two heights were evaluated, because they represent the mould height that is typically used for fabricating test specimens for various laboratory tests such as the Overlay Tester (OT), dynamic modulus, repeated load permanent deformation, and DT (Walubita et al 2010c). A minimum of two cylindrical replicate specimens were scanned with the X-ray CT per sample height per mix type.

<table>
<thead>
<tr>
<th>Table 1 HMA Mix-Design Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMA Mix</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Type B</td>
</tr>
<tr>
<td>Type D</td>
</tr>
</tbody>
</table>

*Legend: NMAS = nominal maximum aggregate size; PG = performance-graded asphalt-binder*

---

**Figure 3 Aggregate gradations of the Type B and D mixes**

**Figure 4 The Servopac SGC and some moulded samples**

<table>
<thead>
<tr>
<th>Table 2 Typical as-built <em>in situ</em> HMA mix-design characteristics for the PP field cores</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMA mix designation</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>PFC</td>
</tr>
<tr>
<td>SMA</td>
</tr>
<tr>
<td>%” SFHMA</td>
</tr>
<tr>
<td>1” SFHMA</td>
</tr>
<tr>
<td>RBL</td>
</tr>
</tbody>
</table>

*Legend: PFC = permeable friction course; SMA = stone mastic (matrix) asphalt; SFHMA = stone-filled hot mix asphalt; RBL = rich-bottom layer; PG = performance-graded asphalt-binder; S = styrene-butadiene-styrene polymer (added as 5% by weight of asphalt-binder)*
Sample moulding was accomplished using a Servopac Superpave gyratory compactor (SGC) to a final target total AV content of 7±1%. This is the AV content that is typically specified for laboratory performance testing in Texas (TxDOT 2004; TxDOT 2011; Walubita et al 2010c). The SGC and some examples of the moulded samples are shown in Figure 4. Standard SGC moulding parameters were used, namely 600 kPa vertical pressure and 1.25° compaction angle at a rate of 30 gyrations per minute (Walubita 2006). Moulding and compaction temperatures were consistent with the asphalt-binder PG grade (Table 1) (TxDOT 2004; TxDOT 2011). Note that the SGC mould shown in Figure 4 can handle sample heights of up to 172.5 mm.

As stated in the introductory section, field cores were used to supplement and relate the total AV distribution of the laboratory-moulded samples to in situ field conditions. To achieve this, cores from PP sections that represent thick multi-layered HMA pavement structures with numerous layer-lift thicknesses and constructed with both fine- and coarse-graded HMA mixes were used (Walubita et al 2010a). The typical as-built in situ HMA mix-design characteristics of these PP cores are summarised in Table 2.

Table 2 shows that the PP core consists of a wide spectrum of aggregate gradations ranging from open- (PFC) to coarse- through to fine-graded (RBL) HMA mixes. Therefore, these PP cores were deemed satisfactory to provide a field correlation to the total AV distribution structure of the laboratory-moulded samples for the mixes listed in Table 1. Accordingly, the laboratory X-ray CT results of these PP field cores are also presented and discussed in the subsequent sections of this paper. More details about PP structures are published elsewhere (Walubita et al 2010a).

**RESULTS, ANALYSES AND DISCUSSION**

This section presents the results of the AV characteristics analysis for both laboratory-moulded samples and field cores. The results are presented in terms of the following: distribution of AV content and AV size for laboratory-moulded samples, distribution of AV content for field cores, and effects of AV distribution on the tensile failure mode of laboratory-moulded samples.

### Distribution of air voids (AV) content: laboratory-moulded samples

The vertical distribution of the total AV content as a function of the SGC sample height (or depth) is shown in Figures 5 and 6 for the 164 and 110 mm high samples respectively. These vertical distributions of AV content are hardly uniform and exhibit considerably high values (i.e. low density) at the top and bottom ends of the SGC samples, particularly in the top and bottom 20 mm zone, where the range of the AV content is from 8% to as high as 25%. For a final target AV content of 7±1% for the laboratory samples, both Figures 5 and 6 suggest staying away from these end zones. Based on this data, trimming a minimum of 20 mm off the top and...
bottom ends of the SGC laboratory-moulded samples may thus be warranted.

Additionally, the taller 164 mm SGC sample (Figure 5) also exhibited much higher AV content values in the middle zone, and their AV content distribution is generally more heterogeneous than that of the shorter 110 mm high samples. Although lower than the 7±1% target in the central portion of the specimen, the AV content distribution for the 110 mm high samples (Figure 6) appears to be less heterogeneous, particularly in the middle zone and does not exhibit any significant difference between the coarse-graded Type B and the fine-graded Type D mix (Walubita et al. 2010a).

Overall, these results (particularly Figure 6) indicate that test specimens should generally be cut from the middle zone of the SGC-moulded samples where the AV content is less heterogeneously distributed, i.e. avoid the edges. For example, a single test specimen of 38 mm thickness for the overlay tester would easily be cut from the middle zone of the SGC-moulded sample shown in Figure 6 with reasonable uniformity in terms of AV content distribution.

As indicated in Figure 7, taller moulding heights would be more prone to heterogeneous AV content distribution and variability than shorter moulding heights (having better AV content uniformity in the middle zone of the SGC samples). Thus, wherever possible and under the current compaction SGC parameters, shorter compaction moulding heights should always be given preference to promote the AV content uniformity in the test specimens.

In general, higher sample-mould heights are also comparatively more susceptible to aggregate segregation, particularly for the coarse-graded HMA mixes. In fact, it is normal practice for the HMA mix to be scooped and poured into the SGC mould in more than one lift when the moulding height is greater than 110 mm just to minimise vertical segregation. Figure 8 illustrates an example of a possible vertical segregation of the aggregates within the SGC sample of a coarse-graded Type B mix (22 mm NMAS).

Clearly, the AV content between the sample heights of 20 mm and 70 mm is relatively lower in magnitude than that below 70 mm. One possible explanation for the higher AV content above 80 mm height is that the larger aggregate particles may have vertically segregated to the sample bottom, thereby creating the higher AV content seen in Figure 8. By contrast, this AV content distribution pattern (Figure 8) was not observed with the fine-graded Type D mix or with the shorter mould height of 110 mm (Walubita et al. 2010a). In summary, taller mould heights can have a higher tendency to promote vertical aggregate segregation, particularly in coarse-graded mixes. Additional research should be conducted to determine if this phenomenon exists in other HMA mix types.

In addition, the results shown in Figures 7 and 8 are consistent with field experience, where compaction and density (i.e. AV content uniformity) problems have also been reported with thick layer lifts – synonymous to higher laboratory sample mould heights. In their construction work with PPs, Walubita et al. (2010a) concluded that thin layer lifts of less than 100 mm yielded better compaction and in situ density results than thick layer lifts of 125 mm and higher for coarse-graded SFHMA mixes. The general observation from this study (Walubita et al.
was that compacting at the lower lift thickness such as 75 mm yielded a more constructable mix than using thicker lifts greater than 100 mm, as gauged by attaining the target in-place density and layer interface bonding. Furthermore, compacting at a greater lift thickness tended to cause the mixes to segregate vertically, creating highly voided areas capable of detrimentally trapping moisture. Figure 9 shows a comparative illustration of the compacted lift thickness for a 25 mm NMAS coarse-graded SFHMA layer whose in situ target density was 96% (Walubita et al 2010a).

Figure 9 shows better construction quality for the 75 mm and 100 mm layer lift-thicknesses (cores #1, #8 and #3) with no visual evidence of vertical segregation or debonding. Also, the measured average AV content using the traditional water-displacement method at 7.3% was fairly reasonable as opposed to about 12.6% for the 125 mm layer lift-thickness (cores #7). Synonymous to the conclusions made for the results in Figures 7 and 8, Walubita et al (2010a) also recommended that a thinner layer-lift thickness (i.e. ≤ 100 mm) was appropriate for the coarse-graded SFHMA mixes in terms of yielding acceptable field compaction and in situ density results during construction (Walubita et al 2010a). In either case, however, the aspect ratio and NMAS coverage requirements must still be met as specified below for laboratory-moulded samples and field compaction layer-lift thickness respectively (Bonaquist et al 2003; Cooley (Jr) & Brown 2003):

- Aspect ratio (ar) (longest side divided by the shortest side): 1.5 ≤ ar ≤ 2.0
- NMAS coverage (NMAS_C):
  1.5 × NMAS ≤ NMAS_C ≤ 3.0 × NMAS

**Distribution of air voids (AV) size: laboratory-moulded samples**

Interestingly for SGC samples of similar height, there was no significant difference in the trends of the vertical AV content distribution (Figures 5 through 7) or total AV content values (at an average of 8.3%) between the Type B and D mixes. By contrast, and as shown in Figures 10 through 13, there was a considerable difference in the AV sizes. The more coarse-graded Type B mix exhibited relatively larger AV sizes in magnitude compared to the more fine-graded Type D mix, as theoretically expected.

Based on the vertical distribution of the AV size shown in Figures 10 and 11, the Type B mix exhibited larger AV size values across the entire SGC sample height than that of the Type D mix. In fact, the average AV size was 0.95 and 0.82 mm for the Type B and Type D samples respectively. Also the variability of the AV size, measured in terms of the coefficient of variation (COV), for the Type B mix (i.e. COV = 28%) was higher than that for the Type D mix (COV = 18%). Based on these COV values, it would be expected that the Type B mix would be associated with more AV size variability during SGC sample fabrication as when compared to the Type D mix.

![Figure 11 AV size comparisons – 164 mm high samples](image)

![Figure 12 Vertical distribution of AV size for the Type D mix](image)

![Figure 13 Vertical distribution of AV size for the Type B mix](image)
Compared to the fine-graded mixes, coarse-graded mixes with larger AV sizes would thus be more susceptible to oxidative ageing and/or moisture damage due to the possibilities of easy air or water infiltration. However, future research based on image analysis should focus on the computation of AV connectivity (i.e. content and vertical distribution) to validate this hypothesis.

The connected AV serves as paths for water movement (defining the mix permeability) and air circulation through the mix, which undesirably creates favourable conditions for oxidative ageing and/or moisture damage. In addition, it can intuitively be stated that coarse-graded mixes require more meticulous work and caution during laboratory sample fabrication. As subsequently discussed, similar results were also observed with field cores.

As discussed for the vertical distribution of the AV content, Figures 12 and 13 suggest that taller moulding heights would be more prone to heterogeneous distributions of the AV size and variability than shorter moulding heights. These findings further support the convenience of preferring (based on the current compaction SGC parameters) shorter compaction mould heights to promote uniformity in the internal structure of the SGC samples. One of the possible causes of having higher AV content and AV size at the ends, in addition to aggregate segregation, is the restriction of the top and bottom metallic surfaces of the SGC device that limits the movement of the aggregate particles and decreases the compaction efficiency at these zones of the sample (Thyagarajan et al. 2010).

Distribution of air voids (AV) content: field cores

Figure 14 shows the AV content distribution of a core extracted from a newly constructed PP section prior to opening to conventional traffic (Walubita et al. 2010c).

The SFHMA layers represent coarse-graded mixes of 25 mm NMAS, whose target in situ density was 96%, and was compacted in 75 mm layer lift-thickness. The RBL layer represents a dense- to fine-graded mix of 12.5 mm NMAS, whose target in situ density was 97% at a layer lift-thickness of 50 mm (Walubita et al. 2010a). The aggregate gradations for these HMA mixes are shown in Figure 15.

As echoed by the laboratory AV content determined using traditional methods, Figure 14 shows that the AV content for the RBL (fine-graded) is lower than that for the SFHMA layers in magnitude. This is consistent with theoretical expectations and supports preceding results. Similarly, the AV content seems to be relatively higher in magnitude at the ends – in this case at the bottom zones of the layer and lift interfaces, among other reasons possibly due to vertical segregation of the aggregates, which is illustrated by the example shown subsequently in Figure 16 (Walubita et al. 2010c).

The AV content distribution pattern depicted in Figure 8 for the SGC sample of the coarse-graded Type B mix (22 mm NMAS) suggested the possible occurrence of vertical segregation in the aggregates, i.e. larger aggregate particles gravitationally moving to the bottom. Figure 16 provides this visual evidence with a core from a PP section. This core consists of a coarse-graded SFHMA mix that was compacted in 125 mm layer-lift thicknesses, with a target in situ density of 96%. The aggregate gradation was shown in Figure 15. Clearly, this figure highlights the potential construction problems of thicker layer-lift thicknesses, with a target in situ density of 96%. The aggregate gradation was shown in Figure 15. Clearly, this figure highlights the potential construction problems of thicker layer-lift thicknesses, particularly for coarse-graded mixes such as the SFHMA mixes. If water infiltrates through and gets trapped within this high AV content mix, it could undesirably damage the HMA pavement (Walubita et al. 2010a).
Figure 17 shows another comparison of the vertical distribution of AV content for different layers of an approximately 525 mm thick multi-layered PP core, with the fine-graded RBL exhibiting the most AV uniformity (i.e. the lowest AV content distribution variability).

Because of its functionalities within the PP structure, such as preventing bottom-up fatigue cracking and upward water intrusion from the base/subgrade, the RBL is typically compacted to a relatively higher in situ target density than other layers (i.e. 97% versus 96% specified for SMA and SFHMA, or 80% specified for PFC mixes). Combined with its fine aggregate gradation and rich asphalt-binder content (average 5.4%), this partly accounts for its superior AV content uniformity compared to the other layers (Walubita et al. 2010a). The comparative aggregate gradations for these PP cores are illustrated in Figure 18, with the RBL having the finest aggregate gradation.

Just like the RBL in Figure 18, Table 1 indicated a finer aggregation gradation (Figure 3) and more asphalt-binder content for the Type D mix compared to the Type B mix. Conversely, and just like in Figure 17, the fine-graded and rich Type D mix exhibited better AV uniformity than the coarse-graded Type B mix (refer to the preceding results in Figures 5 through 13).

The PFC, as theoretically expected, shows the highest AV content values, ranging from about 10% to 25%. Again, the AV content peak at the layer- and lift-interfaces is clearly visible, particularly for the coarse-graded SFHMA mixes. Actually, if the SFHMA layers are analysed separately, their AV distribution patterns do not differ significantly from Figures 5 through 8 for the laboratory-moulded samples. These high AV content peaks suggest potential for debonding problems, which is undesirable. In fact, debonding problems were experienced with some SFHMA cores – see Figure 9 (Walubita et al. 2010c).

In addition to possible vertical aggregate segregation, the heterogeneous AV content distribution exhibited in Figures 14 through 17 may have been contributed by many other factors, including but not limited to the following:

- Temperature segregation within the HMA mix during construction. At the bottom of the HMA mix, the existing layer (which is at a relatively lower temperature) may be promoting rapid cooling and therefore low density at the interface. At the surface, the compactor wheels, as well as the effects of the environment, may have a rapid cooling effect on the top portion of the HMA mix.
In the case of the coarse-graded SFHMA mixes, conjecture is that the large angular aggregate particles do not receive sufficient compactive energy at the bottom of thicker lifts to promote adequate reorientation, and thus contribute to the poor AV distribution structure. This problem is exacerbated by the low asphalt-binder content (Table 2) that limits the effective lubrication to allow reorientation of the aggregate particles (Walubita et al. 2010a).

### Effects of air voids (AV) distribution on the tensile failure mode

Figure 19 shows a typical vertical distribution of AV content in a 164 mm high SGC compacted cylindrical sample (Type D mix) and a side by side comparison including the tensile failure zones for a specimen tested under DT loading.

In Figure 19, the red horizontal boundaries represent the AV content distribution for cutting the sample to 150 mm high, while the green boundaries represent the AV content distribution for cutting the sample to 100 mm high HMA test specimens. With respect to the target AV content tolerance, the dashed blue and red lines represent the lower and upper allowable limits respectively; i.e. 7±0.5% for this particular case for DT testing (Walubita et al. 2010a).

Clearly, Figure 19 shows that the AV content distribution is not homogeneous, leading to weaker zones at the ends of the sample. For DT testing, the tensile failure zone should be in the middle, as exhibited by the 100 mm high test specimen in Figure 19 (Walubita et al. 2010b). End-failures, such as the one exhibited by the 150 mm high test specimen (on the right side of Figure 19), are undesirable. For the 150 mm high specimen, Figure 19 in fact shows a maximum trim depth of about 10 mm, which yielded undesired results with the DT tensile testing. A minimum trim depth of 20 mm would likely yield the desired failure mode as evidenced by the 100 mm high test specimen. Similarly, the irregular distribution of the AV content in the HMA mix specimens can lead to problems for testing and evaluation of other HMA characteristic properties, such as permanent deformation resistance, stiffness, and/or moisture damage susceptibility.

### CONCLUSIONS AND RECOMMENDATIONS

Based on the research work that was conducted and the data presented in this paper, the following conclusions and recommendations were drawn:

- In general, there is potential for aggregate vertical segregation and AV content variability in SGC samples moulded to taller heights in the laboratory. Where possible, shorter mould heights that would allow sufficient trim depth should be used so as to optimise the AV content uniformity of HMA test specimens. For the SGC compactor and sample types considered in this study, a sample height of less than 164 mm – that does not compromise both the aspect ratio and NMAS coverage requirements – would be preferred.

- For the type of SGC samples discussed in this paper (i.e. 110 mm and 164 mm high with 150 mm diameter), trimming a minimum of 20 mm on either side is recommended. This is envisaged to optimise the total AV content uniformity. In general, test specimens should always be cut from the middle zone of the SG-moulded samples where the total AV content is less heterogeneously distributed.

With shorter mould heights, however, caution should be exercised to satisfy the test specimen aspect ratio and NMAS coverage requirements.

- As evidenced by the field core data presented in this paper, there is also a high potential for compaction problems, including aggregate vertical segregation and in situ density variations, when opting for thicker layer-lift thicknesses in the field during construction. Where possible, thin layer-lift thicknesses should be used that will optimise the construction quality without compromising the NMAS coverage requirements.

- For the materials and HMA mixes evaluated in this study, the coarse-graded mixes exhibited more susceptibility to aggregate vertical segregation and AV content variability with larger AV sizes than fine-graded mixes. Thus, care should be exercised both in the laboratory and field to optimise density/AV content uniformity. Fine-graded mixes on the other hand are...
more compactable with better density/AV content uniformity characteristics. Overall, this study has demonstrated that the X-ray CT scanner, coupled with image analysis, can be satisfactorily used to characterise the AV content distribution of both laboratory-moulded samples and field cores. This non-destructive system has the potential and capability to measure both the AV content and AV size as a function of depth or sample height. However, the accuracy of the image analysis is also considered to be a function of the maximum scanning resolution of a specific X-ray CT device that is used, and where applicable, some data correction adjustments may be warranted to optimise the accuracy of the test results.

**ACKNOWLEDGEMENTS**

The authors are thankful to all those who provided support (financial [TxDOT] or technical) in the course of this research work. In particular, special thanks are due to Gautam Das, Tony Barbosa, Lee Gustavus, Rick Canettlla and Jeff Perry for their assistance with the laboratory and field work. Special thanks also go to Geoffrey S Simate (University of the Witwatersrand, South Africa) for his technical insights to the paper.

**DISCLAIMER**

The contents of this paper reflect the views of the authors who are solely responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of any agency or institute. This paper does not constitute a standard or specification, nor is it intended for design, construction, bidding, contracting or permit purposes. Trade names were used solely for information purposes and not for product endorsement.

**REFERENCES**


Masad, E, Kassem, E & Chowdhury, A 2009. Application of imaging technology to improve the laboratory and field compaction of HMA. Report No FHWA/TX-09/0-5261-1, College Station, TX, US: Texas Transportation Institute, Texas A&M University.


TxDOT (Texas Department of Transportation) 2004.

Standard specifications for construction and maintenance of highways, streets, and bridges. Austin, TX, US.

TxDOT (Texas Department of Transportation) 2011.

TxDOT – online manuals. Available at: http://www.onlinemanuals.txdot.gov/manuals/.

Walubita, L F 2006. Comparison of fatigue analysis approaches for predicting fatigue lives of hot-mix asphalt concrete (HMAC) mixtures. PhD thesis, Texas A&M University, College Station, TX, US.

Walubita, L F, Liu, W & Scullion, T 2010a. Texas perpetual pavements – Experience overview and the way forward. Technical Report No 0-4822-3, Texas Transportation Institute, Texas A&M University, College Station, TX, US.

Walubita, L F, Simate, G S & Oh, J 2010b.


New generation mix designs: Laboratory testing and construction of the APT test sections. Report No 0-6132-1, Texas Transportation Institute, Texas A&M University, College Station, TX, US.

Weak layers, interlayers, laminations and/or interfaces in the upper structural layers of road pavements are specifically prohibited in most road-building specifications. However, such layers are extremely common and often lead to premature pavement distress. In Part 1 of this two-part set of papers, it is shown that from experience with heavy vehicle simulator (HVS) and dynamic cone penetrometer (DCP) testing, the presence of such layers and/or conditions at any depth in the structural layers of a flexible or semi-flexible pavement is far more deleterious than is commonly appreciated. In Part 2 the effects of these weak layers are further modelled and discussed using various examples based on an HVS testing and mechanistic pavement analyses. In particular, a weak upper base course of a cemented pavement under a thin bituminous surfacing may lead to severe surfacing (and upper base) failure within a matter of weeks after opening to traffic, not excluding failure even during construction. In this paper (Part 1), the causes of weak layers, interlayers, laminations and/or interfaces, together with simple methods for their detection during construction and analyses of their effects on the structural capacity of flexible and semi-rigid (cemented) road pavements, are briefly discussed.

**INTRODUCTION**

Premature distress in the form of rippling, arcuate (curved) slippage cracking or shoving of the surfacing and shallow base failures of pavements with bituminous surfacings is not rare in southern Africa. The authors know of over 100 such cases that have occurred over the last 50 years and have investigated a number of them. Such distress is usually due to the presence of a weak interlayer between the bituminous surfacing and the base course. Weak interlayers are in fact quite common in spite of current precautions specified to prevent them. Pumping of fines through cracks from a weak interlayer (or laminated interface) between a concrete, asphalt or cemented base and the subbase (De Beer 1985) is another well-known form of distress. Photo 1 shows pumping from a cracked cemented base layer, and Photo 2 shows some fatigue cracking and pumping from a recently constructed road. Examples of delamination (possibly due to construction) are shown in Photos 3, 4, 5 and 6.

The aim of this paper (Part 1) is to show by means of case histories, HVS and DCP testing that the presence of weak layers, interlayers, laminations and/or interfaces at any depth in the structural layers (but especially the upper base) of a flexible or semi-flexible pavement is far more deleterious than is generally assumed or appreciated. In Part 2 of this two-part set of papers (De Beer et al 2012 – see page 43 of this edition) this effect is further discussed and demonstrated in more detail using the well-known mechanistic analysis applicable to the structural design of road pavements.

**DEFINITIONS**

For the purposes of these papers, a weak layer or interlayer is regarded as any layer that is weaker than was assumed in the design and that has a practical measurable thickness, \( t \), of \( > 1 \) mm. Such layers are often referred to as “laminations”, “biscuits” or, less often, “false layers” (Bergh 1979) or “pie crusts” (Gray 1979) (see Photos 5, 6, 7, 8, 9, 10 and 11).

In addition, a weak interface, on the other hand, is a condition of minimum friction...
Photo 2 Fatigue cracking and pumping on a recently constructed main road – possible weak interlayer directly under the surfacing seal (courtesy of Dr P Paige-Green).

Photo 3 Loss of most of prime under light traffic on same road as in Photo 4.

Photo 4 General view of weak, carbonated, primed 3% cement-stabilised weathered granite base disintegrating under light traffic and showing chain drag in foreground.

Photo 5 General view of primed, 3% cement-stabilised, weathered granite base after loose to weak upper 25 mm of carbonated material had been hand-broomed off.

Photo 6 Close-up of Photo 5 showing exposed harder base after loose to weak upper 25 mm of carbonated base had been hand-broomed off.

Photo 7 Roller-induced shear plane dipping to right with overturned shear “biscuits” (uncarbonated on lower surface) lying at left. Loose to very weak (could be broomed or easily scratched off) to 20 mm, and carbonated and weak to 40 mm below top of primed, 3% cement-stabilised, weathered granite base. Detected by chain drag.
(slip) between two layers, or parts of a layer, normally in the upper 75 mm or so in the cemented base, and has no virtual thickness, i.e. \( t \sim 0 \) mm. Such weak layers, interlayers, laminations and/or interfaces are illustrated in the afore-mentioned photos.

In terms of consistency, such interlayers may be loose, i.e. with no unconfined compressive strength (UCS), and at a relatively lower density, be relatively weak, i.e. with a reduced UCS, or relatively soft (moist and/or plastic) (see Photos 10, 11, 12, 13 and 15).

In practical terms of soil structure, they may be intact (structureless), sheared or laminated. The term “laminations” or, colloquially, “biscuits”, is often applied to all of these forms, in spite of the former meaning thinly layered (i.e. geologically up to 10 mm) and the latter implying a relatively hard but brittle material. However, a smooth or striated flat surface (i.e. interface as defined above with thickness \( t \leq 1 \) mm) is usually found immediately below most forms of weak interlayers. Therefore, smooth, flat, inclined or curved surfaces representing interfaces between compacted layers or shear planes within layers are also regarded as forms of undesirable weak interlayers. The detection and laboratory/field testing of interlayers and interlayer shear transfer between layers (including asphalt layers) are, among others, discussed by Oba and Partl (2000), Diakhate et al (2006), and Canestrari et al (2005). The term “chemically stabilised” is used here to mean a material modified or cemented with cement or lime as envisaged in Technical Recommendations for Highways (TRH) 13 (NITRR 1986a).

Regarding carbonation in terms of road structure, a weak layer may or may not be directly related to the existence of a lamination or a weak interface, as it could develop during construction or because of physicochemical breakdown (Netterberg 1991, 1994), or it could be due to a traffic-associated crushing failure as discussed above. However, carbonation of road layers (or parts thereof) should be seen in the context of the layer’s use within a road pavement. A strongly cemented layer might
show signs of carbonation, but the strength of the carbonated material is still adequate for its use and purpose in the pavement. This could be described as “non-deleterious” carbonation (or simply carbonation), whereas when carbonation causes the properties of the material to deteriorate to the extent that the layer cannot fulfil its intended function, this is known as “deleterious” carbonation, in the context of this set of papers.

**CRITICAL IMPORTANCE OF A STRONG UPPER BASE**

Although it has generally been appreciated for many years (e.g. Bergh 1979) that the upper base must be relatively strong and intact, it is not generally appreciated just how critical this is, especially in the case of a pavement with a cemented base under a surface treatment or thin asphalt surfacing. Experience has shown that such pavements constructed in the last 20 to 30 years do not distress so much by cracking as by rippling and shoving of the thin surfacing associated with a weak interlayer, usually some 3 to 20 mm thick, between the base and the surfacing.

The following section demonstrates the importance of a strong upper base layer, and is based largely on the original work by De Beer (1989a, 1989b, 1990) and De Beer et al (1989).

---

**Heavy Vehicle Simulator (HVS)–DCP correlation**

Testing of such pavements under the HVS and correlation of the results with special DCP testing of the same pavements resulted in the prediction model given in Equation 1 (De Beer 1989b, 1990; De Beer et al 1989):

\[
R_L = \frac{DSN_{200}}{10} - \frac{3.83 - DN_{50}}{1.39}
\]

where:

- \( R_L \) = Linear rate of deformation (rutting) in mm/million equivalent standard axles (MESA) (See note on MESA after Equation 2)
- \( DN_{50} \) = Average penetration rate in upper 50 mm of pavement, including surfacing in mm/blow (penetration depth measured after every blow)
- \( DSN_{200} \) = Total number of blows in upper 200 mm of pavement, including the surfacing (DCP penetration depth measured after every blow)

Probability of regression = 50%, \( R^2 \) = 76%, \( n \) = 29.
As the rate of deformation on lightly cemented pavements appears to be linear (De Beer 1989a, 1990), Equation 1 can be rewritten as:

\[ CAP_{T20} = \frac{10^{\left(\frac{2.83-\text{DSN}_{50}}{1.39}\right)} \times (20 - R)}{\text{DSN}_{200}} \]  

Where:

- \( CAP_{T20} \) = Structural capacity in terms of MESA to an assumed additional rut depth of \( (20 - R) \) mm, measured from an existing rut depth \( R \) in mm.

Note that MESA = Million Equivalent repetitions of a Standard 80 kN (8 200 kg) Axle with four tyres at 520 kPa tyre pressure.

In this paper, the research based on the DCP and HVS testing was done in such a way that the relative damage coefficients for permanent deformation \( d \) could be determined. These values ranged between 1.2 and 1.8 for the pavements investigated in the relatively dry state, and were used to convert actual HVS tyre loading to equivalent loading repetitions on which Equations 1 and 2 are based. The usual value of \( d = 4 \) was therefore not used in this study. (For the original work see Chapter 4 in De Beer 1990.)

This model is the same as that used in the CSIR computer software (CSIR 2007) for the analysis and classification of DCP survey data on lightly cemented (C3 and C4 materials) pavements as defined in TRH 4 (COLTO 1996). As stated, the above model (Equation 2) was derived using actual HVS loading repetitions of a given wheel load, converted to MESA or Equivalent Standard Axle Load (ESAL) or E80 units. However, a convenient nomogram for an additional rut depth of \( (20 - R) \) mm (modified from the original Figure 13 in De Beer et al 1989) is provided in Figure 1. This shows the approximate structural capacity \( CAP_{T20} \) in MESA for an additional rut depth of 20 mm, hence the term \( (20 - R) = 20 \) mm, with \( R = 0 \) mm.

The empirical model given in Equation 2 applies strictly to pavements incorporating C3 or C4 cemented bases, i.e. UCS between 0.5 and 3 MPa (NITTR 1986a, 1986b), with the following DCP characteristics: \( \text{DSN}_{50} \) 200 to 750 blows, \( B = 0 \) and \( A \leq 3 \, 000 \) (see De Beer 1989b, 1990), a 20 mm maximum terminal rut depth, \( \text{DN}_{50} \) 0.5 to 4 mm/blow, and an equivalent structural capacity of up to 20 MESA. However, note that relatively higher values of \( \text{DN}_{50} \) (up to 9.5 mm/blow) are also shown in Figure 1, illustrating relatively low associated MESA’s. Figure 1 also supplies the \( \text{DSN}_{50} \) (total number of blows to penetrate upper 50 mm) associated with \( \text{DN}_{50} \) on the vertical axis. It is, however, clear that even under conditions of relatively high DCP penetration rates in the top 50 mm (if the above model is extrapolated beyond \( \text{DN}_{50} \) of 4.0 mm/blow), relatively low structural capacities (MESA) are indeed indicated, which carry some engineering value. For this reason Figure 1 is extrapolated beyond \( \text{DN}_{50} = 4.0 \) mm/blow.

Based on field observations by the authors and applying the methodology given above (HVS and DCP), the effect of a weak upper base layer was investigated. The results of this analysis are summarised in Table 1.

Three important observations are apparent from this analysis (Table 1):

Firstly, for the upper 200 mm of the pavement with an average strength of UCS = 1.5 MPa, a reduction in strength of only 300 kPa in the upper 50 mm to 1.2 MPa will roughly halve the structural capacity from 5.2 to 2.7 MESA. A reduction to 1.0 MPa will reduce the capacity to about 2 MESA, while a strength of 0.7 MPa (~ equivalent to a California Bearing Ratio – CBR – of about 80%) will reduce the capacity to about 0.3 MESA. Reductions of strength of this order are not readily apparent to inexperienced personnel and this pavement would probably be surfaced. Although Figure 1 should be regarded as only approximate, in situ DCP-derived UCS strengths as low as 0.3 to 0.5 MPa have been measured in the upper 50 to 75 mm in un-distressed areas of pavements which suffered premature distress within weeks to

---

**Figure 1** Nomogram for the estimation of the structural capacity of weakly cemented pavements (modified from De Beer 1989b, 1990) (Note that the UCS en CBR figures are the exact equivalents derived from the DN according to CSIR 2007 and would be rounded off in normal practice)
than increases the capacity. For example, of the cemented layer decreases rather than increases the capacity. For example, for a 1.2 MPa weak layer, increasing the average strength of the upper 200 mm from 1.5 to 2.0 MPa decreases the structural capacity from 2.7 to 1.9 MESA. This is most probably due to some “strength balancing” in the top 200 mm of the pavement.

Although any limitations of the model have been ignored in the simplistic analysis presented in Table 1, the results should be at least semi-quantitatively valid and serve to illustrate the critical importance of avoiding a weak interlayer between the base and surfacing layers.

Relatively weak layer on strong sublayer

The failure mode assumed in this analysis is that of weakening and/or crushing of the weak layer on a relatively stronger sublayer. The apparent anomaly in Figure 1, whereby increasing the strength of the under (sub) layer decreases the structural capacity, can be explained in terms of the structural pavement strength balance. In other words, it is easier to crack a nut by placing it on a harder surface.

CAUSES OF SURFACE WEAKENING

Surface weakening and even surface disintegration of compacted pavement layers – especially those chemically stabilised – during construction is not new and has been recorded from a number of countries in southern Africa and elsewhere (e.g. Netterberg et al 1987, 1989, Netterberg 1991, 1994). However, it is not widely known and is poorly understood. Known or suspected causes of, contributory factors towards or mechanisms causing a loose or weak layer at the top of a completed pavement layer include the following (mostly after Netterberg 1991, 1994):

- Very poor-quality gravel or soil: may shear during compaction
- Organic matter, including penetrating primer, weed killer, sugars: may retard set and/or prevent hardening
- Aggregate strength: may break down during compaction, especially near the top of a layer
- Aggregate durability: some aggregates disintegrate spontaneously
- Stabiliser type and quality (“fit for purpose”): suitability for the soil, soundness, age, rate of setting and hardening, slow slaking of quicklime, generation of gas by carbide lime (SANS 824-2006)
- Conversion of metastable calcium alumininate hydrates formed during chemical stabilisation of materials containing free or exchangeable aluminium (some laterites and tropical soils) (Sherwood 1993)
- Excessive or delayed mixing of chemically stabilised material
- Insufficient mixing (“brief-mix set”): a type of false set (Netterberg, in preparation)
- Aeration during mixing: another type of false set (Netterberg, in preparation)
- Poor vertical mixing through layer thickness

### Table 1 Calculated effect of a weak upper base on the structural capacity of a lightly cemented pavement in comparison with a pavement of uniform strength (also see Figure 1, with R = 0)

<table>
<thead>
<tr>
<th>DN&lt;sub&gt;20&lt;/sub&gt; (mm/blow)</th>
<th>Approximate equivalent strength [1]</th>
<th>Approximate structural capacity (CAP&lt;sub&gt;p/p&lt;/sub&gt;) for 20 mm additional rut (R = 0). See Figure 1 with decimal places as indicated (MESA)</th>
<th>Pavement (with and without 50 mm layer) [UCS in MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 to 50 mm</td>
<td>0 to 200 mm</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UCS (MPa)</td>
<td>CBR (%)</td>
<td>UCS (MPa)</td>
</tr>
<tr>
<td>1.8</td>
<td>1.5</td>
<td>180</td>
<td>1.5</td>
</tr>
<tr>
<td>2.2</td>
<td>1.2</td>
<td>150</td>
<td>1.5</td>
</tr>
<tr>
<td>2.6</td>
<td>1.0</td>
<td>120</td>
<td>2.0</td>
</tr>
<tr>
<td>3.6</td>
<td>0.7</td>
<td>80</td>
<td>1.5</td>
</tr>
<tr>
<td>5.0</td>
<td>0.5</td>
<td>50</td>
<td>0.7</td>
</tr>
<tr>
<td>8.0</td>
<td>0.3</td>
<td>30</td>
<td>2.0</td>
</tr>
</tbody>
</table>


### Notes

- See Figure 1 with decimal places as indicated (MESA).
Vibrating roller set at too high an amplitude

Reversing or stopping vibrating roller while it is vibrating

Delayed compaction: breaking down of upper part of layer due to compacting partially set, chemically stabilised or partially dried material

Inadequate compaction

Overcompaction: breaking up of already compacted, semi-hardened material, and production of new faces uncoated by cement or bitumen

Grading across to fill slacks instead of cutting to level (see Photos 5 and 12)

Incomplete removal of fines to expose good mosaic after slushing and before sealing

Spreading of fines on the base to improve the surface finish

Slushing or washing out of stabiliser: prohibited on a chemically stabilised layer, but might occur if emulsion-treated finishing methods (light slush, surface enrichment) or over-watering during finishing or curing are used

Primer type and quality: penetrating primers applied too soon, hindering cementation (NITRR 1986a, 1986c)

Frost damage (has occurred in the eastern Free State and in Lesotho)

Bacterial attack: deterioration of bitumen, generation of gas

Soluble salt crystallisation and/or hydration (previously common in southern Africa)

Acid attack on chemically stabilised material: acidic material, acid rain (personal communication with Johnson (1989) by main author)

Sulphate attack on chemically stabilised material

Detrimental carbonation of chemically stabilised material (common in southern Africa) from top and bottom of cemented layer (Sampson et al 1987; Netterberg 1991) (see e.g. Photos 7, 8, 10 and 15 with associated failures

Drying out of chemically stabilised material (incomplete hydration of cement): even one drying cycle can halve the UCS – probably a common cause (De Wet & Taupe 1985)

Wetting and drying cycles during curing of a chemically stabilised layer or due to rain: probably a common cause

Positive pore water or pore gas (usually water vapour) pressure developed in very hot weather

Positive air pressure developed due to a sudden drop in barometric pressure

Positive air pressure developed due to capillarity (air breakage, slaking)

Thermal buckling

Roots

Material variability (variable optimum moisture content (OMC), variable UCSs)

Emulsion quality and variability (“fit for purpose”): breaking too soon in plant or road, possibly even when within the specification

Variability in cement content (even within the specification)

Shearing by recycler and vibrator tyres, tamping roller and vibrator drum

Faster setting and/or hardening of the cement used than that calculated in the design

False setting of cement induced by high temperature

Reluctance of water to mix with the material: (hydrophobic material)

Difficulty in compacting 300 mm in a single lift while recycling (Netterberg, in preparation)

Use of a tamping or grid roller too high up in the base layer: disturbance and footprint compaction planes too deep to skim off – see Photo 14 (Netterberg, in preparation)

Very high ambient temperatures: water vapour or cutter blistering (entrapment of volatiles), accelerated carbonation, flash setting of cement (possibly aggravated by hot undiluted emulsion), accelerated setting and hardening causing compaction problems and poorer ultimate strength, lower pH of cement, accelerated breaking of emulsion, drying out during compaction, softening of emulsion “primer”, increased pore sizes due to expansion, conversion of reaction products (Netterberg, in preparation)

Almost sole reliance on a recycler for mixing of the materials, water, cement and emulsion in one pass only

Absorption of water and bitumen by the aggregate: premature breaking of an emulsion and incorrect OMCs

Smectite in the gravel: premature coagulation or retardation of an emulsion

Amorphous silica in the gravel: delayed expansion and weakening or decomposition of the cement (Botha et al 2005), later considered by Paige-Green (2009) to be extremely unlikely

Initial consumption of lime (ICL) or initial consumption of cement (ICC) of material not satisfied

Sodium carbonate in the gravel acting as an accelerator to a cement and/or as a retarder to an emulsion

Variation in proportions of old road, borrow material and sand during recycling

Problem of laboratory acceptance control by UCS testing of a material stabilised with a mixture of both cement and emulsion: uncertainty of best curing method

Use of laboratory mixes for UCS testing instead of road samples

Necessity to consider the optimum total fluid content (OTFC), not just the OMC, for compaction of bitumen-stabilised materials

Possible interaction between stabilisers: higher cement contents might accelerate the breaking of an emulsion, while higher emulsion contents might hinder hardening of the cement

Incompatibly charged dust particles along the base course–seal interface (McNally 1998) – hypothesis now apparently withdrawn (personal communication with McNally 1999 by main author)

Detrimental carbonation of chemically stabilised layer from below or from sides after construction

Punching (embedding) of chippings during surfacing or by traffic into a cemented base

Traffic (crushing of an already weak layer and/or sliding on smooth compaction planes): the actual cause of the failures, aggravated by overloading and especially high actual tyre–road contact stresses

Pick-ups of the seal by tyres

Discussion on possible causes of surface weakening

The causes, prevention of and remedial measures for all such weakening cannot be discussed in detail here – for some of them see Grant & Netterberg (1984), Kleyn & Buckle (1989) and Netterberg (1991, 1994). However, many are obvious and covered by the standard precautions and specifications (COLTO 1998 Section 8). Some have not been previously identified and will be discussed elsewhere (Netterberg, in preparation) and some are speculative and unproven. The most common causes are believed to be the filling of slacks during the cutting of final levels, overcompaction of a weak aggregate or partially hardened cemented material, salt crystallisation, surface carbonation of a chemically stabilised layer, surface drying out and/or cycles of wetting and drying of a chemically stabilised layer, water vapour pressures, the former use of a tar primer on an unhardened chemically stabilised material, and construction during very hot weather, especially with cementitious stabilisers. Some materials perform well in laboratory tests, but have a tendency to form a soft surface or soft base in the field (Bergh 1979).

The use of single-pass mixing is no longer recommended in Australia. At least two passes are recommended, especially for depths exceeding 250 mm (Aust Stab 2000). Weak
interlayers probably cause the most problems with chemically stabilised bases, although
unstabilised and even G1 crushed stone bases
(G1 material as defined in COLTO 1998) are
not immune (e.g. Netterberg et al 1989; Wyatt
& Thomson 1994). From published (e.g. Paige-
Green 1984, 1991, 2009; De Wet & Tautze
1985; Bagonza et al 1987) and unpublished
results, cement and lime-stabilised materials
lose, on average, about half their strength
either on carbonation – even after seven years
of ideal laboratory curing (Paige-Green 1991)
or even after one drying cycle (which may also
induce partial carbonation). Some materials
may even disintegrate. Weakening due to
detrimental carbonation, drying out and/or
wet-dry cycles is probably the most common
cause of surface weakening of chemically
stabilised layers.

DETECTION OF WEAK LAYERS
The current general use of nuclear methods
for compaction control and the practice of
sampling only before compaction mean that
holes are dug in most modern pavements
only when failures are investigated or when
the pavements are patched. These modern
methods have deprived the engineer of an
invaluable method of quality control for the
detection of a variety of defects, as well as for
the best measurement of layer thickness.
It is also clear that specifications and
engineering judgement have failed to prevent
the large number of cases of premature distress
known. It is therefore considered that addition-
al means are required. These could include:

- Visual inspection of trial holes for lamin-
ations, compaction planes and, after
spraying with phenolphthalein solution
(Netterberg 1984), for adequacy of mix-
ing and depth of stabilisation, including
filling of slacks with untreated material or
deteriorated window material.
- Testing for carbonation by spraying of
phenolphthalein solution and diluted
hydrochloric acid (Netterberg 1984; Paig-
ne-Green et al 1990) on the surface of
completed chemically stabilised layers.
(Note that in chemical soil stabilisation,
carbonation almost invariably weakens the
stabilised material.)
- Visual inspection of the surface of
completed layers aided by simple test
methods, such as:
  - Tapping or dragging a hammer or pick
handle across the layer and listening
to the tone of the sound produced.
  - Dragging a simple 2 m length of chain
across the layer (see Photo 4).
  - An electromechanical sounding device
or, even better, a simple chain drag as
specified in ASTM D 4580-03 (2009b).
  - Possibly a rapid vehicle-mounted
infrared method as used on bridge
decks (ASTM D 4788-03 2009a) for
longer lengths of road.
  - More frequent measurement of the
density at various depths, such as
from 0 to 50 mm, 50 to 100 mm and
100 to 150 mm, instead of just the
usual 150 mm. In this connection it is
important to note that sand has to be
calibrated for a particular depth of the
hole (NITRR 1986b). A nuclear gauge
does not determine moisture content to
the depth of the density probe when the
direct transmission method is used, but
only to a single depth which may be as
much as 200 mm (or more) by the flush
backscatter method (NITRR 1986a).
Gravimetric moisture content measure-
ments may therefore be necessary for
each depth. The calculation method for
measuring the density of different layers
by taking depth probe readings at dif-
terent depths is only approximate.
- If a chemically stabilised layer has
been badly cycled – even allowed to
dry partially only once – the upper
layer has probably been weakened.

Discussion on detection methods
Because the standard tests for grading,
plasticity and strength are all carried out
in the laboratory on a sample from the full
thickness of the layer, they are insensitive
to the presence of a thin layer of deleterious
material even when taken after compaction.
Even gradings carried out on samples taken
from the full thickness of distressed areas
do not necessarily indicate the presence of
an interlayer of fines between the base and
surfacing (Sampson et al 1985).

Standard deflection surveys are not
appropriate for the detection of a weak base
(Grant & Netterberg 1984) or a weak upper
layer of a cemented base, as a low (maximum
surface) deflection may be maintained by
the rest of the layer (Kennedy & Lister 1978).
However, the radius of curvature, which is
largely affected only by the upper 0,4 m of
the pavement (Grant & Walker 1972), will
generally be lower (Grant & Netterberg 1984)
and has been used by the main author with
more success. In the case of falling weight
deflectometer (FWD) surveys, the use of the
base layer index (BLI) (Horkar 2008) might
also prove successful.

Such techniques have, of course, been
used for many years, but should probably be
a requirement, not an option. Once a weak
layer at any depth has been detected, its
strength can be measured by means of the
DCP if it is not too thin (say t > 10 mm) and
its effect on structural capacity then assessed
mechanistically or, if in the upper cemented
base, by means of the CSIR WinDCP com-
puter program (CSIR 2007), or manually by
means of Figure 1 or Equation 1. When deal-
ing with thin layers it is advisable to take DCP
measurements after every blow close to such
layers rather than after the usual five blows.
How well the DCP evaluates hard biscuits
separated by smooth planes is not known.

Other possible methods for the determi-
nation of the strength of thin layers include the
rapid compaction control device (RCCD)
(De Beer et al 1993), a Clegg hammer (Clegg
1983; ASTM D 5874-02 2008) and other types
of portable falling weight testers (Hildebrand
2003). In the case of the Clegg hammer, it may
be necessary to consider also readings after one
or two blows in preference to the usual stan-
ard four blows. The bond strength between
cemented layers can probably be measured on
cores by using ASTM C 1245-06 (2007), or as
was reported by Canestrari et al (2005).

Importance of correct zero point
measuring with the DCP
The importance of taking the correct zero
point (the top of the vertical shoulder of the
DCP level with the top of the surfacing,
which is not removed, as in TMH 6 (NITRR
1984)) and also the ”softness” of the surfac-
ing into account can be illustrated by means
of a sensitivity analysis. An investigation of
a new road with a rich seal on a C3 base,
which had suffered premature distress due to
a weak upper base, is presented in Table 2.

Table 2 Calculated effect of removing soft
surfacing on the structural capacity
estimated by the DCP method (extract
from Netterberg, in preparation)

<table>
<thead>
<tr>
<th>Site No.</th>
<th>Depth of upper pavement used in analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0–50 mm [1]</td>
</tr>
<tr>
<td></td>
<td>Redefined [2]</td>
</tr>
<tr>
<td></td>
<td>Capacity (MESA)</td>
</tr>
<tr>
<td>1</td>
<td>0,06</td>
</tr>
<tr>
<td>3</td>
<td>0,00,01</td>
</tr>
<tr>
<td>5</td>
<td>0,05</td>
</tr>
<tr>
<td>6</td>
<td>1,3</td>
</tr>
<tr>
<td>7</td>
<td>0,07</td>
</tr>
<tr>
<td>8</td>
<td>0,002</td>
</tr>
<tr>
<td>9</td>
<td>0,3</td>
</tr>
<tr>
<td>10</td>
<td>0,8</td>
</tr>
<tr>
<td>11</td>
<td>0,4</td>
</tr>
<tr>
<td>12</td>
<td>0,2</td>
</tr>
</tbody>
</table>

[1] Including soft surfacing
[2] Excluding soft surfacing (i.e. directly on base layer)
This table shows that, in the redefined case, taking the zero at the point of inflection, i.e. effectively removing the generally new, hot, soft surfacing from the DCP analysis, resulted in an overall increase in the apparent structural capacity, in many cases a substantial increase. This was originally shown by Netterberg (in preparation).

In Part 2 of this set of papers (De Beer et al 2012 – see page 43 of this edition), the so-called traffic-associated "crushing failure" of lightly cemented layers is also discussed, which demonstrates the effect of tyre contact stresses on these pavement layers.

PREVENTION OF WEAK INTERLAYERS

Most weak layers, interlayers, laminations and/or interfaces can be prevented by good construction practices. For example, the South African standard specifications for road and bridge works (COLTO 1998) include a number of specific precautions clearly designed to prevent such interlayers:

- The minimum thickness of compacted pavement layers shall be 100 mm (para 3207 (b) (iii)) except in restricted areas where it may be as thin as 75 mm (para 32707 (b)(iii)).
- The underlying layer shall be free of loose material before a layer of plant-mixed, paver-laid material is placed upon it (para 3703).
- Mixing in of a chemical (para 3503 (d)) and, presumably also intended, a bituminous (para 3505 (b)) stabiliser shall be carried out over the full depth to be stabilised.
- Compaction of pavement layers shall be carried out so as to ensure that specified densities are obtained without damaging the lower layers (para 3207 (b)(i)).
- During compaction, the crust of a chemically stabilised layer shall be lightly harrowed or scarified before final rolling, if required by the engineer, in order to prevent the formation of laminations near the surface of the layer (para 3503(f)).
- The specified density of a crushed stone base shall be obtained throughout the entire layer, except when only water-rolled and, when so required by the engineer, shall be tested at various prescribed depths (para 3604(b)).
- In restricted areas the required densities of pavement layers shall be obtained throughout the thickness of the layer (para 3208 (b)(ii)).
- A completed crushed stone base layer shall be free from surface laminations after compaction (para 3604 (b)) and after slushing (para 3604(c)(i)).
- All excess fines shall be removed from the surface of a gravel (para 3403 (b)(i)) or crushed stone (para 3604 (c)(i)(ii)) base.
- Curing of a chemically stabilised layer for at least seven days is carefully specified and it is stated that drying out or wet-dry cycles may be cause for rejection if the layer is damaged thereby (para 3503(h)).
- No priming shall be carried out on a base which is visibly wet or which is at a moisture content in excess of 50% of the OMC (para 4104).
- Before priming, the base shall be broomed and cleaned of all loose material (para 4105).
- Asphalt shall not be placed if free water is present on the working surface or if the moisture content of the underlying layer, in the opinion of the engineer, is too high, or if the moisture content of the upper 50 mm of the base exceeds 50% of the OMC (para 4205(b)).
- Before applying a tack coat or asphalt, the surface shall be broomed and cleaned of all loose or deleterious material (para 4205(c)(i)).
- Before applying a seal, the moisture content of the upper 50 mm of base shall be less than 50% of the OMC (para 4304(b)) and shall be cleaned of all dust, dirt, dung, oil, etc (para 4304 (d)(i)).
- No asphalt overlay (para 4205 (b)) or reseal (para 4304(b)) shall be placed immediately after a rainy spell on an existing cracked or highly permeable surfacing resulting in the trapping of moisture.
- In addition to the specific precautions listed above, the aggregate strength and durability requirements for G1 – G3 crushed stone (para 3602a(a)) and G4 – G6 gravel materials (para 3402a(a)), and also the soluble salt requirement for both crushed stone (para 3601(b)) and gravel (para 3402(c)) are also intended in part to prevent the formation of weak interlayers due to aggregate degrada-
tion and salt damage respectively.
- However, in spite of all these precautions and earlier versions of many of them dating back many years, cases of premature distress due to weak interlayers continue to occur. This may be due to insufficient appreciation of just how deleterious such interlayers really are, difficulty in complying with some of the precautions and requirements, and/or inadequate supervision. In addition, so many cases are known where the interlayers must have been built-in as such that it must be concluded that they were simply not noticed. Currently, reliance is placed solely on the supervisory staff to detect such layers by whatever means are at their disposal, and no specific tests are specified or in general use. It is clear that reliance solely on visual inspection has proved inadequate in many cases.
- Suggested improvements to the above specifications include the following:
- Underlying layers should always be free of loose material.
- Holes should be dug in compacted layers to check for laminations, and/or weak interlayers, adequacy of mixing and, by spraying with phenolphthalein solution (Netterberg 1984), for an even distribution of stabiliser throughout the layer, from top to bottom (i.e. – pH > 10).
- The engineer should have the power to test the density, grading and plasticity at any depth in all pavement layers.
- A tar primer should not be applied to a chemically stabilised layer sooner than about seven days after compaction, nor, indeed, any highly penetrating primer until the layer has been cured by other means (NITRR 1986b: 50).
- A fluid cutback primer such as MC=30 should not be permitted as a curing membrane as it is ineffective at prevent-
ing both moisture loss (Bofinger et al 1978 in TRL & ODA 1993) and carbonisation (Netterberg et al 1987) for more than a few days. Even a prime such as MC-30 can weaken a cemented material (HRB 1949; Netterberg 1994) possibly due to hindrance of the cement hydration over the depth of penetration (RRL 1952). The surface moisture content should be such as to prevent penetration (HRB 1949).
- However, this may cause water vapour blistering in hot weather (Netterberg 1994).
- Additional precautions may be required when utilising marginal or substandard materials (Netterberg et al 1989).
- Embedment seldom occurs on a crushed stone base. Embedment due to a false layer on a base can be prevented by either carrying out the final compaction with a diluted emulsion or ripping up the top 20 mm after compaction and treating it with diluted emulsion (Bergh 1979). In both cases the prime is omitted.

CONCLUSIONS

The experience of the authors, HVS testing and careful DCP testing and carefult testing and careful DCP testing have shown that the presence of weak layers, interlayers, laminations and/or interfaces is far more deleterious to structural road pavement performance than is commonly assumed. In particular, a thin interlayer of weaker material between a cemented base course and a thin surfacing can lead to extensive distress within weeks or months of opening to traffic. Such distress
can usually be predicted by means of careful inspection, as well as by careful DCP testing. The creation of such weak layers, interlayers and/or laminations must therefore be prevented at all costs, even at the expense of other aspects such as surface finish. The causes of weak layers, interlayers, laminations and/or interfaces are many and some may have more than one cause, but most can be prevented simply by application of known good construction practices. However, such conditions are not always prevented by the application of the standard precautions. Engineering judgement and stricter application of visual inspection, supplemented by simple field tests, are therefore required.

ACKNOWLEDGEMENT

This paper is published with the approval of the Executive Director of the CSIR Built Environment Unit.

REFERENCES


Kley, E G 2012. TPA Graph 82.03. Personal communication.


NITRR (National Institute for Transport and Road Research) 1986c. *Prime coats and bituminous curing membranes*. TRH 1, Pretoria: CSIR, NITRR.


RRL (Road Research Laboratory) 1952. *Soil mechanics for road engineers*. London: HMSO.


TRL & ODA (Transport Research Laboratory and Overseas Development Administration) 1993. A guide to the structural design of bitumen-surfaced roads in tropical and sub-tropical countries, 4th ed. Overseas Road Note 31, Overseas Centre, TRL, Crowthorne, UK.

Mechanistic modelling of weak interlayers in flexible and semi-flexible road pavements: Part 2

M de Beer, J W Maina, F Netterberg

This paper (Part 2 of a two-part set of papers) discusses models and illustrates the adverse effects of weak layers, interlayers, laminations and/or weak interfaces in flexible and semi-flexible pavements, also incorporating lightly cemented layers. The modelling is based on mechanistic analyses for pavement design and evaluation. In Part 1, the effects of these relatively weak layers, interlayers, laminations and/or weak interfaces were discussed. It was shown that methodologies are available to detect and investigate the existence of these weak layers in cemented pavement layers. In Part 2, seven cases of the above conditions for different road pavement types are discussed, with field examples. Mechanistic analyses were done on a typical hot mix asphalt (HMA), several cases of a cemented base pavement and a granular base pavement, with and without these weak layers and interface conditions to demonstrate their adverse effects. The analyses focus on the strain energy of distortion (SED) as a pavement response parameter to indicate the potential for structural damage expected within the pavement structure or layer. Generally, the higher the SED, the higher the potential damage in the pavement layer. SED shows the potential for structural damage expected within the pavement structure or layer. Without these weak layers and interface conditions to demonstrate their adverse effects, the analyses focus on the strain energy of distortion (SED) as a pavement response parameter to indicate the potential for structural damage expected within the pavement structure or layer. Generally, the higher the SED, the higher the potential damage in the pavement layer.

Note
The 17 photos and 12 figures are numbered continuously throughout Part 1 and Part 2 of this two-part set of papers. However, the references and equation numbers are specific to each part.

INTRODUCTION
Low-strength, weak layers, interlayers, laminations and/or weak interfaces in the upper layers, such as bases and/or subbases, of flexible and semi-flexible road pavements are specifically prohibited in most specifications, as was discussed in detail in Part 1 of this paper (Netterberg & De Beer 2012 – see page 32 of this edition). For detailed definitions of the weak layers etc. and their associated conditions, see Part 1. As discussed in Part 1, premature pavement distress in the form of ripping, arcuate slippage, pumping, cracking or shoving of the bituminous surfacing and shallow base failures of cemented base pavements are not rare in southern Africa (see Photos 1 to 11 in Part 1). It is therefore important to discuss these observations and associated conditions in more detail in order to improve the quality of road construction in southern Africa. In particular, the aim of Part 2 is to discuss and quantify the adverse effects of weak layers, interlayers, laminations and/or weak interfaces in pavement layers. This is done by means of selected examples of full-scale testing with the heavy vehicle simulator (HVS), including the so-called traffic-associated “crushing failure” found in the top of lightly cemented base layers. The discussion further includes detailed mechanistic analyses of various types of road pavement and interlayer conditions, used (and proposed) here to study the adverse effects of such layers.

As indicated in Part 1, the presence of these layers and/or conditions at any depth in the structural layers of a flexible or semi-flexible pavement (but especially the upper base) can therefore be quantified, albeit on a relative basis.

FULL-SCALE HEAVY VEHICLE SIMULATOR (HVS) TESTS ON SELECTED PAVEMENTS WITH CEMENTED LAYERS
Equipment such as the HVS is very useful in identifying the modes of failure of full-scale pavement structures, as is indicated in the following sections.
Kleyn et al (1985) reported on one of the first HVS tests on a shallow but thin (< 100 mm) strongly cemented base pavement on Road P95/1. It was shown that the sensitivity of the pavement increased dramatically when the traffic loading was increased, with a further increase in the rate of deformation when the layers were in a wet state. The layerworks of this pavement were not ideal, as shown in Photo 12 (see Part 1). It is postulated here that, in addition to the fact that this pavement was of relatively strong but thin and therefore shallow design, the rather poor layerworks (with some weaker layers present) might have influenced the structural performance of this pavement. The former should be seen relative to the behaviour of deeper pavement structures incorporating cemented base/subbase layers, based on the rate of deformation, which was about 50 times lower than that of the shallow pavement above (Kleyn et al 1985, Table 3).

Opperman (1984) and later Kleyn et al (1985) discussed the structural performance of a strongly cemented gravel base and subbase pavement (Road P30) at Hornsnek, outside of Pretoria. This pavement started pumping after six years in service, and an investigation in 1984 showed the presence of a weak interlayer, unbound and probably a result of poor construction according to Opperman (1984), and not a case of carbonation (for a definition of “deleterious” carbonation in cemented road layers, see Part 1). Photo 1 (see Part 1) indicates the pumping under normal trafficking at that time. In this case, there was a weak interlayer with a thickness of between 25 and 80 mm between the cemented base and subbase, 105 mm to 215 mm from the surface. Owing to its thickness, this weak interlayer was also identifiable with the dynamic cone penetrometer (DCP). During HVS testing in the wet state, severe pumping of fines occurred on the pavement surface from this weaker interlayer, as shown in Photo 16. The cemented base also “collapsed” on the cemented subbase after the weaker interlayer had been “removed” by pumping (Kleyn et al 1985, Figure 6). This created a “deeper” pavement structure, demonstrating a “re-balancing” of the pavement under the action of traffic loading, especially in the wetter state. However, this took place at the cost of a rougher pavement surface (maximum deformation > 20 mm within 750 000 standard 40 kN load repetitions in the wet state) (Opperman 1984, Figure 8) with an associated lower functional riding quality as shown in Photo 17. This situation could have been avoided if the weaker layer had not been present in the initial construction of this pavement.

Asphalt base pavements on lightly cemented base and subbase layers: KwaZulu-Natal
This phenomenon was also observed and modelled during HVS tests on flexible asphalt base pavements, incorporating lightly cemented (C3/C4) subbase layers in KwaZulu-Natal (De Beer 1985, Chapter 5). In a case study by De Beer (1985) with the HVS on Van Reenen’s Pass, it was found that the full thickness of a 150 mm upper cemented subbase of an asphalt base pavement was relatively weak and fully carbonated. This particular layer was constructed during snowfalls in 1981, which hampered the construction control on site. This caused a “sandwiched” type of pavement structure, in which this weaker layer dominated the structural fatigue failure of the rather stiff recycled asphalt base of this pavement, even in the dry state (see Photos 18 and 19). Again, this illustrates the adverse effect of a weak layer within the upper supporting pavement layers. In another study De Beer (1985) showed the effect of erosion in the wet state of a cemented Berea Red Sand upper subbase underneath an asphalt base. In this case, a weak layer was created outside the wheel path, between the upper subbase and the asphalt base. In order to avoid
Fully carbonated and relatively weak upper cementitious subbase layer directly underneath the recycled asphalt base layer, resulting in a "sandwich" structure of a weak upper subbase on the N3 at Van Reenen’s Pass (De Beer 1985)

Stellenbosch and Somerset West, Jordaan (1988) found that weak interlayers developed within the cemented base after traffic loading with the HVS, and also found a second weak interlayer just beneath the surfacing. In addition, crack activity measurements by Rust (1987) complemented the analysis of the traffic-associated failure of this pavement. Photo 20 indicates severe pumping during the HVS test on this road section in the wet condition. In addition, the adverse effect of moisture in the upper layers (i.e. wet state) was also demonstrated here, in that the bearing capacity of this pavement was reduced from 8 million standard axles in the dry state to a mere 1.6 million standard axles in the wet state (Rust 1987). This is another illustration of the combined action of traffic loading and moisture on these pavements, with reported evidence of weak interlayers.

**Discussion**

The above examples (as well as those discussed in Part 1) illustrate the importance of the actual pavement situation in the field versus the pavement structure as designed, specifically the occurrence and effect of weak layers, interlayers, laminations and/or weak interfaces in actual pavement structures. It should also be noted that, once surface failure is initiated for whatever reason, accelerated pavement failure occurs. This, of course, is also a strong function of road roughness and associated vehicle suspension, and therefore of axle and tyre dynamics, as indicated by Lourens (1995) and also Steyn (2001). This situation is aggravated when the pavement is in a wet condition. In addition to the examples given above, this accelerated damage progression on various types of pavement has been adequately demonstrated and documented through many years of HVS testing and associated observations of pavement failures in South Africa (Maree 1982; Freeme et al 1987; Rust et al 1997; Du Plessis et al 2008).

To aid in pavement evaluation for the purposes of rehabilitation, Freeme (1983) and Freeme & De Beer (1983) published documents defining the structural behavioural states of different flexible and semi-flexible pavement types (by catalogue) to inform engineers about specific evaluation and rehabilitation design strategies (see also Jordaan 1988). These documents demonstrate that relatively weak layers between structural layers in a pavement are highly detrimental to the pavement’s structural performance, and become even more crucial in wet conditions, as was discussed above. During wet conditions, the pavement easily changes its structural behaviour into what is referred to as the “moisture accelerated distress” (MAD) state, as reported by De Beer & Horak (1987). In the next section the so-called “crushing failure” of cemented base pavements is discussed.

**CRUSHING FAILURE MECHANISM FOUND FOR CEMENTED BASE PAVEMENTS**

**Effect of tyre–pavement contact stress and associated crushing failure mode of a cemented layer**

Lightly cemented base/subbase layers may also suffer from traffic-associated “crushing failure”. This failure mode could result from a relatively weak upper portion in cemented bases, but could even occur in relatively dry conditions (De Beer 1990). This “weakness”
in the layer is to be seen relative to the applied contact stress it needs to carry from especially truck tyres. In this paper, it is assumed that the tyre inflation pressure is equal to the vertical tyre contact stress. However, a multitude of studies have indicated that three-dimensional (3-D) tyre–pavement contact stresses are not simple, and special provision should ideally be made to incorporate these complex contact stresses into pavement design (De Beer et al. 1997, 1999, 2002, 2004; De Beer 2006, 2008). In these days of higher tyre inflation pressures (and hence tyre vertical contact stresses of up to 1 000 kPa or more), the importance of avoiding a relatively weak upper base is well illustrated by Equations 1 and 2 (also see Figure 2). Figure 2 and these equations show the empirical relationships found for tyre contact stress-associated crushing failure in terms of the “initiation” of crushing failure (Crush Initiation, $N_{ci}$), as well as “advanced” crushing failure (i.e. Advanced Crushing, $N_{ca}$) (De Beer 1989a, 1989b, 1990; De Beer et al. 1997, 1999).

Crush Initiation ($N_{ci}$)  
$$= 10^{8.21 \times \left(1 - \frac{\sigma_t}{2.2 \times UCS_{50}}\right)}$$

Advanced Crushing ($N_{ca}$)  
$$= 10^{9.20 \times \left(1 - \frac{\sigma_t}{2.2 \times UCS_{50}}\right)}$$

Where:
- $N_{ci}$ = initiation of average crushing life span (units in repetitions of tyre–pavement contact stress, $\sigma_t$)
- $N_{ca}$ = advanced average crushing life span (units in repetitions of tyre–pavement contact stress, $\sigma_t$)
- $\sigma_t$ = tyre contact stress (~ tyre inflation pressure in this paper) in kPa
- $UCS_{50}$ = in situ DCP-derived unconfined compressive strength (UCS) of the top 50 mm of base layer in kPa

The probability for $N_{ci} = 50\%$, $R^2 = 0.89$, $n = 23$ (De Beer 1990).

For example, Figure 2 indicates that about one pass of a tyre resulting in an average vertical contact stress of $\sigma_t = 1 000$ kPa on a cemented base layer would immediately initiate the crushing mode of failure ($N_{ci}$) of the upper base (~ 50 to 75 mm) with a similar in situ $UCS_{50}$ strength. Further, a $UCS_{50}$ of about 3 MPa would be needed for 1 million contact stress repetitions before initiation of crushing failure in the upper 50 to 75 mm. The crushing (or compres- sion) failure, even in the dry state, of a cemented base layer, as was observed under HVS testing, is illustrated in Photos 21, 22 and 23. The existing C3 base strength
requirement of UCS = 1.5 MPa (strength in the field) should therefore be adequate to carry about 1 million contact stress repetitions at $\sigma_t = 500$ kPa, and approximately 10 000 stress repetitions at a tyre contact stress of $\sigma_t = 1000$ kPa. These relationships (Equations 1 and 2) were empirically derived from DCP tests at HVS sites (De Beer 1990). Further, as stated in the original work on the crushing failure mode, the initiation of crushing in thinly sealed cemented base pavements almost corresponds to the fatigue life of the thin asphaltic surfacing seal (De Beer 1989a, 1990). Advanced crushing ($N_{cr}$) corresponds to approximately 10 mm of deterioration (downward displacement) in the top of the cemented base, which is a strong warning sign of a shallow failure that may need urgent maintenance.

As shown in Part 1, it is therefore concluded that the creation (or existence) of relatively weak (crushed) upper layers, interlayers, laminations and/or interfaces due to any cause should be avoided at all costs, even at the expense of other aspects such as the quality of the surface finish.

**MECHANISTIC PAVEMENT ANALYSIS WITH AND WITHOUT INTERLAYER SLIP**

It is generally accepted that weak layers, laminations and/or weak interfaces of any kind in the upper base layer, including compaction planes, are highly detrimental to the structural and functional performance of a pavement, as shown by their general prohibition in nearly all specifications. Indeed, it can be quantitatively shown on a relative basis, based on the mechanistic pavement analysis methodology, that relatively smooth interfaces (see Photos 6, 7 and 8 in Part 1, and 24 and 25 alongside) and relatively weak interlayers (see Photos 12, 13 and 15 in Part 1) are highly detrimental at any depth in the base layer, and even at depth such as between the base and subbase layers. This is largely due to a redistribution of stresses and strains (or energy) as a direct result of a degree of “imbalance” caused by these weaker layers and/or weak interface conditions (e.g. AUSTROADS 1998; De Beer 1985, Chapter 5; Romanoschi & Metcalf 2001a, 2001b; Nageim & Hakim 1999).

For the case of a weak interface, a simple analogy would be the improved elastic deflection characteristics obtained by gluing the laminations of a laminated timber beam. An extreme analogy, applicable also to horizontal sliding failure in the upper base, would be the difference between a glued and an unglued ream of paper. Both crushing failure of a weak upper base and

**Photo 23** Final state of the lightly cemented base layer after HVS testing and crushing, in the wet state (De Beer 1990). Note the amount of material lost from the wheel path in centre of photo

**Photo 24** Field cores from lightly cemented base layer – left core: solid; right core: weak interlayers

**Photo 25** Signs of delamination (possible construction plane) in a lightly cemented layer (De Beer 1990)
sliding on interfaces between laminations or weak interfaces appear to have contributed to many “shallow” base failures of cemented bases in southern Africa, though both have not necessarily occurred at each failed site.

In the following section examples of various mechanistic analyses show the relative effects of interlayers (weak layers, laminations and/or weak interfaces), based on the theoretical interlayer slip models developed by Maina et al. (2007). The relative importance of a laminar or weak interface of thickness $t = 0$ versus relatively weak layers, or interlayers of a definite thickness $t > 0$ versus simple weakening of a layer under the action of trafficking, is discussed using the mechanistic approach.

**Strain energy of distortion (SED) as pavement response parameter**

According to Timoshenko & Goodier (1951), the quantity of strain energy stored per unit volume of the material can be used as a basis for determining the limiting stress at which failure occurs. For this to be applied to isotropic materials, it is important to separate this energy into two parts – one due to the change in volume and the other due to the distortion – and to consider only the second part in determining the strength. Whatever the stress system, failure occurs when the SED reaches a certain limit. The mechanistic analyses discussed in this paper focus on SED as a pavement response parameter in terms of its value as an indicator of potential damage in a pavement layer. Generally, the higher the SED value, the higher the potential for damage in the pavement layer. SED shows some potential for quantifying the relative effect of these deleterious layers/conditions of interlayers in flexible and semi-flexible pavements. Theoretically, according to Timoshenko & Goodier (1951), the total strain energy per unit volume, $V_0$, is expressed by Hooke’s law, as follows:

$$V_0 = \frac{1}{2E}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) - \frac{\nu}{E}(\sigma_x \sigma_y + \sigma_y \sigma_z + \sigma_z \sigma_x)$$

$$+ \frac{1}{2G}(\tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2)$$

so that SED can then be expressed as follows:

$$S E D = V_0 = \frac{1 - 2\nu}{6E}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2)$$  \hspace{1cm} (4)

Where:

- $E = $ Young’s Modulus (MPa)
- $\nu = $ Poisson’s Ratio
- $G = $ shear modulus (MPa)
- $\sigma = $ compressive or tensile stress (MPa)
- $\tau = $ shear stress (MPa)

Using Equations 3 and 4 it is anticipated that locations within the pavement structural layers that have relatively higher values of SED (i.e. so-called “hot spots”) will potentially fail first before points with relatively lower SED values. Note that the unit of SED is N-m/m³.

**Example 1:**

**Pavement type: hot mix asphalt (HMA) base: smooth vs rough interface between top of subbase and bottom of asphalt base – vertical tyre loading only**

In this section, an example of a lamination or interface (modelled as “smooth” or “full slip”) at the bottom of a typical HMA base airport pavement is demonstrated, relative to the case of full-friction (modelled as “rough” or “no slip”) condition by way of mechanistic modelling. Figure 3 shows an evaluation of interlayer slip conditions, i.e. a weak interface, which is modelled as a layer (or boundary condition) without any thickness, $t = 0$, for a typical flexible airport pavement structure (Maina et al. 2007). For the purpose of this paper, the results of utilising “Slip Model 3” (Equation 13 in Maina et al. 2007) are given since it seems to be more appropriate than other slip models investigated by Maina et al. 2007. The variation of SED with depth under the front outermost tyre of a Boeing 747-400 gear (of 16 tyres) (after Maina et al. 2007) for three slip rate values (indicating the percentage of horizontal movement allowed theoretically) is clearly

**Figure 3 Variation of maximum SED with depth under the front most outside tyre with vertical and horizontal tyre loading on 16 tyres of a Boeing 747-400 – modelled on a typical asphalt base airport pavement with and without laminated at 50 mm depth (after Maina et al 2007)**

**Table 3 Variation of maximum SED with depth under the front most outside tyre with vertical and horizontal tyre loading on 16 tyres of a Boeing 747-400 – modelled on a typical asphalt base airport pavement with and without laminated at 50 mm depth (after Maina et al 2007)**
Table 1 Pavement structure and associated engineering properties used for the mechanistic analysis – Road Category B, ES3 (TRH 4, COLTO 1996)

<table>
<thead>
<tr>
<th>Layer</th>
<th>Thickness (mm)</th>
<th>E modulus (MPa)</th>
<th>Poisson’s ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surfacing (S2)</td>
<td>15</td>
<td>1 000</td>
<td>0.44</td>
</tr>
<tr>
<td>Cemented base (C3)</td>
<td>50</td>
<td>3 000/30 **</td>
<td>0.35</td>
</tr>
<tr>
<td>Cemented base (C3)</td>
<td>75</td>
<td>3 000</td>
<td>0.35</td>
</tr>
<tr>
<td>Cemented subbase (C4)</td>
<td>200</td>
<td>1 500</td>
<td>0.35</td>
</tr>
<tr>
<td>Selected layer (G7)</td>
<td>150</td>
<td>250</td>
<td>0.35</td>
</tr>
<tr>
<td>Subgrade soil</td>
<td>=</td>
<td>100</td>
<td>0.35</td>
</tr>
</tbody>
</table>

* For material and pavement codes, see TRH 4 (COLTO 1996)
** Representing a relatively weak layer approximated with a rather low effective elastic modulus (i.e. 30 MPa), compared with a “solid” and intact layer with an effective modulus of ~ 3 000 MPa.

Figure 4 Layout for mechanistic analysis of a cemented base pavement (this study)

shown in Figure 3. The trend of maximum SED indicates that, whatever the slip rate value, relatively higher SED values are found at the surface of the pavement structure. Note that a slip rate of 0.0 represents “full friction” (no slip or full bonding) between subsequent horizontal layers, and a slip rate of 0.9 represents “full slip” (almost no friction or no bonding for practical purposes). Figure 3 shows that as the slip rate increases, SED at the bottom of the top layer (where the interface slip occurs) increases to almost 75% of the SED obtained at the top of the pavement layer and this may result in a potential “bottom-up” failure condition.

Based on the SED analysis, interlayer slip may therefore influence the position and type of potential asphalt layer failures, such as cracking and deformation. Less drastic changes in SED occurred at the bottom of layer 2 as a direct result of the “re-balancing” of the stress and strain distribution within the pavement layers. The effect of this re-balancing is that the potential for distortion (i.e. SED) is shifted to the top layers above the lamination (or weak interface), which may result in cracking or stress concentration.

The general trend in SED with depth in Figure 3 is that, as the slip rate increases, SED tends to decrease through the thickness of the layer up to a certain depth within the asphalt base (in this particular case at a depth of approximately 30 mm), and then starts increasing again towards the bottom of the layer. In summary therefore, damage is expected to occur on the road surface (most probably rutting), as well as from the bottom of the asphalt base (most probably dominated by fatigue cracking, not excluding rutting from the bottom). This theoretical finding, however, needs further field verification.

Example 2: Pavement type: cemented base: smooth vs rough single interface within the cemented base – vertical tyre loading only

In this section an example of a single lamina- or weak interface (“smooth” boundary condition, also referred to as “full slip” or “no friction”) at a depth of L = 65 mm (i.e. 15 mm + 50 mm) in the cemented base (Road Category B, ES3, TRH 4 dated 1996, see COLTO 1996) is demonstrated. Slip Model 3 from Maina et al. (2007) was also used for this analysis and the engineering model parameters are given in Table 1. As already indicated, the design capacity of this pavement is ES3 = 3 million standard axles (MISA) over 15 years. Note that MISA = million repetitions of a standard 80 kN axle with four tyres at 520 kPa tyre pressure, as defined in TRH 4 (COLTO 1996).

The load configuration is the standard loading per tyre, which is 20 kN and a value of 520 kPa for the contact stress. The layout of the pavement and its tyre loading used in the mechanistic analysis is illustrated in Figure 4. This figure also shows the different layer thicknesses and the depth of the lamination in the cemented base layer, which was modelled at a depth of L = 65 mm (i.e. 15 mm + 50 mm).

For the purposes of illustration, only one phase (i.e. the initial “pre-cracked” phase) of the structural behaviour of cemented layers is used here (for the different phases in the “life” of a cemented layer, see Theyse et al. 1996). The SED results with vertical loading (V) only, under one tyre (Figure 4, Position A) and between the dual tyres (Figure 4, Position B) for this pavement condition are illustrated in Figure 5. Note that the various mechanistic analyses were done at two different positions, i.e. under the tyre (Figure 4, Position A) and between two tyres (Figure 4, Position B), and the SED values are reported through the depth of the pavement. Note that the case of vertical tyre loading (V) only, is considered in Figure 5.

The highest relative SED value of approximately 130 N-m/m³ (see red-dot curve) occurs at a position directly underneath the tyres at the depth (L) of the lamination (or weak interface), which is at L = 65 mm, from the surface of the pavement. The position of the second highest SED value (~ 28 N-m/m³, see
In Figure 5 (weak interlayer in the base, also indicated for the ideal case of full friction and no SED values (~ 5 N-m/m³ at 65 mm) found this is in sharp contrast to the relatively low "smooth" or full-slip condition at this depth. Most of the other SED values are less than about 25 N-m/m³. Note that the higher SED values (E = 30 MPa) are situated. This is also in sharp contrast to the relatively low effective elastic modulus of 30 MPa) is quite dramatic. In this case, the maximum SED increased to a value close to 3 500 N-m/m³, which is approximately 30 times larger (Figure 6) than the case of laminating (full-slip weak interface) only, as shown in Figure 5.

The dramatic 30-fold increase in SED caused by this 50 mm weak layer clearly illustrates the higher potential for distortion (fatigue, rutting or further crushing) that may occur in the top 65 mm of this cemented layer. This, of course, could lead to premature failure, even during the construction period, if such a condition exists. Note also that most of the SED is concentrated in the top part of the cemented base (under the tyres, Position A) where the lamination, weak interface and weak layer are situated. This is also in sharp contrast to the ideal case of full friction and no weak layer within the base as indicated in Figure 5 (green curve), with maximum SED – 5 N-m/m³ at 65 mm.

For this reason, a relatively weak layer of 50 mm thickness (t = 50 mm) in the top part of the base layer. This condition is modelled with and without a lamination (or weak interface) at the bottom of this weak layer, with depth L = 65 mm. Again, for demonstration purposes only, vertical tyre loading (V) is considered here.

Compared with the results shown in Figure 5, the effect of adding a 50 mm relatively weak layer (modelled here as a layer of relatively low effective elastic modulus of 30 MPa) is quite dramatic. In this case, the maximum SED increased to a value close to 3 500 N-m/m³, which is approximately 30 times larger (Figure 6) than the case of laminating (full-slip weak interface) only, as shown in Figure 5.

The dramatic 30-fold increase in SED caused by this 50 mm weak layer clearly illustrates the higher potential for distortion (fatigue, rutting or further crushing) that may occur in the top 65 mm of this cemented layer. This, of course, could lead to premature failure, even during the construction period, if such a condition exists. Note also that most of the SED is concentrated in the top part of the cemented base (under the tyres, Position A) where the lamination, weak interface and weak layer are situated. This is also in sharp contrast to the ideal case of full friction and no weak layer within the base as indicated in Figure 5 (green curve), with maximum SED – 5 N-m/m³ at 65 mm.

Example 3: Pavement type: smooth vs rough single interface within the cemented base, including a relatively weak layer – vertical tyre loading only

In this example, as Figure 6 illustrates the SED results of the same cemented base layer pavement discussed above, but including a relatively weak layer of 50 mm thickness (t = 50 mm) in the top part of the base layer. This condition is modelled with and without a lamination (or weak interface) at the bottom of this weak layer, with depth L = 65 mm. Again, for demonstration purposes only, vertical tyre loading (V) is considered here.

Compared with the results shown in Figure 5, the effect of adding a 50 mm relatively weak layer (modelled here as a layer of relatively low effective elastic modulus of 30 MPa) is quite significant. In this case, the maximum SED increased to a value close to 3 500 N-m/m³, which is approximately 30 times larger (Figure 6) than the case of laminating (full-slip weak interface) only, as shown in Figure 5.

The dramatic 30-fold increase in SED caused by this 50 mm weak layer clearly illustrates the higher potential for deformation (fatigue, rutting or further crushing) that may occur in the top 65 mm of this cemented layer. This, of course, could lead to premature failure, even during the construction period, if such a condition exists. Note also that most of the SED is concentrated in the top part of the cemented base (under the tyres, Position A) where the lamination, weak interface and weak layer are situated. This is also in sharp contrast to the ideal case of full friction and no weak layer within the base as indicated in Figure 5 (green curve), with maximum SED – 5 N-m/m³ at 65 mm.

Example 4: Effect of horizontal (H) loading – position between the dual tyres (Position B in Figure 4)

To simulate the additional effect of turning/scuffing of tyres around sharp corners or curves, dual tyres with both vertical (V)
loading and with and without horizontal (H) loading in the same direction under both tyres, are considered. For demonstration purposes, H was assumed to be about 0.5 of V, i.e. H = 0.5 V (Figure 4). So, with V = 20 kN and H = 10 kN, the pavement was re-analysed in this example. Figure 7 illustrates the effect of SED for the given loading conditions at a position between the dual tyres (Figure 4, Position B), and represents a case where no slip (“rough”) and full slip (i.e. lamination or weak interface) are considered. Although the SEDs obtained are relatively low for the “no slip” case (yellow and dark blue curves), the effect of SED can be seen within the cemented base layer, peaking at the top and bottom of the layer. Also note the increased SED within the cemented base with the horizontal load (compare the yellow and dark blue curves) between the depths of approximately 15 and 140 mm in Figure 7, suggesting some additional distortion energies within the cemented base layer under conditions of horizontal loading between the two tyre loadings. This is a direct result of the horizontal loading under the tyres, which has a potentially adverse effect (increased SED and therefore potentially increased damage) midway between the tyres for the case of no slip (or full friction).

As before, in the case of lamination or a weak interface (full slip) at L = 65 mm, SED peaks at this depth, but with a much-reduced SED of approximately 28 N-m/m³ (light blue and red curves), compared with the case under the tyres where the SED was approximately 130 N-m/m³, shown by the red curve in Figure 5. It is interesting to note that at this depth and position midway between the tyres (Position B, Figure 4) the additional effect on SED from horizontal loading seems to be relatively small (compare red and light blue curves in Figure 7).

**Example 5:**

**Effect of horizontal (H) loading – position directly under tyres (Position A in Figure 4)**

Figure 8 illustrates the SED results under one of the dual tyres (Position A). Here the peak SED value is also at the depth of lamination or full slip at L = 65 mm, as before, and is approximately 130 N-m/m³ for vertical loading only. This value increases to approximately 147 N-m/m³ if horizontal loading is also included. Note further the increased SED on the surface of the pavement, directly under the tyres, from a value of less than 5 N-m/m³ to approximately 100 N-m/m³, if horizontal loads are included. This suggests the relatively high potential for additional failure potential in the top 65 mm layer under conditions of turning (i.e. “scuffing”), acceleration or deceleration of the tyres on the surface of the pavement. This again shows that a condition of slip in the cemented base layer is to be avoided and, as discussed in Part I, this should be assured through proper inspection during construction and suitable repairs before these cemented base layers are sealed.

In Figure 9 the same situation as in Figure 6 (pavement with lamination together with a 50 mm weak layer) is considered, but with the additional horizontal (H) loading included in the analysis. Note the relatively small increase in SEDs compared with those given in Figure 6 for all cases investigated here. Based on the foregoing, it seems that the weak layer in itself overshadows the effect of additional horizontal loading under the tyres in this example. Again, the SED is concentrated in the top part (top 65 mm) of the pavement as modelled here, suggesting shallow failure potential from the top of the layer.
and L = 140 mm, together with a weak layer in the same base layer) at depths L = 65 mm.

In this example, the effects of a double slip layer in depth – vertical tyre load only – was investigated. The layout of the pavement and its tyre loading used in the mechanistic analysis are shown in Figure 11. The objective here is to illustrate the effect of a 20 mm relatively weak layer in the top of the upper part of the cemented subbase supporting the G1 base layer by way of analysing the associated SED patterns. First, the pavement was analysed without any weak layer (i.e. solid C3 subbase). The results indicate a peak in SED ~ 100 N-m/m³ towards the middle of the G1 base layer at a depth of approx. 60 mm (blue curve in Figure 12). However, when a 20 mm weak layer is introduced directly underneath the G1 base layer, the SED increases greatly at the bottom of the G1 base layer, as well as within the 20 mm weaker layer (purple curve in Figure 12).

From experience, owing to pavement "strength balancing" under the action of tyre loading (or traffic moulding), the G1 granular layer will start to de-densify (or deteriorate) to a relatively lower quality and therefore become a less dense granular layer. This case is modelled here by a reduced effective elastic modulus of 200 MPa (changed from the original 500 MPa (~ G1)), representing a "weakened" G1 granular layer for illustration purposes. The 20 mm weak layer at the top of the upper cemented subbase (C3) is represented here by a weak interlayer (due to poor construction or deleterious chemical breakdown), modelled with an effective elastic modulus of 30 MPa. When the weak layer on top of the subbase, together with this "weakened" G1 base layer, is analysed, the SEDs in the granular at L = 65 and 140 mm, and the 75 mm thick weak layer at the bottom of the cemented base, appear to be only one third of those obtained for the weak layer only in the top 50 mm and with lamination at L = 65 mm (~ 1 000 N-m/m³ vs ~ 3 500 N-m/m³) (Figure 9).

**Example 7: Pavement: granular base on cemented subbases – with and without weak layer and/or slip**

In this final example, a relatively strong pavement with a typical high-quality granular (G1 material (COLTO 1996) base layer supported by two lightly cemented layers is investigated. The layout of the pavement and its tyre loading used in the mechanistic analysis is shown in Figure 11. The objective here is to illustrate the effect of a 20 mm relatively weak layer in the top of the upper part of the cemented subbase supporting the G1 base layer by way of analysing the associated SED patterns. First, the pavement was analysed without any weak layer (i.e. solid C3 subbase). The results indicate a peak in SED ~ 100 N-m/m³ towards the middle of the G1 base layer at a depth of approx. 60 mm (blue curve in Figure 12). However, when a 20 mm weak layer is introduced directly underneath the G1 base layer, the SED increases greatly at the bottom of the G1 base layer, as well as within the 20 mm weaker layer (purple curve in Figure 12).
base and the 20 mm weak layer increase to levels almost double those of the case of the 20 mm weak layer only (compare red with purple curve in Figure 12). Therefore, the adverse effect of a weak layer at the bottom of the G1 layer (or in the top 20 mm of the C3 upper subbase layer) will potentially result in a reduced pavement life, and/or premature failure of the G1 base layer itself. Such a situation could be greatly aggravated when moisture is added to the “weakened” granular base, leading directly to, for example, potholes and base deterioration with further traffic loading. As before, based on the theoretical considerations in this paper, it is therefore concluded that relatively weak layers and interlayers within the supporting structural layers of flexible pavements should be avoided as far as possible in order to ensure quality structural performance of these pavements. It is, however, accepted that further field evaluation and validation are needed to confirm the value of SED as an adequate descriptive mechanistic parameter for the quantification of structural pavement behaviour.

**SUMMARY, CONCLUSIONS AND RECOMMENDATIONS**

In this paper (Part 2 of the two-part set) it was shown that modelling of flexible and semi-rigid pavements, including weak interlayers, laminations and/or weak interfaces, by way of mechanistic analysis using strain energies of distortion (SED) is quite insightful. In addition, experience and HVS testing have shown that the presence of relatively weak interlayers of any kind (i.e. laminations and/or interlayers of relatively weak quality) is far more deleterious to structural pavement performance than is commonly assumed. Especially the crushing mode of shallow failure and its quantification show good promise for the evaluation and prediction of the structural capacity of pavements incorporating lightly cemented layers under these adverse conditions.

It is concluded that the effects of weak interlayers, laminations and/or weak interfaces can be modelled mechanistically to illustrate and quantify potential failure conditions and their locations. The effect of potential crushing failure of cemented pavement base/subbase layers can also be quantified, albeit empirically. These conditions in the structural pavement layers should (and can) be avoided, especially during construction, in order to ensure quality structural performance of pavements of this nature in the long term.

For the design and modelling of durable flexible and semi-rigid pavements in the context of this paper, the following aspects are recommended:

- In medium to high-risk situations, the mechanistic methodology should include an analysis based on the SED parameter, which may greatly improve engineering understanding of the failure mechanisms, structural behaviour and associated potential for premature failure of flexible and semi-rigid pavements. This is
especially important in cases where the upper layers of these pavements may contain weak layers, interlayers, laminations and/or weak interfaces.

- Full-scale pavement studies should be performed and evaluated to calibrate the SED as a potential failure parameter, especially for roads in the conditions described above.

- Emphasis should be placed on ensuring that the construction quality of these pavements is maintained. If a relatively weak layer, interlayer, lamination and/or interface is indeed found in the upper layers of these pavement structures, it should be removed and rectified immediately before any further construction is done, as was discussed in Part 1.

- Pavements incorporating cemented (stabilised) base layers should be evaluated for potential traffic-load-associated crushing failure related to higher tyre contact stresses as proposed in this paper. This is to avoid the shallow crushing failure of these layers when their strength is substandard relative to their strength specifications and the current demand of higher truck tyre contact stresses.

ACKNOWLEDGEMENT

This paper is published with the approval of the Executive Director of the CSIR Built Environment unit.

REFERENCES


The effect of parameters on the end buffer impact force history of the crane

T N Haas, P Mainçon, P E Dunaiski

An overarching investigation was conducted to provide engineers with guidelines for designing crane supporting structures. The focus of this study was to determine whether the identified parameters had an effect on the end buffer impact force history when the electric overhead travelling crane collides with the end stops of the supporting structure. Seven design codes which were reviewed do not consider the crane and its supporting structure as a coupled system. This simplification ignores some of the parameters which have a significant influence on the impact force, which could lead to the codified estimates being sometimes unconservative. During the experimental tests it was discovered that some of the parameters could not be accurately controlled and/or monitored. This led to the development of a finite element (FE) model of the full-scale experimental configuration which was used to conduct advanced simulations. The FE model considered the crane and the supporting structure as a coupled system, in which the parameters were individually varied to obtain its effect on the impact force history. The results showed that some of the individual parameters do have a significant effect on the impact force history.

INTRODUCTION
Electric overhead travelling cranes (cranes) are used predominantly in industrial buildings to move heavy or cumbersome equipment, sometimes under very demanding conditions. Cranes enhance the operational process in industrial buildings, thereby improving production and ultimately reducing the production cost of the manufactured item. Hoist loads in these environments range from inconsequential (less than half a ton) to several hundred tons. Therefore the members of the crane and the crane supporting structure must be designed to have sufficient strength and stiffness to prevent failure at ultimate limit state and to prevent excessive deflection and vibrations at serviceability limit state.

In order to prevent local or global failure of the crane supporting structure, the following forces must be accurately assessed:
- Horizontal longitudinal forces due to the acceleration and braking of the crane, and the crane colliding with the end stops of the supporting structure.
- Horizontal lateral forces due to skewing of the crane.
- Vertical wheel loads due to the hoist being lifted.

The work reported in this paper focuses on determining the horizontal longitudinal forces when the crane collides with the end stops of the supporting structure. Previously, crane operators believed that it was good practice to run the crane into the end stops for realignment. Although the practice continues, it is less frequent nowadays. The constant collision between the crane and the supporting structure weakens the connection at the end stops. It is thus important that the horizontal longitudinal force resulting from the collision must be resisted by the end stops of the supporting structure. Failure of the end stops will result in the disastrous consequence of the crane running off the crane rails, especially when heavy loads are being hoisted. The consequences are even more disastrous if the hoist load is molten steel, as happened recently in China, when a crane which lifted molten steel ran off the rails, causing the load to be splintered on the ground. Several workers were killed when they were engulfed by the molten steel.

Several codes of practice and guidelines for the design of crane supporting structures were reviewed, namely:
- Manufacturers’ guidelines: DEMAG
- Eurocode 1, Part 3, EN 1991
- South African National Standard: SANS 10160, Part 6
- Australian Standard: AS 1418.1, 1994
- Association of Steel and Iron Engineers’ technical report, Aise No 13, 1997

All the above codes of practice and guidelines consider the collision between the crane and the supporting structure as an accidental condition. This implies that this
condition is seldom expected to occur during the life of the crane supporting structure. These codes estimate the end buffer impact force using a decoupled approach. This approach reduces the complexity of determining the member forces in the crane supporting structure. The codes consider several parameters that are used to estimate the end buffer impact force, as shown in Table 1.

Table 1 shows the parameters which the design codes explicitly use to estimate the end buffer impact force. It is evident that the design codes predominantly use the impact velocity, the mass of the crane and the end buffer’s resilience to estimate the end buffer impact force. Other critical parameters which are omitted from the code specifications and guidelines are:

i. The mass of the hoist load and its vertical and horizontal positions at the moment of impact.
ii. The dynamic effects of the crane during impact.
iii. The longitudinal misalignment of one of the end stops.
iv. The effect of continuously running longitudinal motors during impact.

From Table 1 it is clear that the design codes do not consider the effect of all the critical parameters to estimate the end buffer impact force, and therefore, by ignoring these critical parameters, the design codes can substantially underestimate or overestimate the end buffer impact force. This study investigated which of the parameters listed in Table 1 had an effect on the end buffer impact force history. This information made it possible to determine whether the parameters should be included in a codified assessment of the end buffer impact force.

Besides the codes of practice, no other literature was found which directly relates to either experimental or numerical evaluation of end buffer impact forces.

This paper describes the experimental configuration, the codified end buffer impact force estimates, FE modelling of the crane and the supporting structure, and experimental and FE impact force history responses. Reasons for the discrepancies between the experimental and FE impact force history results are given. The paper ends with a summary and conclusions.

**Table 1 Parameters of the design codes used to estimate the end buffer impact forces**

<table>
<thead>
<tr>
<th>Code / Guideline</th>
<th>Impact speed</th>
<th>Impact speed reduction factor</th>
<th>Crane and crab mass</th>
<th>Payload mass</th>
<th>Dynamic factor to account for dynamic effects</th>
<th>Vertical position of the payload during impact</th>
<th>Horizontal position of the payload during impact</th>
<th>Damping characteristics (energy absorption of end buffers)</th>
<th>Longitudinal misalignment of the end stops</th>
<th>Power off (minimum load during impact)</th>
<th>Power on (maximum load during impact)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SABS 0160-1989</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Method (a) &amp;</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Method (b)</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>DEMAG</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EN 1991:3 &amp; SANS 10160-6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AS 1418.18 : 2001</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AS 1418.1 : 1994</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AISE No 13: 1997</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Figure 1** A view of the experimental model configuration

**Figure 2** Numerical layout of the crane supporting structure

**DESCRIPTION OF THE EXPERIMENTAL CONFIGURATION**

Figure 1 shows the full-scale experimental configuration of the 5-ton electric overhead travelling crane and the supporting structure. A brief description of the experimental configuration with reference to Figure 1 is now presented. The crane consists of a 305 × 305 × 118 H-section crane bridge which is 8.485 m long and two 203 × 203 × 60 H-section end carriages which are 4.140 m long with a lifting capacity of 5 tons (yellow structure). The hoist load consists of an in-fill lead concrete block which has a mass of 5.128 kg. DPZ 100 elastomeric cellular plastic buffers manufactured by DEMAG, are attached to the ends of the end carriages to absorb the impact during the collision (black sections at the ends of the end carriages). The elastomeric cellular buffer has an energy absorption capacity of 800 Nm with a corresponding deformation of 73 mm and a final end buffer impact force of...
An experimental investigation was conducted to determine the codified end buffer impact forces. As a result of this discrepancy, an investigation was conducted to gain a better understanding of the impact forces between the various members of the supporting structure. During the initial experimental investigations it was difficult to accurately control and measure the physical parameters, i.e. the pendulum action of the hoist load during acceleration and the constant velocity phases, misalignment of the end stops, the flexibility of the structure, the differential power output from the motor to the wheels and the differences in responses of the end buffers during impact. This difficulty can be attributed to the complex phenomena involved during the collision between the crane and the end stops of the supporting structure. A finite element (FE) model was therefore developed which considered the crane and supporting structure as a coupled system. The advantage of the FE model was that the parameters could easily be individually adjusted and controlled. Also, the entire experimental configuration was modelled to allow the other load models to be individually adjusted and controlled. Also, the entire experimental configuration was modelled to allow the other load models to be individually adjusted and controlled.

Determining the Codified End Buffer Impact Forces

The codified end buffer impact forces were determined using the full-scale experimental crane configuration parameters described earlier. In Figure 3 the codified end buffer impact forces are presented as a function of the impact velocity of the crane.

From Figure 3 it is clear that there is a large discrepancy between the codified end buffer forces, which led to the conclusion that the codified estimates are not properly understood. This is due to the different analysis philosophies and parameters considered by the various design codes of practice. As a result of this discrepancy, an investigation was conducted to determine which parameters contribute significantly to the end buffer impact force.

Methods

Experimental and FE Models

An experimental investigation was conducted to gain a better understanding of how the force resulting from the hoist load is transferred from the crane to the supporting structure. During the initial experimental investigations it was difficult to accurately control and measure the physical parameters, i.e. the pendulum action of the hoist load during acceleration and the constant velocity phases, misalignment of the end stops, the flexibility of the structure, the differential power output from the motor to the wheels and the differences in responses of the end buffers during impact. This difficulty can be attributed to the complex phenomena involved during the collision between the crane and the end stops of the supporting structure. A finite element (FE) model was therefore developed which considered the crane and supporting structure as a coupled system. The advantage of the FE model was that the parameters could easily be individually adjusted and controlled. Also, the entire experimental configuration was modelled to allow the other load models to be studied, i.e. vertical wheel loads and horizontal lateral loads.

Figure 4 places the measuring equipment of the experimental set-up in context in relation to the entire configuration. The measuring equipment was used to determine the impact force and displacement histories when the crane collided with the end stops.

Figure 4 shows the linear variable displacement transducer used to measure the compression of the buffer with a range of 100 mm and a resolution of 6.25 μm, the load cell used to measure the compression force induced by the crane on the end stops with a capacity of 50 kN with a resolution of 3N, DEMAG’s DPZ 100 end buffers which absorbed the energy during impact, and the end stops which prevented the load cell from running off the rails.

FE Model

Due to the complexity of the 5-ton crane and the supporting structure, many simplifications were required to obtain a computationally efficient FE model, i.e. an FE model that can conduct the simulations in the shortest possible time. A computationally efficient model was developed that properly captured the relevant physical system. Commercially available FE analysis software, ABAQUS version 6.5.4, manufactured by Dassault Systemes, was used to develop the FE model.

The purpose of the FE model was to generate accurate global forces and deflections of the members, as well as the contact forces between the various members of the coupled system. Since the stresses and strains...
A non-zero wheel friction in the transverse longitudinal direction (rolling direction), with modelled near zero wheel friction in the pulley. A contact formulation was used that and crane rails, as well as for the cable and the interactions between the crane wheels defined contact properties were required for DPZ 100 cellular plastic buffers. Correctly used to model the highly non-linear DEMAG elastic and damping characteristics were model. Spring elements with appropriate computed from the computationally efficient solid (brick) elements using the global forces of a specific element which consists mainly of it could be obtained using a detailed FE model and strains within the elements were required, model the columns and beams. If the stresses within the members were not of interest, the FE model consisted predominantly of three-dimensional shear flexible (Timoshenko) quadratic beam elements which were used to model the columns and beams. If the stresses and strains within the elements were required, it could be obtained using a detailed FE model of a specific element which consists mainly of solid (brick) elements using the global forces computed from the computationally efficient model. Spring elements with appropriate elastic and damping characteristics were used to model the highly non-linear DEMAG DPZ 100 cellular plastic buffers. Correctly defined contact properties were required for the interactions between the crane wheels and crane rails, as well as for the cable and pulley. A contact formulation was used that modelled near zero wheel friction in the longitudinal direction (rolling direction), with a non-zero wheel friction in the transverse direction. The purpose of this paper was to describe the effect of the parameters on the impact force history of the crane, and thus a detailed description of the FE model is not presented. However, the reader is referred to Haas (2010) for a detailed description of the modelling techniques used to obtain a computationally efficient FE model. The simplifications resulted in an FE model which had 1 642 elements and 3 391 nodes with approximately 20 350 degrees of freedom (DOF). The average analysis time, including hoisting, acceleration and impact steps, was approximately 20 minutes on a Pentium 4 desktop computer with 3 GB of RAM. Figure 5 shows a schematic view of the FE model.

**Description of the experimental tests and FE simulations**

Experimental tests and FE simulations were conducted to determine the histories of the end buffer impact force for the following hoist load conditions:

1. No hoist load.
2. Hoist load raised to 0.15 m above ground level (the minimum distance the hoist load could be lifted to clear any obstructions on the floor).
3. Hoist load raised to 2.20 m above ground level (the maximum distance the hoist load could be lifted).

For each of the three conditions, the hoist load was lifted to its respective height except for the condition of “No hoist load” before any tests were conducted. Once the hoist load and crane were free of any vertical vibrations, the crane was accelerated at 0.2 m/s² for 2.75 s to attain an impact velocity of 0.55 m/s. At the moment of impact the operator released the longitudinal acceleration button on the control pendant which allowed impact to occur as a result of the inertia of the crane and hoist load.

**Experimental Impact Force**

**Condition: No hoist load**

A proper understanding of the experimental impact force history was necessary before any calibration of the FE model could be done. Figure 6 shows the impact force histories for three different cases when the crane collided with the end stops without a hoist load, namely:

i. “Power-Off with residual torque”: For this case, the acceleration of the longitudinal motors of the crane was disengaged at the moment of impact. Most modern cranes have a step-down/step-up torque function which controls the torque transmitted to the longitudinal motors, thus preventing the crane from stopping immediately when the longitudinal acceleration is disengaged. For this case, the torque was transmitted in a decreasingly linear fashion to the wheels of the crane when the acceleration button on the pendant was released.

ii. “Power-On”: For this case, the acceleration of the longitudinal motors of the crane was engaged throughout the impact phase.

iii. “Power-Off without residual torque”: For this case, the acceleration of the longitudinal motors of the crane, as well as the torque step-down function, was disengaged at the moment of impact, i.e. no power was transmitted to the wheels on impact.

The impact force history is also affected by the disc brakes of the wheels which engage the moment the operator releases the longitudinal acceleration crane motor button. The disc brakes were also disengaged for all cases during the experimental and FE tests. At least three tests were conducted for each
case. Thus the impact force histories shown in Figure 6 represent a series of experimental tests that were conducted to obtain the end buffer impact force histories for the three cases. All the other experimental results presented represent a series of tests conducted per case.

The expected end buffer impact force history for each case is shown in Figure 6 and is discussed individually.

Case (i) “Power-Off with residual torque”
Only one impact was expected when the crane without the hoist load collided with the end stops. Figure 6 shows two additional peaks which occurred after the first impact. The secondary peaks are due to the variably adjusted step-down torque present in the longitudinal drive motors of the crane on impact, which propelled the crane back into the end stops.

Case (ii) “Power-On”
To determine the effect of the residual torque on the impact force history, a second series of tests were conducted when the crane collided with the end stops with the longitudinal drive motors fully engaged over a period of time. Figure 6 shows a comparison of the experimental tests for the “Power-Off with residual torque” and “Power-On” conditions. The difference between the magnitudes of the first peaks is 7.8%, while the magnitudes of the second peaks differ by 16.6%. The time difference between the first peaks is insignificant, whereas the time difference between the second peaks was 26.3%. Therefore the residual torque had a significant influence on the impact force history.

Case (iii) “Power-Off without residual torque”
A third set of experimental tests was conducted by disengaging the residual torque to eliminate its effect on the impact force history. As expected, only one impact occurred since there was no residual torque to drive the crane back into the end stops. The first impact force was reduced by 20.7% when compared to the corresponding peak of case (i): “Power-Off with residual torque”.

All further experimental tests were performed using case (iii): “Power-Off without residual torque”, since this case yielded the expected impact force history, i.e. only one impact peak during the collision between the crane and the end stops of the crane supporting structure.

RESULTS

Calibration of the FE model to the experimental impact force history (no hoist load)
Figure 7 shows the end buffer impact force history of test case (iii) with the time reduced to 0.5 s. The FE simulations were conducted in the same way as for the experimental tests. When the original damping characteristics were used in the FE model, it resulted in a slight discrepancy in the impact force histories between the experimental and FE results. An improved FE impact force history was obtained by adjusting the buffer’s damping characteristics by less than 5%. Superimposed on Figure 7 is the FE end buffer impact force history. The impact forces and occurrence of the peaks varied by less than 3%, proving that a good correlation was achieved between the experimental and FE impact force histories.

Comparison of the experimental and FE impact force histories

Condition: Hoist load raised 0.15 m above ground level
Except for the addition of the hoist load, the same experimental and FE models were used as for the condition “No hoist load” to obtain the impact force histories when the hoist load was lifted 0.15 m above ground level.
The hoist load was symmetrically positioned on the crane bridge and lifted 0.15 m above ground level for both the experimental tests and the FE simulations. Figure 8 shows the superimposed experimental and FE impact force histories for this case.

From the experimental tests it was observed that, after the first impact, the buffers lost contact with the end stops for 0.42 s before impacting the end stops for two consecutive collisions. The second and third impacts were due to the hoist load’s inertia during its pendulum motion as the hoist load yanked the crane into the end stops. The secondary impacts occurred as expected. The experimental history resulted in three impacts which occurred at 0.17 s, 1.02 s and 1.48 s with magnitudes of 6.68 kN, 4.43 kN and 2.61 kN respectively.

The FE simulations followed the same trend as the experimental history, but with some discrepancies. The FE simulations resulted in three impacts which occurred at 0.15 s, 0.88 s and 1.48 s with magnitudes of 6.35 kN, 4.43 kN and 2.61 kN respectively.

The differences in the magnitudes of the first and second impact peaks between the experimental and FE histories were 4.9% and 0.9%, while the differences in time were 11.7% and 13.7% respectively. A negative shift of 0.14 s occurred between the second peaks of the experimental test and the FE simulation. The reasons for the discrepancies were not obvious and required further investigation. Additional impact tests were conducted with the hoist load raised to 2.20 m above ground level to determine the differences between the experimental and FE impact histories.

**Condition: Hoist load raised 2.20 m above ground level**

The hoist load was raised to 2.20 m above ground level, instead of 0.15 m as in the previous case. Figure 9 shows the superimposed experimental and FE impact force histories for this case.

In the experimental impact tests, the buffers did not lose contact with the end stops for the entire duration of the tests. The experimental histories resulted in three impacts which occurred at 0.16 s, 0.73 s and 1.04 s with magnitudes of 7.08 kN, 2.74 kN and 3.89 kN respectively. Surprisingly, the second impact was smaller than the third impact. A possible reason for this is the cancellation of various modes during impact.

The FE simulation predicted the first impact reasonably accurately, but thereafter the FE simulation results deviated substantially from the experimental test history. The numerical impacts occurred at 0.16s, 0.75 s and 1.30 s with magnitudes of 6.59 kN, 6.88 kN and 2.37 kN respectively. In the FE simulations the buffers lost contact with the end stops for 0.12 s after the first impact.

**DISCUSSION**

**Possible reasons for the discrepancies between the experimental and FE impact force histories**

After careful observation of the video footage of the experimental tests and the FE simulations of the impact force histories, it was discovered that certain parameters had a significant influence on the end buffer impact force history. A slight change in the magnitude of the parameters can lead to significantly different impact force histories. Table 2 lists the parameters which were identified as having a significant influence on the end buffer impact force history, and gives a comparison of the discrepancies between the FE and the experimental test parameters.

The discrepancies between the parameters of the experimental tests and the FE models led to the surmise that this could be the reason(s) for the differences in the impact force histories when the hoist load is included in the analysis. FE simulations were

**Table 2 Comparison of the FE and experimental parameters discrepancies**

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Experimental Configuration</th>
<th>FE Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal hoist load lag with respect to the crane bridge at the moment of impact</td>
<td>Not measured (difficult to measure)</td>
<td>Known lag angle = 0° Can be varied or kept constant</td>
</tr>
<tr>
<td>Crab position on the crane bridge</td>
<td>At midspan</td>
<td>At midspan</td>
</tr>
<tr>
<td>End stop misalignment in relation to each other</td>
<td>No measurable misalignment</td>
<td>No misalignment</td>
</tr>
<tr>
<td>Crane supporting structure's flexibility during impact</td>
<td>Not measured</td>
<td>Known from FE simulations Exact</td>
</tr>
<tr>
<td>Crane velocity at the moment of impact</td>
<td>0.55 m/s ±</td>
<td>0.55 m/s Exact</td>
</tr>
<tr>
<td>Elastic characteristics of buffer</td>
<td>Similar up to 5 kN at 35 mm compression, then differs significantly</td>
<td>Identical for both buffers</td>
</tr>
<tr>
<td>Damping characteristics of buffer</td>
<td>Slight discrepancy in buffers</td>
<td>Identical for both buffers</td>
</tr>
<tr>
<td>Release of the longitudinal acceleration button at the moment of impact</td>
<td>Done by operator, can vary significantly Not exact</td>
<td>Released at the moment of impact Exact</td>
</tr>
</tbody>
</table>

![Figure 9 Comparison of the FE and experimental end buffer impact force histories with hoist load hoisted 2.20 m above ground level for “Power-Off without residual torque”](image-url)
conducted using the parameters identified in Table 2, together with carefully chosen parameter variations which were observed in the experimental and FE simulations. The initial investigation was conducted by individually varying parameters in the FE model. Figure 10 shows arbitrarily selected impact force histories obtained when the magnitude of a single parameter was varied at a time with the remaining parameters kept constant.

Figure 10 shows the large variation in the first and second impact force peaks and the position of the peaks when the individual parameters are varied. This indicates that a change in the magnitude of the parameters does indeed have a significant influence on the impact force histories, suggesting that the differences between the experimental and FE impact force histories can be attributed to the contribution of the individual parameters when the hoist load is lifted. The range of variation of the parameters in the study was based on the parameter variability which was obtained from observation of the experimental and video footage. A detailed list of parameter variations is given in the paper following on this one, i.e. paper 752-B on page 63, titled Estimation of the maximum end buffer impact force for a given level of reliability.

### SUMMARY AND CONCLUSIONS

The FE impact force histories show that, by adjusting the magnitude of the parameters individually, the impact force histories are significantly affected compared to the base impact force history. An improved match between the experimental and FE histories with the hoist load attached could be achieved through extensive parameter fitting of the FE model. This was abandoned as irrelevant since the magnitudes of the parameters obtained for the improved match would be for a particular situation only, e.g. the magnitudes of the parameters would differ when the position of the hoist load was altered.

The codes of practice for the design of structures yield very different results, as different analysis philosophies and factors are taken into account for the estimation of the end buffer impact force. These approaches are based on a decoupled analysis of the crane and the crane supporting structure. This results in important parameters being omitted in the estimation of the end buffer impact force, which can lead to substantially under- or overestimated end buffer impact forces.

Although the particular combinations of the parameters and their magnitudes which caused the discrepancies between the FE and experimental histories were not found, it is believed that the sensitivity to the identified parameters (mechanisms) indicate that there is significant room for error in the codified end buffer impact forces.

Evidence was provided that the parameters omitted by the codes of practice do indeed have a significant influence on the end buffer impact force history. Crane failures could easily occur if these parameters are not properly accounted for. The effect of varying the magnitudes of the parameters in the FE simulations was investigated in detail through a sensitivity study and is presented in the paper following on this one, i.e. paper 752-B on page 63, titled Estimation of the maximum end buffer impact force for a given level of reliability.

### BIBLIOGRAPHY

ABAQUS. Personal communication and www.abaqus.com


DEMAG. Cranes and components, buffers. DPZ 100 cellular plastic buffers, DPG rubber buffers, DPH hydraulic buffers.

European Committee for Standardisation 1991. EN 1991-3:2003, EUROCODE 1 – Actions on structures,
Part 3: Actions induced by cranes and machinery, CEN/TC250/SCI, Clause 2.11.1, pp 1–44.


The first paper in this set of two, titled *The effect of parameters on the end buffer impact force history of the crane* (see page 55), examined the effect of a change in the magnitude of the parameter on the end buffer impact force history. This paper investigates to what degree a change in the magnitude of the parameter alters the impact force history. This was accomplished through a sensitivity analysis performed by individually varying the magnitude of the parameter in the FE model. For each case individual maximum impact forces were obtained. The maximum impact force could not simply be selected by choosing the greatest value from the sensitivity study. A constraint optimisation technique for a given level of reliability ($\beta$) using the FE simulation data was used to determine the maximum impact force. A comparison between the constraint optimisation and codified results showed that SABS 0160-1999 underestimates the impact force by 18%, while SANS 10160-2010 substantially overestimates the impact force by 64% for a level of reliability of $\beta = 3$. If the relevant clauses of SANS 10160-6 that pertain to end stop design are used in their present form, this will result in a conservative design, whereas SABS 0160 has a probability of 2.3% of being exceeded.

### INTRODUCTION

Underestimation of the end buffer impact forces as a result of a collision between the crane and the supporting structure can lead to disastrous consequences. This could result in the crane running off the rails during impact if the end stops fail. Although the cost of increasing the end stop connections is minimal compared to the overall cost of the structure, the cost of failure if the crane ran off the crane rails would be significant and could lead to fatalities. Some structural engineering professionals who were consulted increased the impact force because they are uncertain whether the codified estimations would prevent a major catastrophe. The guidelines and design codes considered in this study are:

- Manufacturer’s guidelines: DEMAG
- Eurocode 1, Part 3, EN 1991
- South African National Standard: SANS 10160 – Part 6
- Australian Standard, AS 1418.1 – 1994
- Association of Steel and Iron Engineer’s technical report, AISE No 13 – 1997

The design codes of practice use various approaches to estimate the impact force as described in the accompanying paper on page 55. Table 1 of the accompanying paper shows the limited number of parameters which the design codes take into account to estimate the impact forces. These approaches are followed to simplify the calculations. Also, all the design codes consider the crane and the supporting structure as a decoupled system to estimate the impact force. This can lead to significant errors in the estimation of the impact force.

In the accompanying paper, evidence was provided that the parameters do have an effect on the impact force histories. This paper describes a sensitivity study conducted to determine the influence of individual parameters on the end buffer impact force history. From this information the maximum impact force was determined for a given level of reliability using a constraint optimisation technique.

This paper also determined whether the design codes yield reasonable impact force estimates when compared to the constraint optimisation results for a given level of reliability ($\beta$). The results of this study provide a tool which structural engineering professionals can use to assess the codified end buffer impact force results.

Several papers have been published on the control mechanism to prevent the hoist load from oscillating during longitudinal travel. However, apart from the design codes, no papers were found in which the impact force is directly estimated when the crane collides with the end stops.

The sections below examine the methods used in the sensitivity study – only the
The horizontal lag of the hoist load is reviewed and discussed; the maximum end buffer impact force is estimated, which includes the probability of the parameters, the design point and the probability of exceedance, and the results of the constraint optimisation technique are given. The paper ends with a conclusions section.

**METHODS USED IN THE SENSITIVITY STUDY**

The impact force histories shown in Figure 9 of the accompanying paper were obtained without a detailed sensitivity analysis. They were obtained by simply choosing a reasonable variation of the magnitude of the parameter for the FE simulations. In the present paper, the FE model described in the accompanying paper and the variation of the magnitude of the parameters were used to conduct a detailed sensitivity analysis. The range of variation of the parameters was obtained by carefully examining the video footage of the experimental tests and the FE simulations. Table 1 shows the parameters with their corresponding base state, range of variation and interval of variation.

The impact force history was obtained by varying the magnitude of a single parameter while keeping the remaining parameters constant. This approach allowed the impact force history of the individual parameter’s mode of vibration to be obtained, i.e. the response of only one parameter on the impact force history. Besides adjusting the magnitude of the parameters, FE simulations were also conducted for the following cases:

1. “Power-Off hoist load bottom”, i.e. the impact occurred as a result of the crane’s inertia when the hoist load was raised 0.15 m above ground level.
2. “Power-On hoist load bottom”, i.e. during impact the longitudinal motors were constantly engaged with the hoist load raised 0.15 m above ground level.
3. “Power-Off hoist load top”, i.e. the impact occurred as a result of the crane’s inertia when the hoist load was raised 2.20 m above ground level.
4. “Power-On hoist load top”, i.e. during impact the longitudinal motors were constantly engaged with the hoist load raised 2.20 m above ground level.

Due to limited space and to prevent repetition, only one parameter, i.e. the horizontal lag angle of the hoist load, is discussed in detail.

### Table 1 Parameters identified for the FE sensitivity analysis which could have a significant effect on the impact history

<table>
<thead>
<tr>
<th>Parameter (Variable)</th>
<th>Base Value</th>
<th>Range of Variation</th>
<th>Interval of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lag of the centre of gravity (COG) of the hoist load with respect to the crane bridge</td>
<td>0°</td>
<td>2.50° ±</td>
<td>1.25° ± 2.50° ±</td>
</tr>
<tr>
<td>Crab and hoist load eccentricity on the crane bridge</td>
<td>0 m</td>
<td>3.39 m ±</td>
<td>1.695 m ± 3.390 m ±</td>
</tr>
<tr>
<td>End stop misalignment</td>
<td>0</td>
<td>150 mm</td>
<td>25 mm 50 mm 150 mm</td>
</tr>
<tr>
<td>Flexibility of the crane supporting structure</td>
<td>Rigid</td>
<td>Weak, intermediate and strong spring</td>
<td>Spring stiffness varied</td>
</tr>
<tr>
<td>Crane velocity on impact</td>
<td>0.55 m/s</td>
<td>– 0.165 m/s</td>
<td>0.05 m/s ± – 0.165 m/s</td>
</tr>
<tr>
<td>Elastic characteristics of buffer</td>
<td>Stiffness curve used in FEA</td>
<td>20% ±</td>
<td>10% ± 20% ±</td>
</tr>
<tr>
<td>Damping characteristics of buffer</td>
<td>Damping used in FEA</td>
<td>Without damping</td>
<td>Without damping</td>
</tr>
</tbody>
</table>

**Figure 1 Impact force: hoist load bottom with “Power-Off” for the hoist load lag**

**REVIEW OF PARAMETER: HORIZONTAL LAG OF THE HOIST LOAD**

**Impact force history:**  
Parameter = horizontal lag of the hoist load

This parameter was investigated as all the codes of practice, except for SANS 10160-6 and EN 1991:3–2003, ignore the effect of the hoist load if it is not rigidly restrained (fixed) to the crane bridge. To study the horizontal lag effect of the hoist load on the impact force history, the cable and hoist load were inclined at angles of 1.25° ± and 2.50° ± from the vertical at the moment of impact. A positive lag is defined as the hoist load ahead of the crane bridge at the moment of impact.

**Results of the sensitivity study of the horizontal lag of the hoist load**

The effect of the hoist load lag on the impact force history is shown in Figures 1 and 2 when the hoist load is raised 0.15 m and 2.20 m above ground level for the “Power-Off” conditions.

**Sensitivity study of the horizontal lag of the hoist load**

The following information was extracted from Figures 1 and 2 for the horizontal lag of the hoist load:

**Case: hoist load bottom**

A positive increase in the lag angle resulted in a substantial increase in the magnitude of the first impact force, while the magnitude of the second impact force was only marginally affected.
The opposite occurred for a negative lag angle, except that the second impact force increased proportionately as the negative lag angle increased.

The position of the first impact peak was insignificantly affected, while a significant positive shift of the second peak was observed for a negative lag angle, and a significant negative shift was observed for a positive lag angle.

**Case: hoist load top**
The impact force history for the hoist load top case follows a similar trend as for the hoist load bottom case, except that the magnitudes and position of the second impact force were insignificantly affected.

**SUMMARY OF THE FE SIMULATIONS**
The sensitivity study of the remaining parameters showed similar trends. Refer to Haas (2007) for a complete review of the effect of a change in magnitude of the remaining parameters.

Table 2 presents the significant information that was extracted from the FE simulations when the peak forces were compared to the base states for six of the seven parameters listed in Table 1. The remaining parameter, the elastic characteristics of the buffer, was disregarded due to its insignificant effect on the end buffer impact force histories. It is important to note that, although the impact histories are not significantly affected, the displacement histories show a moderate change. When the magnitude of the parameter was varied, it could yield either a positive or negative change in the first and second impact peaks, as well as a position shift of the impacts. This is clearly illustrated in Figures 1 and 2 for a variation of the lag angle of the hoist load. From Table 2, the maximum percentage positive increase for the first peak when the hoist load was raised 0.15 m and 2.20 m above ground level was 38% and 37% respectively. For the second peak, the maximum percentage positive increase was 211% and 57%. The maximum time difference between the peaks was 32% and 34% respectively when the hoist load was raised 0.15 m and 2.20 m above ground level.

**Impact force histories for arbitrarily selected parameters**
Figure 3 shows the impact force histories of arbitrarily selected simulations for six of the

---

**Table 2 Summary of significant information obtained from the FE simulations**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hoist Load Position for “Power-Off” and “Power-On”</th>
<th>First Peak Magnitude</th>
<th>Second Peak Magnitude</th>
<th>Time Between Peaks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Maximum % Positive Increase</td>
<td>Maximum % Negative Decrease</td>
<td>Maximum % Positive Increase</td>
</tr>
<tr>
<td>Hoist Load Lag</td>
<td>Bottom</td>
<td>+ 38</td>
<td>- 26</td>
<td>+ 32</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 33</td>
<td>- 25</td>
<td>+ 7</td>
</tr>
<tr>
<td>Hoist Load and Crab Eccentricity</td>
<td>Bottom</td>
<td>+ 22</td>
<td>N/A</td>
<td>+ 31</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 26</td>
<td>N/A</td>
<td>+ 18</td>
</tr>
<tr>
<td>Flexibility of the Crane Supporting Structure</td>
<td>Bottom</td>
<td>+ 5</td>
<td>- 31</td>
<td>+ 49</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>0</td>
<td>- 34</td>
<td>+ 14</td>
</tr>
<tr>
<td>Impact Velocity of the Crane</td>
<td>Bottom</td>
<td>+ 24</td>
<td>- 46</td>
<td>+ 53</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 25</td>
<td>- 51</td>
<td>+ 28</td>
</tr>
<tr>
<td>End Stop Misalignment</td>
<td>Bottom</td>
<td>+ 33</td>
<td>N/A</td>
<td>+ 65</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 37</td>
<td>N/A</td>
<td>+ 37</td>
</tr>
<tr>
<td>Damping Characteristics of the Buffer</td>
<td>Bottom</td>
<td>+ 20</td>
<td>N/A</td>
<td>+ 211</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 20</td>
<td>N/A</td>
<td>+ 57</td>
</tr>
<tr>
<td>Maximum Percentage Difference</td>
<td>Bottom</td>
<td>+ 38</td>
<td>- 46</td>
<td>+ 211</td>
</tr>
<tr>
<td></td>
<td>Top</td>
<td>+ 37</td>
<td>- 51</td>
<td>+ 57</td>
</tr>
</tbody>
</table>

Note: N/A means that the first and second impact forces were greater than the base state, and thus no impact forces lower than the base state were obtained.
seven parameters investigated when the hoist load was raised to 0.15 m above ground level.

The results from Figure 3 confirm that the individual parameters do have a substantial influence on the impact force histories in terms of magnitude and position. Improved agreement with the experimental impact force histories could be obtained by adjusting the magnitude of the parameters. However, the magnitude of the adjusted parameters will only be valid for the specific case, as the impact force history is very sensitive to the variation of the individual parameters.

**ESTIMATION OF THE MAXIMUM END BUFFER IMPACT FORCE**

The end buffers must be designed to have some arbitrarily chosen, low probability of failure if an impact occurs. Thus the question arises as to what impact force the end buffers must be designed to withstand. A more convenient way to address the same question is to ask: for a given end buffer capacity ($f_o$), what is the probability of failure under impact?

**Linear load model**

The FE analysis provided information on the effect of various parameters on the impact force. Since only one parameter was varied at a time and only in one increment, only the gradient of the impact force could be assessed, which led to the choice of a linear model. Clearly this assumption of linearity is a weak link in the present work. Reinforcing the link would require a much wider set of FE analyses to be carried out.

The linear model is of the form:

$$f(\Delta P) = f(0) + \sum_{i=1}^{n} \frac{\partial f}{\partial P_i} \cdot \Delta P_i$$

$$= f(0) + \nabla f^T \cdot \Delta P$$

(1)

where:

$f(\Delta P)$ is the end buffer impact force,

$\Delta P = P - P_o$ is the change in the parameters

where $P_o$ is the nominal value of the parameters (at which the gradient was assessed),

$n$ is the number of parameters.

The changes in force $\left( \frac{\partial f}{\partial P_i} \cdot \Delta P_i \right)$ for each parameter for all four cases studied using

---

![Figure 3 Selected impact force response of each parameter compared to base response when hoist load is raised to 0.15 m above ground level](image)

**Table 3** Change in force per parameter when the impact forces are 3σ from the base value for the first impact response

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hoist Load Bottom “Power-Off”</th>
<th>Hoist Load Bottom “Power-On”</th>
<th>Hoist Load Top “Power-Off”</th>
<th>Hoist Load Top “Power-On”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base Impact Force ($f_o$)</strong></td>
<td>6.35</td>
<td>7.26</td>
<td>6.65</td>
<td>7.48</td>
</tr>
<tr>
<td>1. Lag Angle</td>
<td>3.17</td>
<td>3.69</td>
<td>2.50</td>
<td>3.56</td>
</tr>
<tr>
<td>2. Crab and Hoist Load Eccentricity</td>
<td>1.08</td>
<td>1.52</td>
<td>1.53</td>
<td>2.03</td>
</tr>
<tr>
<td>3. Flexibility of Crane Supporting Structure</td>
<td>−2.66</td>
<td>−3.06</td>
<td>−2.63</td>
<td>−1.52</td>
</tr>
<tr>
<td>4. Crane Impact Speed</td>
<td>4.13</td>
<td>4.73</td>
<td>4.43</td>
<td>4.88</td>
</tr>
<tr>
<td>5. One End Stop Misaligned</td>
<td>3.69</td>
<td>4.17</td>
<td>4.85</td>
<td>5.19</td>
</tr>
<tr>
<td>6. Damping Characteristics of Buffer</td>
<td>1.13</td>
<td>1.31</td>
<td>1.21</td>
<td>1.38</td>
</tr>
</tbody>
</table>
FE, i.e. Hoist load bottom “Power-Off”, Hoist load bottom “Power-On”, Hoist load top “Power-Off” and Hoist load top “Power-On” for ∆P = 3σ (a change in parameter of three standard deviations), are presented in Table 3 for the first impact and in Table 4 for the second impact.

**Probability distribution of the parameters**

A probability density can be associated to any value of $\Delta P$. Since only information on standard deviation is available, a reasonable model to use was a multinomial Gaussian distribution:

$$p(\Delta P) = (2\pi)^{-n/2} \det(C)^{1/2} \exp\left(-\frac{1}{2} \Delta P^T \cdot C^{-1} \cdot \Delta P\right)$$

(2)

where:

$C$ is the covariance matrix.

Since no cross-correlation information was available, $C$ was taken as diagonal, with the square of the deviation of each parameter on the diagonal. The standard deviations of each parameter presented in Table 5 were obtained from engineering judgement and a review of video footage of the experimental tests and FE simulations.

**Design point**

Finding the combination of parameters leading to a given load with the highest value of $p(\Delta P)$ is equivalent to finding the combination of parameters leading to the same load, with the lowest value of

$$g(\Delta P) = -\frac{1}{2} \Delta P^T \cdot C^{-1} \cdot \Delta P$$

Hence this leads to Equation 4 which must be solved.

**Find $\Delta P$ that minimises**

$$g(\Delta P) = -\frac{1}{2} \Delta P^T \cdot C^{-1} \cdot \Delta P$$

(3)

under the constraint

$$f_c = f(0) + (\nabla f)^T \cdot \Delta P$$

This is a constrained minimisation problem. One convenient way to solve this is to transform Equation 4 into an unconstrained minimisation problem by means of Lagrange multipliers which can show that the above problem is equivalent (Larson 1995) to solving

**Find $\Delta P$ and $\lambda$ for which**

$$g(\Delta P) = -\frac{1}{2} \Delta P^T \cdot C^{-1} \cdot \Delta P + \lambda(\nabla f)^T \cdot \Delta P + f(0) - f_c$$

is extremal

(5)

This again can be shown that it amounts to solving the linear system of equations:

$$\begin{bmatrix} 0 & (\nabla f)^T \cdot \lambda \end{bmatrix} = \begin{bmatrix} f_c - f(0) \end{bmatrix}$$

(6)

The value $\Delta P$ thus found is the most probable combination of parameters that cause an end buffer impact force equal to $f_c$. This value of $\Delta P$ is known in the theory of first order reliability methods (FORM) as a design point (Ang 1990).

**Probability of exceedance**

FORM provides another important result. The *reliability index* $\beta$ is defined by

$$\beta = -\sqrt{\Delta P^T \cdot C^{-1} \cdot \Delta P}$$

(7)

It can then be shown that the probability that the end buffer impact force exceeds $f_c$ is equal to:

$$p(f > f_c) = \Phi(-\beta)$$

(8)

where:

$\Phi$ is the Gaussian cumulative distribution.

**Results of the constraint optimisation technique**

The solution of the constrained optimisation program for various levels of reliability is presented in Tables 6 and 7 for the “Power-Off” and “Power-On” conditions respectively.

### Table 4 Change in force per parameter when the impact forces are 3σ from the base value for the second impact response

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hoist Load Bottom “Power-Off”</th>
<th>Hoist Load Bottom “Power-On”</th>
<th>Hoist Load Top “Power-Off”</th>
<th>Hoist Load Top “Power-On”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Force (kN)</td>
<td>Change in Force (kN)</td>
<td>Change in Force (kN)</td>
<td>Change in Force (kN)</td>
<td></td>
</tr>
<tr>
<td>Base Impact Force ($f_c$)</td>
<td>4.43</td>
<td>4.61</td>
<td>6.88</td>
<td>8.05</td>
</tr>
<tr>
<td>1. Lag Angle</td>
<td>-1.19</td>
<td>-0.96</td>
<td>0.38</td>
<td>1.09</td>
</tr>
<tr>
<td>2. Crab and Hoist Load Eccentricity</td>
<td>0.72</td>
<td>1.43</td>
<td>1.73</td>
<td>1.28</td>
</tr>
<tr>
<td>3. Flexibility of Crane Supporting Structure</td>
<td>-1.48</td>
<td>-1.61</td>
<td>-2.96</td>
<td>-3.04</td>
</tr>
<tr>
<td>4. Crane Impact Speed</td>
<td>4.16</td>
<td>5.03</td>
<td>4.43</td>
<td>6.26</td>
</tr>
<tr>
<td>5. One End Stop Misaligned</td>
<td>2.46</td>
<td>4.75</td>
<td>3.39</td>
<td>2.58</td>
</tr>
<tr>
<td>6. Damping Characteristics of Buffer</td>
<td>7.74</td>
<td>8.75</td>
<td>3.50</td>
<td>3.99</td>
</tr>
</tbody>
</table>

### Table 5 Estimated standard deviation for each parameter

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated Standard Deviation (σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Lag Angle</td>
<td>0.022 Radians (1.25°)</td>
</tr>
<tr>
<td>2. Crab and Hoist Load Eccentricity</td>
<td>1.13 m</td>
</tr>
<tr>
<td>3. Flexibility of Crane Supporting Structure</td>
<td>0.0025 m (2.5 mm)</td>
</tr>
<tr>
<td>4. Crane Impact Velocity</td>
<td>0.05 m/s</td>
</tr>
<tr>
<td>5. One End Stop Misaligned</td>
<td>0.04125 m (41.25 mm)</td>
</tr>
<tr>
<td>6. Damping Characteristics of Buffer</td>
<td>20%</td>
</tr>
<tr>
<td>7. Elastic Characteristics of Buffer</td>
<td>30%</td>
</tr>
</tbody>
</table>

### Table 6 Estimated maximum end buffer impact force from the first impact response

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Hoist Load Bottom “Power-Off”</th>
<th>Hoist Load Bottom “Power-On”</th>
<th>Hoist Load Top “Power-Off”</th>
<th>Hoist Load Top “Power-On”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated maximum end buffer impact force for $\beta = 1$</td>
<td>7.64</td>
<td>9.05</td>
<td>8.44</td>
<td>9.83</td>
</tr>
<tr>
<td>Estimated maximum end buffer impact force for $\beta = 2$</td>
<td>8.93</td>
<td>10.83</td>
<td>10.23</td>
<td>12.19</td>
</tr>
<tr>
<td>Estimated maximum end buffer impact force for $\beta = 3$</td>
<td>10.22</td>
<td>12.62</td>
<td>12.03</td>
<td>14.54</td>
</tr>
</tbody>
</table>
The maximum end buffer impact force of 14.54 kN occurred for the condition “Hoist load top with Power-On” for the particular crane and crane supporting structure investigated, for a reliability index of $\beta = 3$.

The probability of exceedance is related to the reliability indices calculated using Equation 9 and is given for various reliability indices in Table 8:

$$ P = \Phi(-\beta) \quad (9) $$

Figure 4 presents a comparison of the various codified impact forces with the maximum estimated end buffer impact force for $\beta = 1, 2$ and $3$. From Figure 4 it can be concluded that SABS 0160 underestimates the end buffer impact force by 18%, while SANS 10160-6 overestimates it by 64% for a target reliability index of $\beta = 3$.

It can also be concluded that SABS 0160 corresponds to $\beta = 2$. The code therefore yields an impact force which has a probability of $2.3 \times 10^{-2}$ ($2.3\%$) of being exceeded.

**CONCLUSIONS**

End buffer impact forces are the result of complex behaviour of the structure during an impact, and this behaviour is influenced by a series of parameters. Failure to adequately address these effects can lead to a catastrophe. An estimation of existing forces shows that, except for EN 1991-3 and SANS 10160-6, all other design codes result in a reliability index ($\beta$) lower than 3 as calculated in this paper using constraint optimisation. It is generally accepted that a reliability index of 3 should be used for design purposes. Thus the design codes that yield estimates lower than 3 do not meet international standards.

At this stage it is not possible to make a general recommendation as to the most important parameters, as only one impact velocity was considered. However, the present work clearly highlights the need for a revision of the code requirements. This would require the FE simulations to be repeated for various impact velocities, different masses and different crane configurations.

**BIBLIOGRAPHY**

ABAQUS. Personal communication and www.abaqus.com


DEMAG. Personal communication and www.demag.com


Assessment of the behaviour factor for the seismic design of reinforced concrete structural walls according to SANS 10160 – Part 4

R C le Roux, J A Wium

Reinforced concrete structures, designed according to proper capacity design guidelines, can deform inelastically without loss of strength. Therefore, such structures need not be designed for full elastic seismic demand, but could be designed for a reduced demand. In codified design procedures this reduced demand is obtained by dividing the full elastic seismic demand by a code-defined behaviour factor. There is, however, no consensus in the international community regarding the appropriate value to be assigned to the behaviour factor. The purpose of this study is to assess the value of the behaviour factor currently prescribed by SANS 10160-4 (2011) for the design of reinforced concrete structural walls. This is done by comparing displacement demand to displacement capacity for a series of structural walls. The first step in seismic force-based design is the estimation of the fundamental period of the structure. The influence of this first crucial step is investigated in this study by considering two period calculation methods. It was found that, regardless of the period calculation method, the current behaviour factor value prescribed in SANS 10160-4 (2011) is adequate to ensure that inter-storey drift of structural walls would not exceed code-defined drift limits.

INTRODUCTION
In the 1960s, with the development of inelastic time history analysis (ITHA), came the realisation that well designed structures can deform inelastically without loss of strength (Priestley et al 2007: 1–4). Engineers realised that structures need not be designed for the full elastic seismic demand (seismic load), but could be designed for a reduced demand. This reduced demand is obtained by dividing the full elastic seismic demand by a code-defined behaviour factor. There is, however, no consensus in the international community regarding the appropriate value to be assigned to the behaviour factor. This is evident in the wide range of behaviour factor values specified by international design codes (see Table 1). (These behaviour factor values should, however, not be directly compared, since various other code-related

Table 1 Examples of maximum force-reduction factors for the damage control limit state in different countries (Priestley et al 2007: 13)

<table>
<thead>
<tr>
<th>Structural type and material</th>
<th>US West Coast</th>
<th>Japan</th>
<th>New Zealand**</th>
<th>Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concrete frame</td>
<td>8</td>
<td>–</td>
<td>9</td>
<td>5.85</td>
</tr>
<tr>
<td>Conc. struct. wall</td>
<td>5</td>
<td>1.8–3.3</td>
<td>7.5</td>
<td>4.4</td>
</tr>
<tr>
<td>Steel frame</td>
<td>8</td>
<td>2.0–4.0</td>
<td>9</td>
<td>6.3</td>
</tr>
<tr>
<td>Steel EBF*</td>
<td>8</td>
<td>2.0–4.0</td>
<td>9</td>
<td>6.0</td>
</tr>
<tr>
<td>Masonry walls</td>
<td>3.5</td>
<td>–</td>
<td>6</td>
<td>3.0</td>
</tr>
<tr>
<td>Timber (struct. wall)</td>
<td>–</td>
<td>2.0–4.0</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>Prestressed wall</td>
<td>1.5</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Dual wall/frame</td>
<td>8</td>
<td>1.8–3.3</td>
<td>6</td>
<td>5.85</td>
</tr>
<tr>
<td>Bridges</td>
<td>3–4</td>
<td>–</td>
<td>6</td>
<td>3.5</td>
</tr>
</tbody>
</table>

* Eccentrically braced frame ** S factor of 0.67 incorporated

Keywords: seismic design, behaviour factor, reinforced concrete, structural wall, inter-storey drift
requirements also vary between international codes. Thus, each behaviour factor should be viewed from within the context of the corresponding code).

The purpose of this paper is to assess the current value of the behaviour factor in SANS 10160-4 (2011) for the seismic design of reinforced concrete structural walls. A value of 5 is specified in this standard.

Additionally, this paper evaluates the way in which the fundamental period of a structure is determined. Seismic design codes, including SANS 10160-4 (2011), provide a simple equation by which the fundamental period of a structure may be calculated, subject to certain limitations. It is well known that this equation results in seismic design forces to be overestimated, and lateral displacement demand to be underestimated (Priestley et al. 2007: 11). An alternative period calculation procedure, based on moment-curvature analysis, will also be assessed. This method provides a more realistic estimate of the fundamental period of structures, but due to its iterative nature it is not often applied in design practice.

The influence of the behaviour factor becomes evident in seismic displacement demand. Therefore, in order to assess the current behaviour factor value, a comparison is required between seismic displacement demand and displacement capacity. A series of independent structural walls are assessed in this investigation. A first estimate of displacement demand of these walls is obtained from the equal displacement and equal energy principles. The displacement demand is then verified by means of a series of ITHA applied to these walls. Displacement capacity is defined by seismic design codes in terms of inter-storey drift limits to prevent non-structural damage in building structures. “Displacement capacity” could thus be described as “allowed displacement”.

DUCTILITY DEMAND AND CAPACITY

Displacement ductility is a measure of the magnitude of lateral displacement of a structure, where a displacement ductility of greater than one represents inelastic response. In the remainder of this paper the term ductility will be used with reference to displacement ductility. Both the displacement demand and displacement capacity will be expressed in terms of ductility for comparison purposes.

Ductility demand

The displacement calculation method prescribed by seismic design codes such as SANS 10160-4 (2011) is based on the equal displacement principle. However, the validity of the equal displacement principle has recently been questioned (Priestley et al. 2007: 26–29). Therefore, in this investigation ductility demand is calculated according to either the equal displacement or the equal energy principles (depending on the fundamental period), and then verified by means of ITHA.

Ductility capacity

Priestley et al. (2007: 71) states that it is difficult to avoid excessive non-structural damage when inter-storey drift levels exceed approximately 0.025, and hence it is common for building design codes to specify inter-storey drift limits of 0.02 to 0.025. At these levels, most buildings would not have reached the structural damage-control limit state.

Separating non-structural infill panels from the structural system by means of isolation joints forms part of good conceptual design practice (Bachmann 2003: 40). For such buildings EN 1998-1 (2004) specifies the following drift limit:

\[ d_r \leq 0.01h_s \]  

(1)
SANS 10160-4 (2011: 30) imposes the following drift limits:

\[ d_f = 0.025h_s \text{ if } T < 0.7 \text{ s} \]  
\[ d_f = 0.02h_s \text{ if } T > 0.7 \text{ s} \]  

where:
- \( d_f \) is the relative displacement between the top and bottom of a storey in the structure, obtained from a seismic event with a 10% in 50 year probability of occurrence
- \( h_s \) is the storey height
- \( \nu \) is a reduction factor which is equal to between 0.4 and 0.5, depending on the importance class of the structure.

SANS 10160-4 (2011: 30) imposes the following drift limits:

\[ \frac{\nu T}{H} \leq 0.02 \]  
\[ \frac{\nu T}{H} \leq 0.025 \]  

where:
- \( T \) is the fundamental period of the structure
- \( H \) is the height of the building, in metres, from the top of the foundation or rigid basement (see Figure 3).
- \( \nu \) is a reduction factor which is equal to between 0.4 and 0.5, depending on the importance class of the structure.

It may be seen that for a \( \nu \) value of 0.5, Eq 1 yields a drift limit of 0.02, which corresponds to the SANS drift limit for fundamental periods longer than 0.7 seconds. In this investigation ductility capacity is based on the period-dependent drift limits of Equations 2 and 3. The calculation of ductility capacity from these drift limits is discussed later.

**PARAMETER STUDY**

The following parameters are considered in this investigation:

- Period calculation method
- Wall aspect ratio
- Number of storeys

**Period calculation method**

**Method 1**

According to SANS 10160-4 (2011: 27) the fundamental period of a structure may be calculated using Eq 4:

\[ T_1 = C_T h_w \]  

where:
- \( C_T = 0.05 \) was assumed for this investigation (as per SANS 10160-4 (2011))
- \( h_w \) is the height of the building, in metres, from the top of the foundation or rigid basement (see Figure 3).

Equation 4 has been shown to correspond well to measured building periods (Priestley et al. 2007: 11). These measurements were, however, taken at very low levels of vibration (normally resulting from wind vibration), where non-structural participation is high and concrete sections are uncracked (Priestley et al. 2007: 11). Under seismic excitation, however, sections are allowed to crack and thus structures respond at much higher fundamental periods. It is often argued that using a too low period is conservative, since the acceleration demand is then overestimated (Priestley et al. 2007: 11). This, however, is not true, since an underestimation in period results in an underestimation of displacements (Dazio & Beyer 2009: 5-15).

Because Eq 4 underestimates the fundamental period, Dazio & Beyer (2009: 5-16) suggest that it “should never be used”. Eigenvalue analyses based on the stiffness derived from the cracked section should rather be used (Dazio & Beyer 2009: 5-16–18; Priestley et al. 2007: 11).

**Method 2**

As an alternative approach, the stiffness of a cracked reinforced concrete section can be obtained from a moment-curvature analysis of the section. This is done by drawing a bilinear approximation to the moment-curvature curve as shown in Figure 1 (Priestley et al. 2007: 144).

The fundamental period is then obtained from an eigenvalue analysis, assuming the same sectional stiffness, \( EI_{eff} \), over the height of the wall. The design of a wall, using this method, is unfortunately iterative, since the moment-curvature analysis cannot be done unless the reinforcement content and layout of the section is known, and the demand on the section depends on the stiffness of the section. For structures which comply with the requirements to allow for the use of the equivalent static force method, the iterative method depicted in Figure 2 should thus be followed.

**Wall aspect ratio**

The aspect ratio of the wall, defined as the height of the wall \( h_w \) divided by the length of the wall section \( l_w \) (see Figure 3), is another variable to be considered.

The aspect ratio determines the extent to which a wall responds in flexure or shear. A wall with an aspect ratio of less than three responds predominantly in shear (Paulay & Priestley 1992: 371). A structural wall subject to seismic action should preferably respond in ductile flexural action (Paulay & Priestley 1992: 362).

The aspect ratio should also not be too large. Priestley et al. (2007: 326) have shown that the elastic seismic force should not be reduced at all (behaviour factor \( \leq 1 \)) for walls with an aspect ratio of more than approximatively 9.

For the two above-mentioned reasons it was decided to consider walls with aspect ratios of 3, 5 and 8 in this study.

**Number of storeys**

This investigation focuses on the series of walls shown in Figure 4. The storey height was chosen as 3.23 m. The walls are all independent and free-standing. The behaviour of such a wall is, however, similar to that of a wall forming part of a symmetric structure.

Eq 4 is only applicable for buildings up to a height of 40 m. The 60 m wall is designed according to method 2 only.

The reason that the aspect ratio increases with height is that the wall section lengths need to remain within reasonable limits. The wall section lengths are shown in Table 2. It can be seen that only the shaded cells contain reasonable wall lengths.

Thus, the scope of this investigation is composed of the eight walls shown in Figure 4. These walls are designed according to both period calculation methods discussed earlier. Ground types 1 and 4 of SANS 10160-4 (2011) are used to define the range of seismic ground types. The methodology according to which seismic drift is assessed for these eight walls is presented next.

---

**Figure 3 Definition of wall dimensions**

**Table 2 Wall section lengths**

<table>
<thead>
<tr>
<th>Height [m]</th>
<th>Aspect ratio</th>
<th>3</th>
<th>5</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.230</td>
<td>1.080</td>
<td>0.640</td>
<td>0.400</td>
<td></td>
</tr>
<tr>
<td>6.460</td>
<td>2.160</td>
<td>1.300</td>
<td>0.800</td>
<td></td>
</tr>
<tr>
<td>9.690</td>
<td>3.240</td>
<td>1.940</td>
<td>1.220</td>
<td></td>
</tr>
<tr>
<td>19.380</td>
<td>6.460</td>
<td>3.880</td>
<td>2.420</td>
<td></td>
</tr>
<tr>
<td>38.760</td>
<td>12.920</td>
<td>7.760</td>
<td>4.840</td>
<td></td>
</tr>
<tr>
<td>58.140</td>
<td>19.380</td>
<td>11.620</td>
<td>7.260</td>
<td></td>
</tr>
</tbody>
</table>

The aspect ratio of the wall, defined as the height of the wall \( h_w \) divided by the length of the wall section \( l_w \) (see Figure 3), is another variable to be considered.
The methodology used in this investigation is illustrated in Figure 5 and is listed in steps 1 through 6 below. These steps are applied to each of the eight walls defined in Figure 4 for both ground types 1 and 4. Thus, the steps are applied sixteen times. Steps 1 to 3 describe the design of the walls, while steps 4 to 6 describe the assessment of the walls.

Two period calculation methods were previously introduced. The difference between these two methods will be evaluated by using both these period calculation methods in the design of the walls.

Different period calculation methods would produce different force demands on the structure. In practice, the mass of a structure is fixed, and thus different force demands will be reflected in the longitudinal reinforcement content of the structural wall, or the wall cross-sectional dimensions. For this study, however, the cross-sectional dimensions are fixed (for the purpose of comparison), and thus it was decided to use an “inverse” design method, where the capacity of the cross-section is fixed at the start (step 1) and the associated floor masses are obtained as the final result of the design (step 3).

The methodology steps are the following:

1. The width of the wall section \( b_w \) is chosen such that wall instability due to out-of-plane buckling in the plastic hinge region does not occur (Paulay & Priestley 1992: 403). An amount of reinforcement must be provided to comply with codified criteria. In this study the recommended reinforcement quantities of Dazio & Beyer (2009: 7–12) were used.

2. The moment capacity of the wall cross section at the base of the wall can be determined using either design equations or a moment curvature analysis. The moment capacity calculated using the design equations \( (M_n) \) corresponds to design material strengths. For analysis purposes it is important to predict the most likely response of the wall, thus the nominal yield moment \( (M_y) \) obtained from moment-curvature analysis corresponds to mean material strengths.

3. Given the chosen wall, the purpose of this step is to calculate floor masses \( m_1 \) and \( m_2 \) corresponding to the two period calculation methods respectively.

3.1 Method 1

3.1.1 The fundamental period \( (T_1) \) is calculated using Equation 4.

3.1.2 The design pseudo acceleration \( (a_1) \) is obtained from the design spectrum.

3.1.3 The floor mass \( m_1 \) should be of such a magnitude that the resulting base moment is slightly less than the nominal yield moment \( (M_y) \) obtained from the design equations. This is to take the additional strength, due to reinforcement choice, into consideration.

3.1.4 For analysis purposes a better estimate of the fundamental period at which the wall would respond \( (T_{1\text{real}}) \) is obtained by means of an eigenvalue analysis based on the cracked sectional stiffness obtained from the moment-curvature analysis.

3.2 Method 2

3.2.1 This step starts by assuming a value for the fundamental period \( (T_2) \). A good estimate is \( T_{1\text{real}} \) obtained in the previous step.

3.2.2 The design acceleration demand \( (a_2) \) is obtained from the design spectrum.

3.2.3 Similar to 3.1.3 above, the floor mass \( m_2 \) can be obtained.

3.2.4 A new estimate of \( T_2 \) is calculated using the eigenvalue analysis. Iteration, such as shown in Figure 2, is required until the value of \( m_2 \) does not change significantly between two iterations.

4. The purpose of this step is to estimate the ductility demand according to the equal displacement and equal energy principles. For this purpose the multi degree of freedom (MDOF) wall is converted into an equivalent single degree of freedom (SDOF) wall.

4.1 Firstly, the properties of the equivalent SDOF system need to be calculated. This includes the equivalent SDOF height \( h^* \) and the effective first modal masses \( m_1^* \) and \( m_2^* \). The equivalent height is obtained from Eq 12, while the effective first modal mass can be obtained from finite element modal analyses.

4.2 The shear \( (V_1) \) corresponding to nominal yield moment can be calculated from the nominal yield moment \( (M_y) \) obtained from moment-curvature analysis.

4.3 For both methods the acceleration \( (a_{1\text{real}}, a_{2\text{real}}) \) corresponding to the yield shear can be calculated.

4.4 The elastic acceleration demand \( (A_1 \) and \( A_2 \) can be obtained from the elastic pseudo acceleration spectrum.

**Figure 4 Structural wall range**

**METHODOLOGY**

The methodology used in this investigation is illustrated in Figure 5 and is listed in steps 1 through 6 below. These steps are applied to each of the eight walls defined in Figure 4 for both ground types 1 and 4. Thus, the steps are applied sixteen times. Steps 1 to 3 describe the design of the walls, while steps 4 to 6 describe the assessment of the walls.

Two period calculation methods were previously introduced. The difference between these two methods will be evaluated by using both these period calculation methods in the design of the walls.

Different period calculation methods would produce different force demands on the structure. In practice, the mass of a structure is fixed, and thus different force demands will be reflected in the longitudinal reinforcement content of the structural wall, or the wall cross-sectional dimensions. For this study, however, the cross-sectional dimensions are fixed (for the purpose of comparison), and thus it was decided to use an “inverse” design method, where the capacity of the cross-section is fixed at the start (step 1) and the associated floor masses are obtained as the final result of the design (step 3).

The methodology steps are the following:

1. The width of the wall section \( b_w \) is chosen such that wall instability due to out-of-plane buckling in the plastic hinge region does not occur (Paulay & Priestley 1992: 403). An amount of reinforcement must be provided to comply with codified criteria. In this study the recommended reinforcement quantities of Dazio & Beyer (2009: 7–12) were used.

2. The moment capacity of the wall cross section at the base of the wall can be determined using either design equations or a moment curvature analysis. The moment capacity calculated using the design equations \( (M_n) \) corresponds to design material strengths. For analysis purposes it is important to predict the most likely response of the wall, thus the nominal yield moment \( (M_y) \) obtained from moment-curvature analysis corresponds to mean material strengths.

3. Given the chosen wall, the purpose of this step is to calculate floor masses \( m_1 \) and \( m_2 \) corresponding to the two period calculation methods respectively.

3.1 Method 1

3.1.1 The fundamental period \( (T_1) \) is calculated using Equation 4.

3.1.2 The design pseudo acceleration \( (a_1) \) is obtained from the design spectrum.

3.1.3 The floor mass \( m_1 \) should be of such a magnitude that the resulting base moment is slightly less than the nominal yield moment \( (M_y) \) obtained from the design equations. This is to take the additional strength, due to reinforcement choice, into consideration.

3.1.4 For analysis purposes a better estimate of the fundamental period at which the wall would respond \( (T_{1\text{real}}) \) is obtained by means of an eigenvalue analysis based on the cracked sectional stiffness obtained from the moment-curvature analysis.

3.2 Method 2

3.2.1 This step starts by assuming a value for the fundamental period \( (T_2) \). A good estimate is \( T_{1\text{real}} \) obtained in the previous step.

3.2.2 The design acceleration demand \( (a_2) \) is obtained from the design spectrum.

3.2.3 Similar to 3.1.3 above, the floor mass \( m_2 \) can be obtained.

3.2.4 A new estimate of \( T_2 \) is calculated using the eigenvalue analysis. Iteration, such as shown in Figure 2, is required until the value of \( m_2 \) does not change significantly between two iterations.

4. The purpose of this step is to estimate the ductility demand according to the equal displacement and equal energy principles. For this purpose the multi degree of freedom (MDOF) wall is converted into an equivalent single degree of freedom (SDOF) wall.

4.1 Firstly, the properties of the equivalent SDOF system need to be calculated. This includes the equivalent SDOF height \( h^* \) and the effective first modal masses \( m_1^* \) and \( m_2^* \). The equivalent height is obtained from Eq 12, while the effective first modal mass can be obtained from finite element modal analyses.

4.2 The shear \( (V_1) \) corresponding to nominal yield moment can be calculated from the nominal yield moment \( (M_y) \) obtained from moment-curvature analysis.

4.3 For both methods the acceleration \( (a_{1\text{real}}, a_{2\text{real}}) \) corresponding to the yield shear can be calculated.

4.4 The elastic acceleration demand \( (A_1 \) and \( A_2 \) can be obtained from the elastic pseudo acceleration spectrum.
6. Compare the ductility demand and capacity.

5. The force reduction factors (R₁ and R₂) are calculated as the ratio between elastic demand (A₁ and A₂) and yield capacity (a₁[mean] and a₂).

4.6. The ductility demand can now be calculated as a function of the force reduction factor according to the equal displacement and equal energy principles.

5. The ductility capacity based on code drift limits can be determined. This is discussed later.

6. If the demand is greater than the capacity, choose a lower behaviour factor and repeat from step 3.

6.1 If the demand is less than the capacity, a lower behaviour factor can be chosen, resulting in an ultimate stress (fₚ) of 569 MPa. The ultimate strain capacity assumed was 7.5%. The stress-strain relationship equations used for the steel material model are taken from Priestley et al (2007: 140):

Elastic:  
\[ f_s = E_s \varepsilon_s \leq \varepsilon_y \]  
(5)

Yield plateau:  
\[ f_s = f_y \leq \varepsilon_y < \varepsilon_{sh} \]  
(6)

Strain hardening:  
\[ f_s = f_u - (f_u - f_y) \left( \frac{\varepsilon_{sh} - \varepsilon_s}{\varepsilon_{sh} - \varepsilon_y} \right) \]  
(7)

### Table 3: Material strengths

<table>
<thead>
<tr>
<th></th>
<th>Concrete</th>
<th>Reinforcement yield strength</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cube (design)</td>
<td>Cylinder (moment-curvature analysis)</td>
</tr>
<tr>
<td>Characteristic strength [MPa]</td>
<td>30</td>
<td>25</td>
</tr>
<tr>
<td>Mean strength [MPa]</td>
<td>39</td>
<td>33</td>
</tr>
</tbody>
</table>

### Design Equations

Design equations are used in step 2 of the methodology. The moment capacity of a wall cross section may be determined using an equivalent stress block method such as the
one set out by Bachmann et al (2002: 137). In this investigation the stress block method of SANS 10100-1 (2000) was used.

**DUCTILITY CAPACITY AND DEMAND**

It was stated in step 6 of the methodology that ductility demand will be compared to ductility capacity. This section shows how ductility capacity may be expressed as a function of inter-storey drift limits and how ductility demand may be calculated from ITHA results.

As shown in Figure 7, the displacement of a MDOF wall can be measured by an equivalent SDOF wall (Chopra 2007: 522-532). This equivalent SDOF wall must have the same dynamic characteristics as the first mode of the MDOF wall. In addition, the height of the wall is chosen such that the base moment of the SDOF wall due to the concentrated force \(F^*\) is equal to the base moment of the MDOF wall due to the distributed force (Priestley et al 2007: 316). This height \(h^*\) is referred to as the effective height.

In order to calculate ductility capacity as a function of a drift limit, equations for the drift profile and displacement profile at yield are sought. This is the point at which the curvature at the base of the wall is equal to the yield curvature \(\phi_y\). It is sufficient to assume a linear yield curvature profile (Priestley et al 2007: 317-319):

\[
\phi_{yi} = \phi_y \left( 1 - \frac{h_i}{h_w} \right) \tag{8}
\]

where:

- \(\phi_{yi}\) is the curvature at height \(h_i\)
- \(\phi_y\) is the yield curvature
- \(i = 0, 1, 2, ..., N\) is the storey number, and
- \(h_w\) is the height of the wall, defined in Figure 3.

Integration of Eq 8 with respect to the height produces an equation for the yield drift profile:

\[
\phi_{yi} = \phi_y \left( h_i - \frac{h_i^2}{2h_w} \right) \tag{9}
\]

Integration of Eq 9 produces an equation for the yield displacement profile:

\[
\Delta_{yi} = \frac{\phi_y h_w^2}{2} \left( 1 - \frac{h_i}{3h_w} \right) \tag{10}
\]

**Defining ductility capacity in terms of a code drift limit**

Ductility capacity is calculated in this study using both the plastic hinge method and an approximate equation.

**Plastic hinge method**

The yield displacement can be obtained from Eq 11 (Priestley et al 2007: 96):

\[
\Delta_y = \sum m_i \Delta_{yi} \tag{11}
\]

where:

- \(\Delta_{yi}\) is obtained from Eq 10.

The effective height can be calculated from Eq 12 (Priestley et al 2007: 100):

\[
h^* = \frac{\sum h_i m_i \Delta_{yi}}{\sum m_i \Delta_{yi}} \tag{12}
\]

where:

- \(\Delta_{yi}\) is the \(i^{th}\) value of the first mode shape vector.

The maximum yield drift can be calculated from Eq 9:

\[
\theta_y N = \phi_y \left( h_w - \frac{h_i^2}{2h_w} \right) \tag{13}
\]

Since this would be the maximum yield drift for all values of \(i\), the allowable plastic rotation is the difference between the code drift limit \(\theta_c\) and \(\theta_y N\). Having obtained the allowable plastic rotation, the plastic displacement at the effective height is:

\[
\Delta_p = (\theta_c - \theta_y N) h^* \tag{14}
\]

The ductility capacity in terms of the code drift limit is then

\[
\mu_c = \frac{\Delta_p}{\Delta_y} \tag{15}
\]

**Approximate equation**

Based on the following simplifying assumptions, Priestley et al (2007: 325-326) derived...
Thus, from Eq 13 the maximum yield drift is:
\[ \psi_y = \frac{\varepsilon_y h_w}{l_w} = \frac{\varepsilon_y A_y}{A} \]
(17)
where:
\[ A_y \] is the aspect ratio of the wall.

From Eq 10 the yield displacement profile can be described by:
\[ \Delta_y = \frac{\phi_y h_y}{l_y} \left( 1 - \frac{h_y}{3h_w} \right) \]
\[ = \frac{\varepsilon_y h_y^2}{l_w} \left( 1 - \frac{h_y^2}{3h_w^2} \right) \]
\[ = \frac{\varepsilon_y A_y h_w}{h_w} \left( 1 - \frac{h_y^2}{3h_w^2} \right) \]
(18)

The equivalent yield displacement can be obtained by substituting Eq 18 in Eq 11 and assuming equal floor masses (Priestley et al. 2007: 326):
\[ \Delta_y = \frac{\sum_{i} \Delta_y}{\sum_{i} \Delta_y} = 0.45\varepsilon_y A_y h_w \]
(19)

The effective height at yield, from Eq 12, is \( h_y^* = 0.77h_w \). Thus, by substituting Eq 17 in Eq 14, the plastic displacement is:
\[ \Delta_p = 0.77h_w \left( \frac{\theta_c - \varepsilon_y A_y}{\varepsilon_y A_y} \right) \]
(20)

Hence, from Eq 15, the ductility capacity is:
\[ \mu_d = \frac{\Delta_y + \Delta_p}{\sum_{i} \Delta_y} = \frac{0.45\varepsilon_y A_y h_w + 0.77h_w \left( \frac{\theta_c - \varepsilon_y A_y}{\varepsilon_y A_y} \right)}{0.45\varepsilon_y A_y h_w} \]
\[ = 1 + 1.71 \left( \frac{\theta_c - \varepsilon_y A_y}{\varepsilon_y A_y} \right) \]
(21)

Both the plastic hinge method and Eq 21 are used in this paper to calculate the ductility capacity in terms of the code drift limits prescribed by SANS 10160-4 (2011: 30) (see Figures 16 to 19).

Calculating ductility demand from inelastic time history analysis (ITHA) results

As stated in step 6.2 of the methodology, ITHA is used here to validate the ductility demand obtained from the equal displacement and equal energy principles. For each wall, ITHA is performed for a number of ground motion records. For each ground motion record the peak displacement of each degree of freedom (DOF) is recorded. The equivalent displacement of the average of the peak displacements, obtained from the different ground motions, can be calculated from Eq 22 (Priestley et al. 2007: 96):
\[ \Delta_{eq} = \frac{\sum_{i} \Delta_i^2}{\sum_{i} \Delta_i} \]
(22)
where:
\[ \Delta_i \] is the average of the peak displacement values of the \( i \)th DOF. The yield displacement is known from Eq 11, and thus the ductility demand can be calculated using Eq 23:
\[ \mu_d = \frac{\Delta_{eq}}{\Delta_y} \]
(23)

INELASTIC TIME HISTORY ANALYSIS

Degree of sophistication in element modelling

Line elements are beam-column elements with the ability to form plastic hinges at the ends of the member. With a suitable moment-curvature hysteresis rule assigned to the plastic hinges, the structural response can be predicted with remarkable accuracy (Priestley et al. 2007: 193). In this investigation the student version of Ruaumoko (Carr 2007) was used for ITHA.

Beam properties

The two types of line elements available in Ruaumoko are the elastic beam (Timoshenko beam – shear deformable) and the Giberson beam. The first storey was modelled with a Giberson beam element which, in addition to the elastic beam properties, contains a rotational spring at one end of the member representing the plastic hinge which forms at the base of the wall.

The upper part of the wall is required to remain elastic. Thus all higher storeys were modelled with elastic beam elements. An illustration of a typical finite element model of one of the walls of the investigation is shown in Figure 8.

Elastic properties

The input required for the elastic beam is summarised in Table 4:

As indicated in Table 4, the cracked sectional moment of inertia is obtained from the pre-yield branch of the bilinear moment-curvature relationship. Only one moment-curvature analysis was done for each wall, namely at the base of the wall (Dazio, Beyer & Bachmann 2009). The stiffness obtained from this analysis was applied over the full height of the wall. The properties obtained from the moment curvature analysis are illustrated in Figure 9.

Inelastic properties

In addition to the elastic section properties, the Giberson beam requires the input listed in Table 5.

<table>
<thead>
<tr>
<th>Table 4 Elastic beam properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>( E )</td>
</tr>
<tr>
<td>( G )</td>
</tr>
<tr>
<td>( A )</td>
</tr>
<tr>
<td>( A_s )</td>
</tr>
<tr>
<td>( l_{eff} )</td>
</tr>
</tbody>
</table>

Figure 8 Typical finite element model of a structural wall

a convenient equation which relates ductility to the code drift limit:

From a series of moment-curvature analyses, the yield curvature of a rectangular reinforced concrete structural wall is known to be (Priestley et al. 2007: 158):
\[ \phi_y = \frac{2\varepsilon_y}{l_w} \]
(16)

where:
\[ \varepsilon_y = 0.00225 \] is the yield strain of the longitudinal reinforcement, and
\[ l_w \] is the length of the wall section, defined in Figure 3.

Thus, from Eq 13 the maximum yield drift is:
\[ \psi_y = \frac{\phi_y h_y}{l_y} = \frac{\varepsilon_y A_y}{A} \]
(17)

where:
\[ A_y \] is the aspect ratio of the wall.

From Eq 10 the yield displacement profile can be described by:
Hysteresis rule

The Modified Takeda Rule shown in Figure 10 with a $\beta$ value of zero applies to structural walls (Priestley et al 2007: 201-202).

The unloading stiffness $k_u$ is a function of the elastic stiffness $k_o$ and the ductility at the onset of unloading ($\mu = \frac{u_m}{u_y}$) (Priestley et al 2007: 201):

$$k_u = k_o \mu^{-\alpha} \quad (24)$$

where:

$\alpha = 0.5$ is considered appropriate for reinforced concrete structural walls (Priestley et al 2007: 201). Tables 4 to 6 thus contain all input required for the Giberson beam.

Time step integration parameters

For this study Newmark’s average acceleration time-stepping method with time steps of 0.005 seconds was used (Chopra 2007: 175).

Ground motions

According to Priestley et al (2007: 210) it is sufficient to use the average response of a minimum of seven ground motion records. Spectrum-compatible accelerograms may be obtained through “manipulating existing ‘real’ records to match the design spectrum over the full range of periods” (Priestley et al 2007: 211). It has the advantage over purely artificial records that it preserves the essential character of the original real records (Priestley et al 2007: 211).

Thus it was decided to obtain real records with characteristics similar to that of ground types 1 and 4, and to manipulate these records to match the SANS 10160-4 (2011) elastic spectra. For this manipulation the student version of Oasys Sigraph (Oasys Limited 2010) was used.

Ground motion records were selected based on $v_{s,30}$ values and peak ground acceleration (PGA). The selected ground motions are listed in Table 7. Each earthquake has two orthogonal components. The seven ground motions were thus obtained from both components of the first three earthquakes and one component of the fourth. The records were obtained from the PEER NGA Database (2007).

These fourteen records were manipulated to match the SANS 10160-4 (2011) spectra. The pseudo acceleration spectra of the manipulated records are plotted in Figure 11 with the elastic SANS spectra.

Damping

Tangent-stiffness proportional damping was used with a damping ratio of 0.05 for the first mode (Priestley et al 2007: 207). When applying stiffness proportional damping, one should also be careful that the damping of the highest mode is less than 100%.
(Carr 2007). Thus, the damping in the highest mode was limited to 100%, resulting in some cases in a damping of less than 5% in the first mode.

**RESULTS**

**Design results (Figure 5(3) of the methodology)**

Figures 12 to 15 show the elastic-, capacity-, and design spectra of ground types 1 and 4.

- The design acceleration coordinates \( a_1 \) of the eight walls of this investigation, each with a different fundamental period, are shown on the design spectrum.
- The names of the walls, defined in Figure 4, are included in the figures. It may be seen that for design method 1, the design acceleration values \( a_1 \) are the same for walls of equal height, since Eq 4 depends only on the height of the wall.
- The capacity of the walls is also shown in Figures 12 to 15. For the purposes of this discussion, we refer to this as the capacity spectrum. The pseudo acceleration capacity was calculated from the yield moment capacity as described in step 4 of the methodology.

The relationship between the design spectrum and the capacity spectrum is influenced by three factors, namely over-strength, design conservatism, and period shift. These are briefly discussed below.

**Over-strength**

The capacity spectrum is higher than the design spectrum due to over-strength. The main factors which lead to over-strength are the following (Dazio & Beyer 2009: 3-21):

a. Mean material strengths, which are used to predict the most likely bending moment capacity of a section, are higher than the characteristic material strengths, used to predict bending moment capacity during design.

b. The provided reinforcement is always more than the required reinforcement.

**Design conservatism**

In this paper *design conservatism* is the name given to the assumption made during design that the design force is related to the total mass of the structure. To account in some way for the effect that higher modes inevitably have on the structure, the design seismic force is based on the total building mass, instead of the effective first modal mass. The effect of *design conservatism* is most clearly seen in Figure 13 by the steadily increasing capacity spectrum with increasing period.

<table>
<thead>
<tr>
<th>Record</th>
<th>Earthquake</th>
<th>PGA [g]</th>
<th>(v_{s,30} ) [m/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground type 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGA0023 San Francisco 1957-03-22 19:44</td>
<td>0.107</td>
<td>874</td>
<td></td>
</tr>
<tr>
<td>NGA0098 Hollister-03 1974-11-28 23:01</td>
<td>0.117</td>
<td>1 428</td>
<td></td>
</tr>
<tr>
<td>NGA0146 Coyote Lake 1979-08-06 17:05</td>
<td>0.120</td>
<td>1 428</td>
<td></td>
</tr>
<tr>
<td>NGA0680 Whittier Narrows-01 1987-10-01 14:42</td>
<td>0.102</td>
<td>969</td>
<td></td>
</tr>
</tbody>
</table>

| Ground type 4                  |          |                     |
| NGA0201 Imperial Valley-07 1979-10-15 23:19 | 0.141   | 163                 |
| NGA0780 Loma Prieta 1989-10-18 00:05 | 0.121   | 170                 |
| NGA0808 Loma Prieta 1989-10-18 00:05 | 0.132   | 155                 |
| NGA1866 Yountville 2000-09-03 | 0.150   | 155                 |

**Figure 11 Artificial ground motion spectra**

**Figure 12: Design results for ground type 1, design method 1**
The term period shift here refers to the difference in fundamental period predicted by the code (SANS 10160-4, 2011) in Eq 4 and the “true” period predicted by moment-curvature analysis of the cross section. Period shift only occurs for design method 1. The fundamental period calculated according to design method 2 is based on moment-curvature analysis, and thus no period shift occurs.

The relation of the demand spectrum to the capacity spectrum determines the extent to which the walls respond inelastically. As stated in step 4 of the methodology, the force reduction factor ($R$) is equal to the ratio between acceleration demand ($A_1$ or $A_2$) and capacity ($a_{1(rea)}$ or $a_{2(rea)}$). Thus, if the demand is less than the capacity, the force reduction factor is less than one, and thus no inelastic action is expected. This is illustrated in Figures 12 to 15 by the dividing line which intersects at the intersection of the demand and capacity spectra.

**Analysis results (steps 4 to 6 of the methodology)**

With the force reduction factor ($R$) known, the ductility demand can be calculated according to the equal displacement and equal energy principles and verified with ITHA. As previously discussed, the ductility capacity is based on code drift limits and is calculated according to the plastic hinge method and a simplified equation (Eq 21). Figures 16 to 19 show the comparison between ductility demand and ductility capacity for ground types 1 and 4, and design methods 1 and 2.

It is evident from Figures 16 to 19 that, on the capacity side, the plastic hinge method and the simplified equation (Eq 21) predict similar results. The simplified equation is, however, slightly conservative since it predicts a lower ductility capacity. The effect of the wall aspect ratio ($A_r$) on the ductility capacity is also evident. It was shown in Eq 21 (repeated here as Eq 25) that the ductility capacity reduces as the aspect ratio increases.

$$\mu_c = 1 + 1.71\left(\frac{\theta_c - \epsilon_y A_r}{\epsilon_y A_f}\right)$$  \hspace{1cm} (25)

It may also be seen that the ductility demand predicted by the equal displacement and equal energy principles corresponds to that of the ITHA.

The only wall to which the equal energy principle applied is the single-storey wall on ground type 4. For this wall the ductility capacity is exceeded by the ductility demand. This implies that the drift of the single-storey wall would exceed the code drift...
limits, and would thus suffer non-structural damage in excess of the design limit state. This does, however, only apply to walls with an aspect ratio of three or higher. This wall was only included in the scope of this investigation to obtain structural walls with a very short period. The aspect ratio was limited to three, since flexural response was desired of structural walls. In general, structural walls used for single-storey construction would have aspect ratios of less than three, and would therefore fall outside the scope of this investigation. The reader is referred to Paulay & Priestley (1992: 473) for the design of squat structural walls.

For all the other walls the ductility demand is less than the ductility capacity. Inter-storey drift levels for these walls are thus below code drift limits. It can be seen that the ductility demand reduces as the period increases. This is due to the artificial acceleration plateau of the design spectrum (see Figures 12 to 15). It can also be seen that method 1 produces “safer” structures than method 2 because of the assumption of a short period, and thus higher acceleration demand. Method 1, however, severely underestimates structural displacement.

It is therefore concluded that the current value of 5 of the behaviour factor, as defined by SANS 10160-4 (2011), is adequate to ensure that code drift limits are not exceeded, whether design is done according to method 1 or 2. The designer is, however, still required by the code to calculate structural displacements as the final step in the seismic design process (SANS 10160-4, 2011, p. 30).

**CONCLUSION**

The purpose of this investigation was to assess the value of the behaviour factor currently prescribed by SANS 10160-4 (2011) for the seismic design of reinforced concrete structural walls. The behaviour factor is used in seismic design to reduce the full elastic seismic demand on structures, since well-designed structures can dissipate energy through inelastic response. The behaviour factor was evaluated by comparing displacement demand with displacement capacity for eight structural walls.

Displacement demand was calculated by means of the equal displacement and equal energy principles and confirmed by inelastic time history analyses (ITHA). Displacement capacity was based on inter-storey drift limits specified by SANS 10160-4 (2011). These drift limits serve to protect building structures against non-structural damage.

Displacement demand was evaluated for two period estimation methods. Firstly, the fundamental period may be calculated from
an equation provided by the design code (SANS 10160-4, 2011), which depends on the height of the building. This equation is known to overestimate acceleration demand, and underestimate displacement demand. The second period estimation method involves an iterative procedure where the stiffness of the structure is based on the cracked sectional stiffness obtained from moment-curvature analysis. This method provides a more realistic estimate of the fundamental period of structures, but due to its iterative nature it is seldom applied in design practice.

The conclusion of this investigation is that the current behaviour factor value of 5, as found in SANS 10160-4 (2011), is adequate to ensure that structural walls comply with code-defined drift limits. This applies to both period estimation methods.

NOTE
1 Not to be confused with the ‘Capacity Spectrum Method’ by Freeman (2004).

REFERENCES


Figure 19 Analysis results for ground type 4, design method 2

Ductility
Capacity – simplified equation
Demand – R–μ–T relationship
Capacity – plastic hinge method
Demand – ITHA
A rational approach to predicting the buckling length of compression chords in prefabricated timber truss roof structures braced by means of diagonal bracing

W M G Burdzik, N W Dekker

In South Africa, timber-trussed roofs supporting concrete tiles have for many years often been braced solely by means of diagonal braces. Failures have shown that the diagonal brace was inadequate for larger span roofs, and the use of diagonal bracing has subsequently been limited to spans of less or equal to 10 m. When designing the compression chords of a timber truss in a braced roof, SANS 10163:1 (2003) recommends a minimum effective length for out-of-plane buckling of not less than 15 x b, which is 540 mm for a 36 mm wide member. This effective or out-of-plane buckling length of the top chord was later assumed to be equal to the spacing of the trusses. With the availability of PC-based packages that are able to perform three-dimensional buckling analyses, it is perhaps useful to investigate the validity of using the effective length equal to the truss spacing, and then also the 10 m limit on span for roofs braced by means of diagonal braces.

A common error made when analysing three-dimensional buckling problems is to assume connectivity on the centreline of the members, thereby neglecting eccentricity between the centreline of the bracing and the centreline of the member being braced (see Figure 1). In timber-trussed roofs, the diagonal brace is nailed to the underside of the top chord of a number of adjacent trusses. The brace runs at more or less 45° and triangulation appears to be complete when viewed on plan, as the battens form the other elements of the bracing system triangulation. Trusses some distance from the trusses that are connected to the diagonal brace can, however, only obtain lateral support via the battens that are connected to the top of the compression chords. The authors feel that a more correct way of analysing a timber-trussed roof, braced by means of a diagonal brace, requires that the eccentricity between the centreline of the battens on top of the compression chords and the centreline of the braced points on the underside of the compression chords be taken into account. Furthermore, the connections between the battens and the top chord are not infinitely stiff and this stiffness, together with the low torsional rigidity of the timber members, should be taken into account in the buckling analysis. The analysis can be further improved by taking the out-of-plane bending stiffness of the web members into account. All these factors will influence the buckling length of the compression chords to some degree.

In this paper, the authors show how incorrect assumptions may mislead the designer into believing that the buckling length is equal to or less than the spacing of the trusses. They also show that, even though the bracing members have been placed on the correct sides of the top chord in the analysis, incorrect assumptions about the torsional stiffness of the top chords can lead to buckling lengths that are slightly less than when a more realistic torsional stiffness is used.

INTRODUCTION

Since the introduction of computer-based analysis programs for timber roofs, the pre- and post-processing parts of the software have changed and improved to the extent that the designer is no longer aware of the design process, even though forces, displacements, sizes and assumed effective lengths may be printed for checking by a competent person. Loads are calculated from the layout and these are applied to a two-dimensional analysis of the truss, even though a timber roof structure is a three-dimensional problem constructed out of a brittle material with limited ductility in the connections. Limited ductility can be a problem in cases where construction errors have been made and force-fitting is applied. Assumptions are made about the member sizes and sometimes the connector plate stiffness for an initial

Keywords: effective length, buckling, flexible supports, bracing

Journal of the South African Institution of Civil Engineering • Volume 54 Number 1 April 2012
analysis. More often than not, a centreline analysis using beam elements is used and the forces so obtained are used to size the members in accordance with SANS 10163: Part 1 (2003) or Part 2 (2001). At this stage of the design, assumptions are made about the type of bracing to be used, as well as the effective or buckling length of the compression chord that would result from using that specific type of bracing. In most cases the trusses are re-analysed with the correct sizes once these have been calculated to ensure that the sizing is adequate for the load and that the deflection is not excessive.

SANS 10163: Part 2 (2001) has a rule of thumb for the minimum slenderness, Le/b, equal to 15 for tiled roofs or the spacing of the purlin for sheeted roofs. Many believed, and some still believe, that the effective buckling length of the top chord is equal to the spacing of the battens. This assumption would perhaps be correct if the tiles could be relied on to supply diaphragm action, and that the battens were rigidly connected to the compression member. In such a case, any further bracing would only be required for erection purposes. The authors accept that diaphragm bracing by the tiles will initially be active, but with time the friction between the tiles seems to break and movement does occur. This eventual movement of the tiles has led to failures of roofs.

The assumption of the buckling length equal to $15 \times b$, with $b = 36$ mm, may not be a problem when the spacing of the trusses is equal to 640 mm as is commonly used in Australia, but could become a problem where the spacing of the trusses is as much as 1 050 mm, as is often found in South Africa. The minimum slenderness of $Le/b = 15$ was later changed by the South African Institute for Timber Construction to an in-house rule which suggests, the operative word being suggests, an effective buckling length of the spacing of the trusses, i.e. 750 mm to 1 050 mm. The authors believe that a blanket rule such as effective length $= 15 \times b$ or even buckling length equal to the spacing of the trusses may not be conservative, as the buckling length depends on the boundary conditions, the stiffness of the bracing and the method of transferring loads once buckling is initiated.

For small span timber trusses, i.e. up to 10 m, a diagonal brace is the norm in South Africa. As the limit on the span for the use of a diagonal brace is less than 10 m, only the 10 m span and 7.5 m spans were investigated. Timber sizes for 10 m span roofs would typically be 36 mm $\times$ 111 mm top and bottom chords with 36 mm $\times$ 73 mm web members. When bracing a 10 m span trussed roof, a 36 mm $\times$ 111 mm timber member is fixed to the underside of the compression chords and runs at about 45° when seen in plan. Three 100 mm long nails are used to fix the brace to the underside of the top chord (see Figure 1). Maximum spacing rules are used to ensure that trusses are not too far from the brace. Battens, the smallest nominally being 36 mm $\times$ 36 mm, are placed on top of the compression chord and these are then fixed with one 75 mm long nail to the compression chord.

In this paper it is the intention of the authors to only investigate diagonal bracing, and not all forms of bracing that are currently used by the South African timber roof truss industry. Although this investigation is a theoretical exercise and cannot, at this stage, be validated by test results, the authors are of the opinion that a three-dimensional buckling analysis is an acceptable way of determining the buckling length of a compression chord in a timber roof structure. Buckling and finite element analyses are widely used for many structural systems and materials as the analyses are based on theories that have historically been proven to work for structures.

**EFFECTIVE LENGTH FACTORS IN SIMPLE LATTICE STRUCTURES**

**In-plane and out-of-plane buckling**

The effective length factor, or K-factor, is used to adjust the actual unrestrained length of a compression member to account for prevailing boundary conditions. Many software packages use a default out-of-plane effective length factor of 0.85, implying some form of rotational joint restraint by adjacent members. This is only possible where adjacent members have high out-of-plane bending or torsional stiffness and they themselves are not compression members that could buckle. Some design codes specify effective length factors for compression members in lattice structures.
trusses. BS 5400 Part 3 (BSI 2000), in its Table 11, specifies effective length factors for buckling in the plane of the truss, as well as out of the plane of the truss. In all cases the values given in Table 11 of BS 5400 Part 3 (BSI 2000) are less than or equal to 0.85. Eurocode 5 (CEN 1995) gives the effective column length for members of triangulated trusses with loading at the nodes as the bay length. For strength verification, Eurocode 5 (CEN 1995) states that the calculated force should be increased by 10%.

SANS 10163 Part 1 (2003) has non-mandatory clauses for the calculation of the effective length factor:

1. With regard to the effective length for “in-plane” buckling, in a continuous compression member such as a chord of a truss, take the effective length for “in-plane” buckling as the distance between the node points multiplied by a factor of between 0.85 and 1.0, depending on the degree of fixity and the distribution of the load. In a non-continuous compression member such as the web of a truss, take the effective length for “in-plane” buckling as the actual length of the member multiplied by a factor of between 0.85 and 1.0, depending on the degree of end fixity.

2. With regard to effective length for “out-of-plane” buckling, the following apply:
   a) take the effective length of the compression chords to be equal to the purlin or batten spacing, provided that the purlins or battens are adequately fixed to the chords, properly spliced to transmit the forces and adequately braced against longitudinal movement;
   b) in the case of tiles supported on battens, the battens being spaced less than 400 mm apart and fixed to the chords with one or two plain wire nails, use a minimum slenderness value of Le/b = 15 for calculating the ultimate compressive stresses for the chords;
   c) if the compression chords are braced by means of a bracing frame or a truss that restrains the longitudinal movement of all battens, use the minimum slenderness value given in (b) above; and
   d) in the case of web members, use the distance between the intersection of the centre lines of connecting members.

Boundary conditions influencing the degree of restraint exercised on a compression member, are not merely a function of connection details and continuity, but are influenced by the capacity of adjacent members at the node. Consider the example of a simple lattice truss with a constant section, shown in Figure 2, where lateral supports are provided at nodes, A, B, C, D, E, F and G:

The compression chord ABCDEFG is divided into equal portions. The basic principle that the buckling load for member ABCDEFG is unique shows that an effective length factor of less than one for a particular member is consistent with an effective length factor of greater than one in the adjacent members, albeit with a smaller force, as shown in Equation 1:

\[
\frac{n^2EI}{(k_{AB}L)^2} \cdot P_{AB} = \frac{n^2EI}{(k_{BC}L)^2} \cdot P_{BC} = \frac{n^2EI}{(k_{CD}L)^2} \cdot P_{CD}
\]

If the member is of constant section with a constant force, Equation 1 will be satisfied if the buckling length is taken as \(L = L_{AB} = L_{BC} = L_{CD}\). If the torsional stiffness of the lacing elements is ignored, an effective length factor of less than one is clearly incorrect. For the loading as shown, the force in ABCDEFG will vary over the length of the truss. Once again, the buckling strength of the chord ABCDEF is unique. In order to still satisfy Equation 1, the effective length factors for the members with the lower forces are greater than for those members with the higher forces.

It is significant that the Eurocodes for steel design specifically have discarded the practice of using tabulated effective length factors in the design of both compression and flexural members. Elastic buckling loads are used as a basis of design, and such loads are commonly calculated using computer programs.

Stanway, Chapman & Dowling (1992) have discussed the influence of elastic supports at any position of the length of the strut, thereby considering the influence of unequal bay lengths and the beneficial restraint offered by adjacent subcritical elements having a shorter buckling length. The basis of elastic buckling analysis is subsequently discussed.

Most PC-based analysis packages are capable of performing buckling analyses on framed structures using beam elements, and individual members using shell elements. It is important that the user be aware of the actual process and the premises on which such analyses are based. The method of buckling analysis of a frame structure is presented in Coates et al (1988), as described below.

In the case of a linear elastic analysis of a framed structure, deformation is linearly related to load, or, expressed in matrix form:

\[
P = K \Delta
\]

where:

- \(P\) is the force or load matrix,
- \(\Delta\) is the displacement matrix,
- \(K\) is the stiffness matrix.

The terms of \(K\) are constant for a given structure, provided that second order effects are neglected, therefore \(K\) is independent of \(P\). If, however, the influence of axial forces on member stiffness is included, \(K\) becomes a function of \(P\), or \(K = K(P)\).

In the case where axial loads are not neglected, Equation 2 becomes non-linear, but if the axial loads are known, the deflections may be calculated.

\[
\lambda P = K_s(\lambda P)\Delta
\]

In Equation 3 the term \(\lambda\) has been inserted as a load multiplier. As the loads are progressively increased, a state of neutral equilibrium is achieved where any deflection is possible for a given load level. This state defines instability and may be referred to as \(\lambda = \lambda_{cr}\). The critical state is consistent with the matrix \(K_s(\lambda P)\) becoming singular.

A test of the singularity of the matrix \(K_s\) can therefore be used as a check on stability. If it is non-singular and positive definite, the structure is stable; if it is singular, the structure is on the point of collapse. The value of \(\lambda_{cr}\) is therefore a multiple whereby an arbitrarily chosen load can be multiplied
to achieve a state of collapse. The following comments regarding the value of $\lambda_{cr}$ should be clearly noted:

- $\lambda_{cr}$ is not a safety factor. Even if $P$ is chosen to represent load effects at working loads, the influence of inelastic buckling is not taken into account in an elastic buckling analysis.
- In the case where buckling modes are de-coupled, for example lattice structures consisting of pin-ended members, the value of $\lambda_{cr}$ applies to the member most susceptible to buckling, and has no application to other members.
- If a two-dimensional analysis were to be carried out to determine $\lambda_{cr}$, the value of $\lambda_{cr}$ applies to in-plane and not out-of-plane buckling.

The significance of an elastic buckling analysis is that the value of $\lambda_{cr}P$ is the elastic buckling load of the critical member or portion of a structure, or of the structure as a whole. Dekker and Burdzik (2000) have shown that, in order to calculate the inelastic buckling load of the critical member or to achieve a state of collapse. The following procedure may be followed:

1. Calculate the effective equivalent un-braced length from the relationship:

   \[ \lambda_{cr} = \frac{\pi^2 KL}{P} \]  

   Therefore:

   \[ KL = \frac{\pi^2 P}{\lambda_{cr}^2} \] (4)

   or

   \[ KL = \frac{P^2}{\lambda_{cr} P} \] (6)

   where:

   \( P_e \) = Euler buckling load for a compression member hinged at both ends

2. Calculate the compressive resistance $C_r$ using the value of $KL$ obtained from Equation 5 for the appropriate member size. The resistance equation is given by SANS 10163:1 (2003):

   \[ C_r = \frac{P_{cr} A}{\gamma_m} \times \frac{f_c}{\sqrt{f_{cr}}} \] (7)

   The appropriate slenderness ratio is given in SANS 10163:1 (2003):

   \[ \lambda = \frac{K \times L}{r} \times \frac{f_c}{\sqrt{f_{cr}}} \] (8)

   or for a rectangular section

   \[ \lambda = \frac{f_c}{\sqrt{f_{cr}}} \frac{KL}{P} \] (9)

   and the buckling factor, $\beta_b$, is given by:

   \[ \beta_b = (1 + \lambda^2/n)^{1/n} \] (10)

   with $n = 1.8$ and $\lambda_{cr}$ is the elastic buckling factor, $K$ is the effective un-braced length factor

   The required flexural stiffness EI can be calculated as follows:

   \[ EI_{BA} = K_{ser} \times \frac{L_{AB}^3}{12} \] (12)

   The equivalent diameter of a steel nail that now connects the batten to the top chord can be determined from this equation. If the modulus of elasticity of the steel is 206 GPa and Equation 12 is applied:

   \[ I_{AB} = K_{ser} \times \frac{L_{AB}^3}{12} \times \frac{800 \times 0.0735^2}{206 \times 10^6 \times 12} \]

   \[ = 1.285 \times 10^{-10} \text{ m}^4 \] (13)

   The diameter of the equivalent round nail with the required second moment of area is 7.2 mm. It may be prudent to remember that the theoretical stiffness of the nails is based on tests where the nail is forced into double curvature by longitudinal displacement of the connected members. However, if the chord is free to rotate, this stiffness will be reduced, i.e. ‘soften’, and it would
no longer be correct to assume a spring stiffness of 800 kN/m when using springs in the analyses. It would be more correct in the authors' opinion to have a stiffness of something between fully fixed on both ends and a cantilever. For a cantilever the stiffness was required to achieve this was 267 kN/m and no longer 800 kN/m. This shows that, when using shell or plate elements that are connected to the bracing battens by way of springs, great care should be taken as it may result in misleading answers. If the spring stiffness of 800 kN/m is used, the buckling load factor, \( \lambda \), was 293 with a buckling length of 454 mm. Both the buckling lengths are substantially greater than the spacing of 446 mm. This may explain why the friction between the tiles, leading to bracing by diaphragm action. Tile mass was taken as being 55 kg/m², although the actual mass is not that important, as the buckling analysis is only used to calculate buckling lengths. The different configurations (see Figures 5 and 6) were used to ascertain whether the configuration would play a significant part in the buckling length of the compression chord.

**SHEAR MODULUS**

It is accepted that the shear modulus of South African pine is about equal to MOE/13 (Burdzik & Nkwera 2003). Not all software packages have the facility to input the shear modulus. Prokon (2011), the package used in the following analyses uses a Poisson Ratio, \( \nu \), of 0.2. The shear modulus is then calculated from the following equation:

\[
G = \frac{E}{2(1 + \nu)} = \frac{E}{2.4}
\]  

This shortcoming can be overcome by reducing the St Venant torsional constant of the relevant members by 13/2.4 = 5.42.

**ANALYSES**

In order to demonstrate the principles discussed above, a commonly available PC-based analysis package, Prokon (2011), was used to calculate the effective length factors of the top chord of gable to gable timber-trussed roofs with spans of 7.5 m and 10 m, with pitches of 17.5°, 25° and 35°. The batten spacing was taken as 262 mm and 305 mm respectively in order to simplify the input of the truss and batten geometry. Only the tile weight and the self weight of the timber were used to determine the buckling length of the top chords, as the buckling is a long-term problem, rather than a problem that occurs when imposed load is applied, as imposed load will increase the friction between the tiles, leading to buckling of the compression member. A buckling load factor of 174 and an out-of-plane buckling length of 622 mm resulted.

**Results of the analyses**

In all cases the top and bottom chords were assumed to have dimensions of 36 mm × 111 mm with web members being 36 mm × 73 mm with a 36 mm × 111 mm diagonal brace, although in practice the top and bottom chords may be 36 mm × 73 mm and the diagonal bracing member a
36 mm × 73 mm for small span roofs. A full span, complete roof (see Figure 7) was analysed to ascertain the buckled shape of the roof so that a half-structure, with the correct boundary conditions, could be analysed. From the buckled shape, one can deduce that the apex moves as the brace is flexible and it then becomes apparent that one cannot assume an inflection point at the apex. This then makes it possible to define the boundary conditions for a structure where only the half structure is investigated. If the half structure with the correct boundary conditions is used, it simplifies the input and speeds up the analyses of the various truss layouts and spans. Assume that the truss lies in the X-Y plane and that the Z axis is perpendicular to that plane. The eaves of the truss is supported in Y and Z directions, whereas the apex and the bottom chord of the half truss are supported in the X direction and fixed against rotation about the Y axis. Figure 8 shows the buckled shape of the half-structure with a span of 10 m.

It is noteworthy to see how the top chords of the trusses will buckle, as this buckling shape is sometimes visible on the tile lines of some of the older houses in South Africa. The battens force sympathetic buckling, i.e. all top chords move in the same direction, as all the top chords are tied together by the battens. It is also apparent from the buckled shape that the low torsional rigidity of the top chord and the stiffness of the nails all play a part in the final buckled shape. A buckling analysis of a trussed roof, where centreline connectivity of chords, battens and diagonal brace is assumed, is shown in Figure 9. Note the difference in the buckled shape of the roof shown in Figure 8, where the actual relative position of the members is taken into account, and Figures 9, where centreline connectivity is assumed. The effect of the difference in the buckled shape leads to a difference in the buckling length of the top chord as is given in Table 2.

This difference in the buckled shape and the buckling load factor between centreline connectivity and the buckled shape of the trussed roof where the relative distances between the centreline of the battens, top chord and diagonal brace are taken into account, shows how important it is to take the actual position into account. Even if one allows for the stiffness of the nails in the centreline analysis, one may still land up with what the authors believe to be an incorrect evaluation of the buckling length of the top chord.

The results of the analysis for the different configurations of the 7.5 m span trusses are given in tabular form in Tables 1, 2 and 3. By considering the results shown in Tables 1 to 3 it is clear that the actual
buckling length exceeds the purlin spacing by a factor of between 3.8 and 4.4. The error caused by centreline modelling is shown to be significant in Table 2. No further centreline modelling was undertaken as the authors were convinced that the difference in the buckling factors for the different layouts would not be significant.

The results of the analysis for the different configurations of the 10.0 m span trusses are given in tabular form in Tables 4, 5 and 6.

**ULTIMATE STRENGTH OF TRUSSES**

To see whether the theoretical increased buckling length will negatively influence the design of timber trusses, one each of the 7.5 m and 10 m trusses will be used to illustrate the code requirements between using the effective buckling length based on the truss spacing, and the theoretical buckling length of 1.2 m.

**Fink truss layout at 7.5 m span with 25° pitch**

If one includes the imposed load, the imposed load on a tributary area of 5.63 m² is 0.46 kN/m². A frame analysis on the truss with pin joints between the web members and the chords, but with continuity of the top and bottom chords, was applied. With a total load that includes all self-weight and imposed load, the ultimate axial force in the top chord = 8.96 kN and the ultimate imposed load on a tributary area of 5.63 m² is 0.46 kN/m². A frame analysis on the truss fink truss layout at 7.5 m span with 25° pitch.

The axial resistance using a buckling length of 750 mm can be calculated from Equation 8:

\[
λ = \frac{K \times L}{P} \times \sqrt{\frac{f_c}{\pi^2 \times E_{\text{mean}}}}
\]

\[
= \sqrt[12]{12} \times \frac{K \times L}{b} \times \sqrt{\frac{f_c}{\pi^2 \times E_{\text{mean}}}}
\]

\[
= \sqrt[12]{12} \times 750 \times \frac{18}{36} \times \frac{1}{\pi^2 \times 7,800}
\]

= 1.104

and the buckling factor \(β_b\) is given by Equation 9:

\[
β_b = \left(1 + \lambda^{2/3} \right)^{-1/n}\]

\[
β_b = \left(1 + 1.104^{2/3} \right)^{-1/1.8}
\]

= 0.611

The resistance of a 36 mm × 111 mm Grade 5 member is therefore (Equation 7):

\[
C_r = φ \times A \times \frac{β_b \times f_c}{f_m}
\]

\[
y_{md} = 0.60 + 0.63 \times 0.63 = 0.6 + 0.63 \times 0.54 = 0.94
\]

<table>
<thead>
<tr>
<th>Table 1 Results of analysis: 7.5 m span, 17.5° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Fink</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Queen post</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2 Results of analysis: 7.5 m span, 25° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Fink</td>
</tr>
<tr>
<td>Fink*</td>
</tr>
<tr>
<td>Fink#</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Queen post</td>
</tr>
</tbody>
</table>

* with incorrect torsional stiffness of the top and bottom chords

# with centreline analysis

<table>
<thead>
<tr>
<th>Table 3 Results of analysis: 7.5 m span, 35° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Fink</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Queen post</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4 Results of analysis: 10 m span, 17.5° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Double-W</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Fan</td>
</tr>
</tbody>
</table>

\(P_{\text{assumed}}\) is based on a buckling length of 750 mm

<table>
<thead>
<tr>
<th>Table 5 Results of analysis: 10 m span, 25° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Double-W</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Fan</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6 Results of analysis: 10 m span, 35° pitch, 11 trusses braced by the diagonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truss Type</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Double-W</td>
</tr>
<tr>
<td>Howe</td>
</tr>
<tr>
<td>Fan</td>
</tr>
</tbody>
</table>

Journal of the South African Institution of Civil Engineering • Volume 54 Number 1 April 2012
\[ y_{n3} = 0.82 + 0.023 \times 7.5 = 0.993 \text{ for 7.5 m span} \]

The product of the \( \gamma \) factors is equal to 0.93.

\[ Cr = 0.67 \times 36 \times 111 \times \frac{0.611 \times 18}{0.93} = 31.66 \text{ kN} \]

The moment of resistance,

\[ M_r = \phi \times Z_c \times \frac{f_b}{f_m} \]

(15)

Using a buckling length of 1.2 m the slenderness and the buckling factor are:

\[ \lambda = \frac{K \times L}{r} \times \sqrt{\frac{f_b}{E_{\text{mean}}}} = \frac{\sqrt{12} \times K \times L}{b} \times \sqrt{\frac{f_b}{E_{\text{mean}}}} \]

\[ = \frac{\sqrt{12} \times 7}{36} \times \frac{100}{18} \times \frac{36 \times 111^2}{2 \times 7 \times 800} \]

\[ = 1.766 \]

and the buckling factor \( \beta_b \) is given by:

\[ \beta_b = (1 + 1.766^2 \times 1.8)^{-1/1.8} \]

\[ = 0.30 \]

\[ C_r = 0.67 \times 36 \times 111 \times \frac{0.30 \times 18}{0.93} = 15.55 \text{ kN} \]

The moment of resistance,

\[ M_r = \phi \times Z_c \times \frac{f_b}{f_m} \]

\[ = 0.67 \times 36 \times 111^2 \times \frac{11.5}{0.93} = 0.61 \text{ kN.m} \]

The Interaction Index:

\[ \frac{C_u}{C_r} = \frac{8.96}{15.55} = 0.577 \]

The buckling analyses and calculations would appear to justify limiting the span of trusses that are braced by means of a diagonal brace to less than 10 m, as the capacity of the nailed connections between the battens and the braced trusses may be exceeded once buckling is initiated. Owing to the many uncertainties involved, as well as the number of failures noted, it is proposed that the buckling length should be increased to 1.2 m or 30 \( \times \) b for timber-trussed roofs that are braced solely by means of diagonal bracing. Perhaps there should be two interaction equations for checking the lateral buckling strength of the roof trusses. The first check should be to ascertain whether the truss strength is adequate for permanent load with the increased buckling length, and the second for total load, with the buckling length, however, reduced to 15 \( \times \) b.

**REFERENCES:**


**Double-W truss layout at 10 m span**

If one includes the imposed load, the imposed load on a tributary area of 7.5 m² is 0.425 kN/m². With a total load that includes all self-weight and imposed load, the ultimate force and bending moment in the top chord was found to be \( C_u \) = 13.24 kN and \( M_u \) = 0.23 kN.m.

\[ \gamma_{m1} = 0.60 + 0.63 \times w_{ds} = 0.6 + 0.63 \times 0.54 = 0.94 \]

\[ \gamma_{m3} = 0.82 + 0.023 \times L = 0.82 + 0.023 \times 10.0 = 1.05 \text{ for 10 m span} \]

With an assumed buckling length of 0.75 m and an assumed top chord size of 36 mm \( \times \) 111 mm in a Grade 5 timber, the resistances become:

\[ C_r = 0.67 \times 36 \times 111 \times \frac{0.611 \times 18}{0.93} = 29.83 \text{ kN} \]

The moment of resistance,

\[ M_r = \phi \times Z_c \times \frac{f_b}{f_m} \]

\[ = 0.67 \times 36 \times 111^2 \times \frac{11.5}{0.93} = 0.577 \text{ kN.m} \]

The interaction equation indicates that the compressive force in the lateral chords exceeds the design load. Perhaps there should be two interaction equations for checking the lateral buckling strength of the roof trusses. The first check should be to ascertain whether the truss strength is adequate for permanent load with the increased buckling length, and the second for total load, with the buckling length, however, reduced to 15 \( \times \) b.

**CONCLUSION**

The theoretical buckling lengths of the compression chords, of trusses in a roof braced by means of a diagonal brace, are shown to be in the region of between 1 m and 1.2 m for a gable-to-gable timber roof structure. This increase in the buckling length from 0.76 m to over 1.0 m may not be critical for roofs that have been designed for a buckling length of 0.76 m, or the spacing of the trusses, as the imposed load is very seldom applied to the full roof. Furthermore, the imposed load would increase the friction between the tiles, thereby perhaps leading to diaphragm bracing. The 30% shortfall in capacity should not impact significantly on the probability of failure of the compression chords, provided that the integrity of the connections between the trusses and the battens is maintained.

Ignoring the lack of torsional stiffness of the top chord also has a small effect on the buckling length obtained from the analysis. This may not be true for sections that have a greater depth, i.e. depth of 149 mm and 225 mm. However, a centreline analysis neglecting to consider the distance between the centrelines of the brace, the chords and the battens is shown to under-estimate the theoretical buckling length by a dangerous margin, possibly leading to unsafe member sizes.

The buckling analyses and calculations would appear to justify limiting the span of trusses that are braced by means of a diagonal brace to less than 10 m, as the capacity of the nailed connections between the battens and the braced trusses may be exceeded once buckling is initiated. Owing to the many uncertainties involved, as well as the number of failures noted, it is proposed that the buckling length should be increased to 1.2 m or 30 \( \times \) b for timber-trussed roofs that are braced solely by means of diagonal bracing. Perhaps there should be two interaction equations for checking the lateral buckling strength of the roof trusses. The first check should be to ascertain whether the truss strength is adequate for permanent load with the increased buckling length, and the second for total load, with the buckling length, however, reduced to 15 \( \times \) b.

**REFERENCES:**


PROKON Software Consultants Ltd 2011. PROKON suite of structural analysis programs. Pretoria: PROKON.


Thanks to our Referees!

The SAICE Journal Editorial Panel would like to thank the persons listed below, all of whom served as referees during 2010 and 2011. The quality of our journal is not only a reflection of the technical excellence of participating authors, but certainly also of the high standards set by our referees.

Mr Dotun Adegoke  Mr Trevor Green  Dr Martin Rust
Prof Mark Alexander  Ms Marietjie Griffioen  Mr Matthew Sandham
Mr George Annandale  Mr Joe Grobler  Prof Manus Santhanam
Dr Joseph Anechie-Boateng  Prof Johannes Haarhoff  Mr Jonathan Shamrock
Prof Yunus Ballim  Prof Gerhard Heymann  Mr Rob Sik
Dr Celeste Barnardo  Prof Stephan Heyns
Dr Robert Beale  Dr Emile Horak
Dr Ken Been  Dr SW Jacobsz
Dr Hans Beushausen  Ms Karin Jansen van Rensburg
Mr David Blitenthall  Prof Kim Jenkins
Dr Billy Boshoff  Mr Werner Jerling
Dr Johan Bosman  Dr Gary Jones
Prof Walter Burduzik  Mr Stephen Joseph
Dr Amanda Cassa  Prof Elsabé Kearsley
Dr Nicol Chang  Mr Gerhard Keyter
Dr Byron Daniels  Dr Geoff Krije
Mr Peter Day  Prof Lynesse Laloui
Mr Jaco de Villiers  Dr Kuinian Li
Prof Romano del Mistro  Prof Simon Lorentz
Dr Erik Dennenman  Dr Erik Louber
Dr Mark Dent  Mr Keith Mackie
Dr Gerhardus Diedericks  Dr Jeffrey Mahachi
Dr Christosa du Plessis  Dr Joe Mahoney
Ms Hanli du Plessis  Dr James Maina
Dr Rudi du Preez  Mr Sandy Melvill
Mr Heinrich Elges  Dr Martin Mgangira
Mr Trevor Green  Mr Lionel Moore
Dr Alex Elvin  Ms Santie Gouws
Dr Yves Filion  Dr Philip Gouws
Dr Leon Geustyn  Mr Alan Parrock
Dr Jeremy Gibert  Mr Bryan Perrie
Prof Mitchell Gohnert  Prof K Ramamurthy
Ms Santie Gouws  Dr Hubrecht Ribbens
Dr Philip Goyds  Mr Louis Roodt
Mr Philip Goyns  Prof Chris Roth
Prof Hannes Grabe

†

Dr Fritz Wagener
Dr Pieter Venter
Mr Benoit Verhaeghe
Dr Nico Vermeulen
Prof Alex Visser
Prof Deon von Willich
Dr Pieter Wessels
Prof Jan Wiim
Prof Alphose Zingoni
Guidelines for the preparation of papers and technical notes

The Journal of the South African Institution of Civil Engineering is published biannually in April and October. Articles submitted for publication are reviewed by a panel of referees under the guidance of the SAICE Journal Editorial Panel, a sub-committee of the SAICE Editorial Panel. When preparing articles for publication, authors should please take note of the following and comply with the guidelines as set out:

CLASSIFICATION OF ARTICLES CONSIDERED FOR PUBLICATION

- Technical papers are well-researched, in-depth, fully referenced technical articles not exceeding 6 000 words in length (excluding tables and illustrations). Related papers that deal with softer sciences (e.g. education, social upliftment, etc) are accepted if they are of a technical nature and of particular interest to the civil engineering profession. The latter type of paper will be subject not only to peer-review by civil engineers, but also to review by non-engineering specialists in the field covered by the paper.

- Technical notes are short, fully referenced technical articles that do not exceed 2 000 words. A typical technical note will have limited scope often dealing with a single technical issue of particular importance to civil engineering. An article incorrectly submitted as a technical paper can, with the author’s agreement, be re-categorised by the Panel as a technical note. Alternatively, on recommendation by the appointed referees, the Panel may request an author to condense a technical paper into a technical note.

- Reviews and case studies are considered for publication as either technical papers or technical notes on condition that they are the original work of the author and will assist the reader with the understanding, interpreting or applying of the subject under review.

- Discussion on published articles is welcomed up to six months after publication. The length of discussion contributions is limited to 1 500 words. Where appropriate, discussion contributions will be subject to the normal reviewing process and will be forwarded to the authors of the original article for reply.

POLICY REGARDING LANGUAGE AND ORIGINALITY OF SUBMITTED ARTICLES

- Language: Manuscripts should preferably be presented in English, as the journal is distributed internationally. Articles submitted in any of the other official South African languages should be accompanied by an expanded abstract in English.

- Original work: Papers and technical notes must be original contributions. Authors must affirm that submitted material has not been published previously, is not under consideration for publication elsewhere and will not be submitted elsewhere while under consideration by the SAICE Journal Editorial Panel. It is the responsibility of the authors to ensure that publication of any paper in the journal will not constitute a breach of any agreement or the transgression of any law. The corresponding author should confirm that all co-authors have read and approved the manuscript and accept these conditions. Authors are responsible for obtaining permission to publish experimental data and other information that is confidential or sensitive. Authors are also responsible for obtaining permission from copyright owners when reproducing material that has been published elsewhere.

- Copyright: On acceptance of the paper or technical note, copyright must be transferred by the author/s to the South African Institution of Civil Engineering on the form that will be provided by the Institution.

- Photos of authors: The final corrected version of the paper should be accompanied by recent head and shoulders colour photographs of the authors.

- Footnotes, trade names, acronyms, abbreviations: These should be avoided. If acronyms are used, they should be defined when they first appear in the text. Do not use full stops after abbreviations or acronyms.

FINAL ARTICLE

- Figures, tables and equations should be numbered consecutively and should appear in the text directly after they have been referred to for the first time.

- Figures should be produced using computer graphics. They should be submitted as separate files in JPEG (jpg) format and not embedded in the text document. Lettering on figures should be equivalent to a Times New Roman 9 pt font or slightly larger (up to 12 pt) if desired. Lettering smaller than 9 pt is not acceptable. Hand-drafted figures will not be accepted.

- Mathematical expressions: Equations should be presented in a clear form which can be easily read by non-mathematicians. Symbols should be typed using the Times New Roman symbol set. Equations should be in italics. Each equation should appear on a separate line and be numbered consecutively.

- References: References should be written in the Harvard system. The format of text citations should be as follows: "Jones (1999) discovered that ..." or "recent results (Brown & Carter 1985; Green et al 1999) indicated that ...". References cited in the text should be listed in alphabetical order at the end of the paper. References by the same author should be in chronological order. The following are examples of a journal article, a book and a conference paper:


- Manuscript: Manuscripts should be submitted to the SAICE Journal Editorial Panel, either electronically as a word document in PDF format for verification before publication.