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

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

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Impact of the Gautrain on property development around station precincts</td>
<td>R Mushongahande, C E Cloete, C J Venter</td>
</tr>
<tr>
<td>11</td>
<td>Factors predicting the intention to accept treated wastewater reuse for non-potable uses amongst domestic and non-domestic respondents</td>
<td>J R Adewumi, A A Ilemobade, J E van Zyl</td>
</tr>
<tr>
<td>20</td>
<td>The management of constructability knowledge in the building industry through lessons learnt programmes</td>
<td>V Kuo, J A Wiwm</td>
</tr>
<tr>
<td>28</td>
<td>Evaluation of the strength behaviour of unpaved road material treated with electrochemical-based non-traditional soil stabilisation additives</td>
<td>R J Moloi, A T Visser</td>
</tr>
<tr>
<td>40</td>
<td>The status of basic design ground motion provisions in seismic design codes of sub-Saharan African countries: A critical review</td>
<td>A Worku</td>
</tr>
<tr>
<td>54</td>
<td>Soil-structure-interaction provisions: A potential tool to consider for economical seismic design of buildings?</td>
<td>A Worku</td>
</tr>
<tr>
<td>63</td>
<td>Proposed guideline for modelling water demand by suburb</td>
<td>M L Griffioen, J E van Zyl</td>
</tr>
<tr>
<td>69</td>
<td>Accounting for moment-rotation behaviour of connections in portal frames</td>
<td>H L Albertyn, T N Haas, P E Dunaiski</td>
</tr>
<tr>
<td>77</td>
<td>The application and interpretation of linear finite element analysis results in the design and detailing of hogging moment regions in reinforced concrete flat plates</td>
<td>S A Skorpen, N W Dekker</td>
</tr>
<tr>
<td>93</td>
<td>Towards a systems thinking approach in allocating infrastructure budgets in local government</td>
<td>G N Kaiser, J J Smallwood</td>
</tr>
<tr>
<td>100</td>
<td>Analysing delay and queue length using microscopic simulation for the unconventional intersection design Superstreet</td>
<td>H H Naghawi, W J A Idewu</td>
</tr>
<tr>
<td>108</td>
<td>Discussion: Cradle-to-gate environmental impacts of the concrete industry in South Africa</td>
<td>Comment by Dr Frank Netterberg</td>
</tr>
<tr>
<td>109</td>
<td>Discussion: Structurally efficient housing incorporating natural forms</td>
<td>Comment by Dr Theo van Robbroek</td>
</tr>
</tbody>
</table>
INTRODUCTION

Transit-Oriented Development (also referred to as TOD) is a potentially useful concept around which to promote sustainable development. It seeks to increase accessibility through public transport, and promotes the human mobility approach to transport that encourages planning for non-motorised transport forms where people can walk and cycle more than drive (Department of Housing 2000).

Property development plays a crucial role in achieving the objectives for successful TODs. Marx et al (2006) acknowledged that TOD cannot happen without the involvement and commitment of property developers. Utter in Curtis et al (2009) commented that, while the public sector invests in transit systems and station area plans, it actually falls to private sector developers to implement and build these vibrant TOD areas. Similarly Cervero et al (2004) pointed out that property developers “... occupy the front lines of TOD, organizing the financial, physical, and human resources needed to build projects around transit stations.” However, while property developers involved in TOD neighbourhoods acknowledge TOD as a smart investment in congested and built-out urban areas, their hope is to reduce development costs, increase income and enhance profitability (Dittmar & Ohland 2004).

The Gautrain Rapid Rail project has been positioned as a potential catalyst for TOD in Gauteng (Du Plessis 2010), perhaps showing the way for sustainable urban development elsewhere in South Africa. Anecdotal evidence suggests that developer interest in some station precincts – notably Sandton – has increased sharply in response to the station development. Gautrain reports, for instance, that property transactions within two kilometres of its stations have grown much faster than further away from the stations (Gautrain 2008). In light of the existing widespread advocacy for TOD, the newness of the Gautrain project in South Africa, and the subsequent implementation of the TOD concept at station areas, this paper investigates the impact of the TOD concept on property development in the case of the Gautrain. The paper evaluates the spatial TOD elements implemented at a number of stations, and then evaluates the property development projects undertaken in line with the TOD elements. We seek to understand which elements of TOD – if any – are found attractive enough by local property developers to alter their investment decisions. Even though Gautrain itself has been operational for barely two years (at the time of writing), its planning process has lasted a decade – enough time, we argue, for some early trends in property response to emerge. The paper will be of use to local government, transport agencies and metropolitan municipalities in formulating development.
plans that encourage property development for successful TODs.

**TOD PRINCIPLES**

Cervero (2009, in Curtis et al 2009) pointed out that TOD “… is a straightforward concept: concentrate a mix of moderately dense and pedestrian-friendly development around transit stations to promote transit riding, increased walk and bicycle travel, and other alternatives to the use of private cars.” Four spatial principles characterise TOD worldwide:

**TOD Principle 1: Mix of land uses**
The mix of uses in TOD is crucial in facilitating ridership of public transport, as different uses tend to generate different trip volumes at different times of the day. International studies have shown that residential uses, with the correct density, are key in facilitating trip generation; retail uses generate ten times more trips per unit than office use; office uses induce higher public transport use during peak periods; while industrial uses have a limited impact on public transport use (The Gautrain Project 2011). Cervero et al (2004) commented that industrial uses are generally discouraged in TODs, although not always.

**TOD Principle 2: Availability of reliable public transportation systems**

There should be a reliable public transportation system that links the TOD neighbourhood and the broader region, and meets the demands of the travellers. Newman & Kenworthy (2007) noted that the key to move towards sustainability is to provide transport modes that have a competitive speed advantage for long trips through transit and for short trips through walking and biking.

**TOD Principle 3: Walkability**

Walkable distances, pedestrian-friendly infrastructure, and attractive and safe landscaping are essential to enhance access to public transport stations by users of properties within the neighbourhood and beyond.

**TOD Principle 4: High-density compact development**

Viable public transportation systems require fairly high levels of density which can sustain their operations and improve affordability.

**PROPERTY DEVELOPMENT FUNDAMENTALS**

Wilkinson, Reed & Cadman (2008) defined property development as: “... a process that involves changing and intensifying the use of land to produce buildings for occupation. It is not the buying and selling of land for a profit; land is only one of the raw materials used. Others include the building materials, infrastructure, labour, finance and professional services.”

Property development can be brownfields where existing buildings are redeveloped, or greenfields where new buildings are erected. It can also be distinguished in terms of the various property uses, i.e. residential, commercial (which includes offices and shops) and industrial; as well as various property classes: office property grades, for example, range across Grades P, A, B and C, where P denotes premium with top quality modern space and C denotes the lowest quality space (Cloete 1999).

Key drivers of property development include inter alia:

- property market demand
- availability of land in a good location
- design of buildings
- availability of finance
- management of completed buildings, and
- profitability of the project.

The interplay of these property development fundamentals influences the decision of a developer to engage in a property development project.

**TOD IMPACT ON PROPERTY DEVELOPMENT FUNDAMENTALS**

The four essential principles of TOD can, according to the literature, appeal to the fundamentals of the property development process in several ways. Each TOD principle is discussed in relation to its potential impacts on property development.

**TOD Principle 1: Mix of land uses**

*Availability of land in a good location*

The biggest challenge in implementing mixed-use projects in TODs is the difficulty in assembling a sufficiently large piece of land. Utter in Curtis et al (2009) commented that a large piece of land is required to accommodate an interesting mix of uses as well as the layout plan of the neighbourhood as a whole. Greenfields station areas have the benefit of the availability of land to allow the development of ideal TOD type of properties. This is unlike brownfields station areas where development of TOD type of properties has to be done by refurbishment or demolition and subsequent redevelopment of existing properties; this might not be an easy decision for a property developer to make, in light of other crucial real estate fundamentals.

*Good design*

One of the challenges with mixed-use projects is the integration and coordination of the many uses that should be incorporated within the project in order to maintain its mixed-use nature. For instance, Fruitvale transit village in Oakland, California, experienced a problem of lack of adequate access to the retail area, as commuters could park their vehicles in the park-and-ride lots and then use the transit system without passing the retail area (Dittmar & Ohlson 2004).

*Availability of finance*

Mixed uses incorporated vertically in a property pose challenges in terms of getting funding from financial institutions, who perceive the risks involved in property market cycles for different uses. Most financial institutions provide funding using common standards applicable to all properties, making it irrelevant that a property is in a TOD neighbourhood. In a survey undertaken by Cervero et al (2004) one lender commented that TOD seemed to be largely an irrelevant concept for lenders, as far as financing issues were concerned. Instead the focus was on regarding the TOD property developments simply as mixed-use projects that have the added bonus of being near a transit stop.

**TOD Principle 2: Availability of reliable public transport systems**

*Profitability of property*

Property developers aim to improve profitability of the property by investigating all the possible ways of keeping costs to a minimum while maximising income that can be derived from the property. Wilkinson & Marks (2007) pointed out that expediting the approval process for TOD projects facilitates property development as it helps reduce the developer’s holding costs. The land costs, design of a project, cost of finance and repayment period are crucial aspects that influence project costs. McLinden (2006) commented that construction costs in TODs are high, because construction must be done in a manner that minimises interruption of traffic at infill or redevelopment sites near busy intersections. In addition, developers are required to provide additional services such as park-and-ride facilities and bicycle paths. The incorporation of transit elements is reported by one developer in Pittsburg to have added 10% to 12% to the normal cost of development (McLinden 2006). Development costs are also influenced by an increase in property taxes as a result of increasing...
property value. By way of example, property taxes in Jersey City, Essex Street Corridor, which were estimated at $200 000 to $300 000 per annum before the rail line’s completion, increased to between $4 million to $6 million a year after the opening of the rail line in 2004 (Marx et al 2006).

Even though there are high costs associated with developing property in TODs, there is a potential rental premium that can be achieved from properties in TOD neighbourhoods. This rental premium is due to location efficiency of properties in TODs caused by the synergy of all the elements of TOD. Research findings from literature show that properties that are close to transit stations tend to experience increased values ranging from 6% to 24% for residential properties, and up to 50% for office and retail properties in American cities (Cervero et al 2004).

Staley (2009) argues that the relationship between property investment near rail stations and transit ridership is not direct; in other words, any increases in the values of properties in TOD are not entirely based on a property’s proximity to transit. The underlying causes of property value premiums include public safety, access to jobs, quality housing, tax rates, financing and zoning. Access to transit systems is probably further down the list than these other factors, and so transit ridership alone does not drive property development decisions around light rail stations. It should not therefore be presumed that simply improving access to mass transit will push property development in a significant way. Staley went on to conclude that policies should emphasise transit as an amenity that local residents and businesses require rather than as a driver of market demand and value. A study by Deakin & Mejias in 2005 of property development activity along San Pablo Avenue in Oakland supports Staley’s argument, as it found that developers view transit availability as a bonus, but not necessarily a major development incentive (Blume 2008). McLinden (2006) also supported Staley’s argument when he quoted an architect as saying that “... transit alone is not the secret ingredient to a successful development. Real estate fundamentals have to be there at a transit stop ...”

Nevertheless, it is crucial to note that the type, quality and route of public transport technology influence the property value premiums. TODs served by heavy rail transit tend to experience higher premiums because of higher speeds, more frequent services and wider coverage, and architectural integration. For light rail systems, property value premiums tend to be smaller than for heavy rail. In some instances residential properties within 900 feet of stations actually sell for less because of the transit system’s “nuisance effect.” Value premiums are also influenced by the type of neighbourhood. Proximity to transit tends to increase value for lower-income neighbourhoods, while it may reduce value for high-income neighbourhoods. Property values in TOD neighbourhoods are also enhanced by the availability of developable space. It can therefore be concluded that, although construction costs are high in TODs, the subsequent increased value of property facilitates enhanced profitability for successful property development in TODs.

**Good market demand for property in TODs**

There are many demographic and cultural changes in America that indicate that TOD will be a long-lived trend (Utter in Curtis et al 2009). These indications include ageing of the population, immigration and race, climate change, healthy communities, as well as rising transportation costs. Traffic congestion creates an increasing willingness to pay a premium for housing near rail stations, even if this means living in smaller houses on smaller stands. In South Africa, there is as yet little evidence of this, except for anecdotal indications of a growing demand for urban living.

**TOD Principle 3: Walkability**

**Good design**

The need to provide walkable distances in TODs influences the design elements of property. Designs that provide more pedestrian space on the premises through the use of arcades, atriums and open courtyards are effective in encouraging people to walk (Anderson-Watters 2009). Pedestrians are sensitive to the safety and attractiveness of the route they follow, and normally choose the shortest and most convenient route. In Clarendon in America, sidewalks encouraged pedestrian activity by creating a sense of safety and vitality (Dittmar & Ohland 2004). The differences in property design principles for TOD neighbourhoods and automobile-based neighbourhoods are shown in Figure 1.

**Proper management of property**

A proper management plan should be put in place for a TOD neighbourhood as a whole in order to maintain its characteristics and value. The use of the City Improvement Districts (CIDs) concept goes a long way towards promoting the use of non-motorised modes of transport (Downs 1991). Proper maintenance of the neighbourhood pavements, courtyards and buildings also promotes walking. In Emerson Park in America, a lack of proper maintenance of public infrastructure, open spaces and houses led to the deterioration of the residential units and the neighbourhood as a whole (Dittmar & Ohland 2004). Proper marketing plans of a neighbourhood also attract people to that neighbourhood, and hence the use of walking facilities provided thereto. Jacobson & Forsyth (2008) commented that appropriate programming of events and festivals can improve usage, safety and a sense of place within a TOD neighbourhood.

**TOD Principle 4: High-density compact development**

**Availability of land in good location**

Land is a scarce resource and compact densities in TOD assist, significantly, the supply of developable land. The concept of compact densities is promoted by relaxing zoning and building regulations, assisting particularly in brownfields developments.

**Profitability of the project**

Compact densities allow property developers to reduce property development costs in terms of providing services. The construction
of high-rise semi-detached housing can also be cheaper, as there are shared costs among units – for example, the cost of shared walls is spread among the owners of the units, and even the installation of plumbing services is cheaper when units are clustered.

On the other hand, where development costs increase as a result of building in a TOD neighbourhood, developers have been found to benefit from increased density, as it helps them to offset the extra costs experienced in TOD projects by increasing the lettable space, from which rental income will be realised (McLinden 2006).

Availability of market demand for TOD
There is a strong market demand in the developed world for TOD compact densities and affordable housing that is attractive, liveable and interesting. Cervero et al. (2004) pointed out that the interest of private property developers in TOD is derived from an increasing market for transit-oriented living, working and shopping, particularly in big cities where traffic congestion is a problem. Public policies, such as those aimed at increasing affordable housing in TODs, also further enhance the market for TOD.

METHODOLOGY
The Gautrain stations can be classified into three distinct groups, according to the type of area (inner city, brownfields or greenfields) in which they were developed, as depicted in Table 1. The Pretoria, Midrand and Rosebank stations were chosen as typical representatives of stations in an inner city, greenfields or brownfields type of neighbourhood, respectively.

Study areas have been defined to be the areas within a 400 m radius of the selected Gautrain station (shown in Figures 2 to 4), in line with the typical primary catchment area radius of transit (Vuchic 2005). Our analysis aimed at identifying the extent to which private property development within this radius responded to the TOD concept during the planning and construction phases of the property.

Data was collected from archival records of the land development applications register and building plans register of the metropolitan municipalities of Johannesburg and Tshwane for the period 2000 to 2011. Aerial photographs for the same period were also utilised to assess the changes in property status over the years.

A questionnaire survey was also undertaken with developers in the study area in order to qualify the data provided by secondary data sources. The survey excluded all properties owned by the government and those owned by sectional title, as the aim of the study was to assess the views of private property developers. Table 2 summarises the size of the sample.

The questionnaires collected data on the timing of property acquisition and development in the station area, and reasons for the selection of the particular site. Respondents...
were also asked to rate the significance of the main TOD principles in motivating their choice of property development. Qualitative data was analysed using thematic content analysis, in order to get an understanding of the response of private property developers to the TOD concept. An interview was also conducted with an official from the Gautrain Management Agency in order to gain insight into the role of TOD principles during the development of the station precincts.

RESEARCH FINDINGS

TOD principles implemented at selected Gautrain station areas
The Gautrain Rapid Rail link has been regarded as a catalyst to create a new urban environment, in the form of TOD. The aim is to create long-term sustainability of the Rapid Rail Link by achieving optimal ridership. While the Gautrain bus feeder and distribution system was implemented to enlarge the catchment area of the rail service, true sustainability in the long run requires cultivation of ridership from within the immediate vicinity of the stations. To achieve this goal, the spatial principles of TOD need to be realised incrementally around station precincts. Gautrain seemed to be eminently aware of this.

In an interview with an official from the Gautrain Management Agency the importance of policies emerged as a strong catalyst for the development of TOD at the station precincts. He cited the preparation of local spatial development frameworks for each of the station precincts which inform the implementation of the much needed spatial principles of TOD. He further identified other policies, such as the Rosebank City Improvement District, the Dunkeld Residents Association for Rosebank, Intersite development proposals for the larger Pretoria Station precinct, and the Department of Public Works land assembly for the Pretoria Station precinct.

Pretoria Station
Types and mix of uses
Pretoria Station is situated on the edge of a central business district (CBD) which, although already formally established, includes many fairly old and dilapidated buildings. One of the aims of locating the station here has been the revitalisation of the declining CBD through the upgrading and improvement of the area, as well as the encouragement of business, residential and tourism facilities. The study area comprises 82 properties. The property type is identified by the use of the buildings as observed on the ground by the researcher and verified by the provisions of the town planning scheme for the area. Table 3 shows the types and mix of uses found around the Pretoria Station, indicating residential to be the dominant use.

Most of the residential properties are low-rise stand-alone houses. These constitute 43% of the residential units, while high-rise residential flats constitute 40% and vacant stands constitute 17%.

Hospitality properties in this area are primarily hotels. Parking premises and open spaces are municipal-owned. Mixed-use properties predominately comprise retail use on ground floors and residential flats on upper floors, which constitute 75% of mixed-use properties, while mixed-use properties comprising retail and offices constitute 25% of the mixed-use properties.

Density
The density of property as depicted by the height of the building is presented in Table 3. The tallest building has ten storeys and is residential high-rise. All the mixed-use properties in the precinct. Sidewalks and open spaces are already well provided, and Gautrain invested in improved sidewalks and crossings immediately adjacent to the station. Nevertheless, some existing sidewalks are in a poor state of repair, which negatively affects safety and the quality of the walking environment. The northern side of the neighbourhood comprises a terrain with a relatively steep gradient, which affects the ability to walk and cycle.

The availability of various modes of transport
Pretoria Station is a hub of public transportation systems comprising the Gautrain and its bus feeder system, Metrorail, long-distance buses, metered taxis and mini-bus taxis. Private vehicles are also used, as supported by the provision of parking space within the public open space and railway property.

<table>
<thead>
<tr>
<th>Station</th>
<th>Number of properties</th>
<th>Properties eligible for the questionnaire survey</th>
<th>Questionnaires sent out</th>
<th>Questionnaires completed</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretoria</td>
<td>82</td>
<td>56</td>
<td>44</td>
<td>31</td>
<td>55%</td>
</tr>
<tr>
<td>Midrand</td>
<td>32</td>
<td>13</td>
<td>8</td>
<td>5</td>
<td>38%</td>
</tr>
<tr>
<td>Rosebank</td>
<td>103</td>
<td>86</td>
<td>80</td>
<td>54</td>
<td>63%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Use</th>
<th>% properties in study area</th>
<th>Number of sites developed since 2000</th>
<th>Height of buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential</td>
<td>51</td>
<td>1</td>
<td>up to 10 storeys</td>
</tr>
<tr>
<td>Retail</td>
<td>23</td>
<td>–</td>
<td>1 storey</td>
</tr>
<tr>
<td>Office</td>
<td>4</td>
<td>–</td>
<td>1 to 5 storeys</td>
</tr>
<tr>
<td>Mixed-use</td>
<td>10</td>
<td>1</td>
<td>6 to 9 storeys</td>
</tr>
<tr>
<td>Hospitality</td>
<td>4</td>
<td>–</td>
<td>1 to 8 storeys</td>
</tr>
<tr>
<td>Educational</td>
<td>1</td>
<td>–</td>
<td>1 storey</td>
</tr>
<tr>
<td>Parking</td>
<td>1</td>
<td>1</td>
<td>ground level</td>
</tr>
<tr>
<td>South African Railways</td>
<td>2</td>
<td>1</td>
<td>1 storey</td>
</tr>
<tr>
<td>Open space</td>
<td>4</td>
<td>–</td>
<td>ground level</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
Midrand Station

Types and mix of uses
Midrand is a greenfields station that has been integrated with planning proposals rather than existing development. The location east of the N1 motorway, in a sparsely developed spot, was chosen because it links the emerging Midrand CBD, Gallagher Estate (a conference venue), various planning proposals of the municipality, and a very large proposed mixed-use development named Zonk’Izizwe. The larger Midrand area is slowly transitioning from a low-density rural to a more urban environment, but it generally lacks pedestrian and public transport facilities. The station was thus earmarked for its future potential rather than its current vibrancy.

The Midrand Station study area comprises 32 properties. The type of use dominating this area is business, followed by residential, agricultural, offices and open space in that order (see Table 4). According to the Consolidated Johannesburg Town Planning Scheme (Johannesburg 2011), business use referred to here could include retail, offices, parking, showrooms and restaurants.

Density
The tallest building in the neighbourhood is eight storeys high, while all residential flats are three storeys high. Twenty-four out of 32 properties are undeveloped. The Midrand Urban Development Framework (Johannesburg 2008b) provides the required height of buildings in the study area that can sustain Gautrain as follows: the tallest buildings of up to 13 storeys are expected for properties closest to the station, while seven-storey and ten-storey properties are expected for properties further away from the station, at the edge of the study area. In comparison with the height of existing properties, it is clear that there is still a gap between the existing and the desired height of property. This gap indicates the need for some property development projects to be undertaken to meet the required densities that can sustain the Gautrain.

Use of non-motorised transport
The station precinct has limited walkability due to its location on a steep gradient and right next to a major arterial. The K101 road tends to create a barrier to walking due to its width, high traffic speeds, and insufficient sidewalks and pedestrian crossings. Streets are not labelled, making way-finding more difficult.

The availability of various modes of transport
The modes of transport found in the Midrand Station neighbourhood include private cars, the Gautrain and its bus feeder system, as well as minibus taxis. There are two taxi ranks in close proximity to the study area.

Rosebank Station

Types and mix of uses
Of the three study sites, Rosebank already displayed most TOD principles before the advent of the Gautrain. The area around the station was a well established node comprising a mix of primarily retail and office uses, with a strong tourism component. The buildings in the study area were generally in good condition, except for a few that required refurbishment. The implementation of the TOD concept was thus aimed at revitalising an existing urban node, as well as at promoting a mix of uses, particularly residential densification. The retail core and office components need to be managed and sustained in order to attract passengers for the Gautrain.

The Rosebank study area comprises 103 properties and is characterised by a variety of land uses. Table 5 shows that the dominant use is office use, followed by residential use. Hospitality, mixed-use, retail, educational and parking uses were also noted. The dominance of office use has been boosted by recent construction, as well as by the conversion of some houses into offices, mostly in the Parkwood suburb along Bolton Road.

Residential properties are characterised by low-density and high-density residential apartments. Of all the residential properties, 79% are low-density stand-alone houses, 18% are high-rise apartments and 3% are vacant stands.

Although the statistics seem to show a low percentage of retail use, there is a lot of retail space in this node, which constitutes mixed-use. All the mixed-use properties tend to have a retail component. Similarly, office uses are also incorporated in all the mixed-use properties, while 33% of the mixed-use properties contain hospitality uses. The Rosebank Urban Development Framework (Johannesburg 2008a) describes the dominance of office space, as well as retail and hospitality, to be as a result of the character of Rosebank as a regional node within the Gauteng region.

Density
Table 5 shows that the tallest building in the area is a nine-storey mixed-use property comprising retail, offices and hospitality. The Rosebank Urban Development Framework (2008) provides the height of buildings desired in order to meet the required density and hence promote ridership of the train. The height of mixed-use properties is expected to be up to 20 storeys, offices up to two levels.

<table>
<thead>
<tr>
<th>Use</th>
<th>% properties in study area</th>
<th>Number of sites developed since 2000</th>
<th>Height of buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business</td>
<td>66</td>
<td>–</td>
<td>ground level, undeveloped</td>
</tr>
<tr>
<td>Offices</td>
<td>9</td>
<td>1</td>
<td>up to 8 storeys</td>
</tr>
<tr>
<td>Residential</td>
<td>13</td>
<td>4</td>
<td>3 storeys</td>
</tr>
<tr>
<td>Agricultural</td>
<td>9</td>
<td>–</td>
<td>ground level</td>
</tr>
<tr>
<td>Open space</td>
<td>3</td>
<td>–</td>
<td>ground level</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>5</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 Property uses and density in Midrand Station study area

<table>
<thead>
<tr>
<th>Use</th>
<th>% properties in study area</th>
<th>Number of sites developed since 2000</th>
<th>Height of buildings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Offices</td>
<td>40</td>
<td>6</td>
<td>1 to 8 storeys</td>
</tr>
<tr>
<td>Residential</td>
<td>38</td>
<td>8</td>
<td>1 to 7 storeys</td>
</tr>
<tr>
<td>Retail</td>
<td>3</td>
<td>3</td>
<td>1 storey</td>
</tr>
<tr>
<td>Parking</td>
<td>6</td>
<td>–</td>
<td>ground level</td>
</tr>
<tr>
<td>Mixed-use</td>
<td>9</td>
<td>3</td>
<td>3 to 9 storeys</td>
</tr>
<tr>
<td>Hospitality</td>
<td>6</td>
<td>1</td>
<td>2 to 3 storeys</td>
</tr>
<tr>
<td>Educational</td>
<td>1</td>
<td>–</td>
<td>2 storeys</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100</strong></td>
<td><strong>21</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 5 Property uses and density in Rosebank Station study area
15 storeys and residential up to six storeys. This, in comparison with the current height of buildings provided in Table 5, shows that there is need for some adjustments to the existing densities in order to meet the desired densities for the Rosebank study area. Significant redevelopment of the low-density residential component would have to take place to make this happen.

Use of non-motorised transport
The mix of uses in the Rosebank study area, both horizontally and vertically, promotes walking within the neighbourhood. The Rosebank business district is very pedestrian-friendly, as most of the streets have been pedestrianised and the station is well integrated with the mixed-use properties, forming the heart of the Rosebank neighbourhood. Where the streets have not been pedestrianised, appropriate sidewalks have been provided and intersections are signalised.

The availability of various modes of transport
Rosebank Gautrain Station neighbourhood is served by various modes of transport, including private cars, the Gautrain and its bus feeder services, Metrosbus, metered taxis, minibus taxis, as well as the proposed bus rapid transport. Bus and taxi routes are generally well integrated, and all facilities are within walking distance of the node.

Identification of properties developed since the announcement of Gautrain in 2000

Pretoria Station
The Building Application register reflects all the applications for Building Control received by the municipality over the period 2000 to 2011. It provides an indication of where development is occurring on the ground and where developers are still in the planning stages. The register shows that in the Pretoria precinct 13 building control applications were received since the announcement of the Gautrain project, of which 12 have a direct influence on new developments as they are applications for building plans approval and/or demolition of old buildings. Of these, five applications have been approved and, to date, the construction of four has commenced. Aerial photographs supplied by the Tshwane Municipality for the periods 2001, 2003, 2005, 2007, 2009 and 2010 confirmed that four properties were developed during that period. Apart from the railways-owned land used for the construction of bus and rail facilities, the developed land uses included residential, mixed-use and parking (Table 3).

Midrand Station
Six properties were developed in the Midrand study area since the announcement of the Gautrain project in 2000. Of these six properties, five are new developments that were built and completed (as confirmed by both aerial photographs and building application data provided by the Municipality, as well as an interview with the Gautrain Management Agency official) – four residential and one office building. The Town Planning Application System shows that one property – the mixed-use development of Zonk’izizwe – is at the planning stage, development rights already having been applied for.

Rosebank Station
Forty-seven properties went through some stages of property development, varying from planning the project to actual construction, since the announcement of the Gautrain project in 2000. This is approximately 44% of the properties in the Rosebank study area. Of the 47 properties, 68% made new development applications with the Municipality of Johannesburg in terms of the Town Planning Application System (TAS), representing the planning stage of property development. The remaining 32% were developed in line with the existing provisions of the scheme. The Building Application System also shows that 28 properties applied for building plans approval. Data derived from the researcher’s physical inspection, in conjunction with the analysis of aerial photographs, as well as the interview with the Transit Management Agency official, shows that 20 properties went through the construction phase and one property went through demolition to become a vacant stand. There is clearly a high level of activity in the Rosebank Station node.

Role of TOD in private property developers’ decision-making
Developers were asked to rate the importance of various TOD principles in their decision to invest in property within the station precincts (i.e. within 400 m of the Gautrain stations). Figure 5 shows the results. Results varied slightly between stations, suggesting that different considerations drive property development in different environments. In all three areas the availability of the Gautrain rated very highly, indicating that developers are aware of the potential of the rail system to improve the viability and profitability of their properties. Gautrain rated more highly in the more decaying (in the case of Pretoria) and more emerging (Midrand) nodes, suggesting that developers buy into the authorities’ vision of leveraging nodal (re)generation through investment in the train system. In Rosebank, where many TOD elements were already present, Gautrain is considered equally important to the ability to achieve high densities and mixed-land uses, and supportive policies and incentives.

The importance of supportive policies and incentives was echoed during the interviews. Property developers’ recommendations centred on the need to reduce the time of development and building approvals, the provision of incentives that can reduce the cost of construction for developers, and the possibility of specific financing for TOD property development. One important recommendation that was brought up by individual and small developer firms concerned the need to advertise the Gautrain more, to overcome the problem of lack of knowledge about the train.

![Figure 5 Importance rating of TOD principles in private developers’ decisions to invest in Gautrain station precincts (n = 74)](image-url)
CONCLUSIONS

Extent of TOD principles implemented at selected Gautrain stations

Literature identified a TOD neighbourhood as one that can foster sustainability by creating walkable, mixed-use, compact and vibrant places around high-quality public transport nodes, especially in automobile-dependent societies. Similarly this research identified four key TOD principles informed by spatial planning in the implementation of the TOD concept at the Pretoria, Midrand and Rosebank Gautrain stations. The principles are a mix of various land uses, higher density in terms of building heights, ability to walk and cycle in the neighbourhood, and the availability of various transport modes.

The research found that the amount of development that has taken place towards the TOD concept varies significantly across the three stations. There is evidence of accelerated property development and increasing mixed use at all three sites, but activity is by far highest in Rosebank where elements of TOD design – notably well-integrated mixed-use environments, pedestrianised spaces of high quality, and modal integration – are most present.

The preparation of local spatial development frameworks for each of the station precincts emerged as an essential commitment by the authorities to foster TOD, which could in turn motivate property developers to engage in building in the TODs. This research has shown, however, that none of the three stations has as yet fully met the spatial TOD principles provided by these frameworks. This confirms the findings of literature that TOD neighbourhood development takes a long time, more than a decade, to develop.

The Pretoria Station neighbourhood, being a downtown type of TOD in the CBD of Pretoria, has maintained the mix of uses and densities that existed before the Gautrain Rapid Rail Link project. The Midrand Station neighbourhood is a greenfields type of TOD neighbourhood where, literature shows, implementing a TOD can be fairly easy. However, the Midrand Station neighbourhood has not started to reflect the implementation of the provisions of the Station Development Framework. Property development in this neighbourhood is slow, as most of the land around the station is not developed and the large tract of land surrounding the station is owned by one entity, while most of the land west of the K101 is owned by another. Significant investment is needed to turn this largely vacant, car-oriented area into a TOD-supportive neighbourhood. This illustrates the risks involved in attempting to implement TOD in an area belonging to only one or two property developers. It seems easy for the property development process to be out of sync with the transport investment.

The study also demonstrates the importance of promoting residential land uses in TOD nodes, as these result in increased population density that improves ridership (and reverse ridership share). In both the Pretoria Station and Rosebank Station neighbourhoods there are substantial components of residential use, and some further residential development has occurred in recent years. Yet it seems most of this development is either of a low-cost or a low-density quality, which is not conducive to TOD success.

Rate of property development since the announcement of the Gautrain in 2000

Although the rate of property development varies significantly across the case study areas, there is evidence that the TOD concept does attract new development in the vicinity of Gautrain stations. Property developers consider the presence of the Gautrain to be a major factor influencing their decision to develop in these areas. The Gautrain is especially important in the less developed, less attractive nodes of Midrand and Pretoria. This suggests that, in line with international experience, the benefits offered by a high-quality public transport system in terms of permanence, nodal regeneration potential, and securing property demand, are also present locally.

What seems equally clear, though, is that there are real estate fundamentals other than the TOD principles that drive property development and these will ultimately determine the success of a TOD undertaking. When comparing the vibrant Rosebank with the slower-growing Pretoria and Midrand precincts, a number of factors seem to contribute to its success: (1) prior successful mixed-use properties that demonstrated market demand at this location, (2) good design, (3) land available for development, and (4) the “location, location, location” factor of Rosebank as a regional node in Gauteng. It seems likely that the strong real estate fundamentals already present were further strengthened by the proximity and availability of the Gautrain Station.

CONCLUDING REMARKS

The outcomes of the research have generally tended to match the findings of the surveyed literature on the impact of TOD on property development. TODs in the western world, wherever implemented, take a long time to develop. In the South African context it has been eleven years since the announcement of the Gautrain project; yet, on the ground, the pace of change is slow and, in some cases, non-existent.

There might be merit in recognising that not all stations are good candidates for TOD – whether because of their location, development environment, or intermodal potential. It is important to identify in advance those with highest potential to enable energies to be focused where they will most effectively achieve the synergies needed to make TOD work.

In conclusion, it is possible that the slow manifestation of the TOD principles at some Gautrain stations inhibited private property developers’ interest in investing in station areas. Developers need to see the commitment of government and the public transport agency in creating the TOD neighbourhoods as already applied at successful transit villages in the western world.

ACKNOWLEDGEMENTS

This research was financially supported by the Department of Transport through the Sustainable Public Transport and Sport Project, implemented by the United Nations Development Programme (UNDP), and funded by the Global Environment Facility (GEF).

REFERENCES


Gautrain Management Agency 2008. Gautrain is more than just a train. *Mobility magazine*, June/August.


Factors predicting the intention to accept treated wastewater reuse for non-potable uses amongst domestic and non-domestic respondents

J R Adewumi, A A Ilemobade, JE van Zyl

Water reuse projects can fail if the factors (social, technical, financial, etc) governing their implementation are not adequately addressed prior to implementation. This paper presents the findings of the analysis of questionnaires administered to potential domestic and non-domestic consumers in Capricon and Vhembe (Limpopo Province) where wastewater reuse was being considered. The analysis examined the factors considered to predict the intention to accept wastewater reuse prior to implementation. Structural Equation Modelling (SEM) was employed to test several hypotheses addressing respondents’ intention to accept wastewater reuse for non-potable end uses. In other words, SEM tested whether certain factors (e.g. trust, attitudes and control) measured intention. Intention was measured as a second order factor. For non-domestic respondents, the factors that predicted intention to accept wastewater reuse in order of significance were their attitude towards wastewater reuse, the degree of control over the source of water and its application, the respondents’ knowledge of the advantages of reuse and the respondents’ trust in the service provider. For domestic respondents, the factors were their knowledge of the advantages of reuse, the degree of control over the source of water and its application, attitude towards wastewater reuse, trust in the service provider, and the subjective norms of the respondents. Physical quality satisfaction (for both respondent categories) and subjective norms (for non-domestic respondents only) could not be assessed because a reliable scale was not formed. The above findings have implications for future wastewater reuse in South Africa, i.e. that decision-makers contemplating reuse for non-potable uses would profit from addressing the various factors predicting intention to accept reuse prior to implementation.

INTRODUCTION
The increasing demand for water in South Africa is driven by growing populations, the connection of previously un-served households to municipal water supply, growing industrial development, urban in-migration and a host of other factors. Consequently, water resources planners are continually looking for additional sources of water to supplement the limited resources available (Adewumi et al 2010). It was predicted that unless the water consumption patterns in South Africa change significantly, the country would not be able to meet the growing demand for water, and the country would be extremely severe within 20–25 years (DWAF 2007). As shortages increase, allocation of water to irrigated agriculture, for example, may result in downstream urban areas facing water shortages, leading to water use restrictions and increased general discontent. It is thus within the context of freshwater constraints that the South African government is faced with the challenge of implementing sustainable alternatives, including wastewater reuse, for potable and/or non-potable requirements.

Wastewater reuse has become an attractive option for conserving and extending available water supplies. Other benefits of reuse include the decrease in the diversion of freshwater from sensitive ecosystems, replenishment of soil nutrients in agriculture due to irrigation, enhancement of groundwater recharge and delay in the future expansion of water supply infrastructure (Angelakis & Bontoux 2001; Joksimovic 2006).

Despite the benefits mentioned above, reuse should not be contemplated where there is non-compliance in treated wastewater effluent quality, crops to be irrigated have not been proven to be tolerant to the salts in the effluent, and there is no risk of salts from the effluent resulting in the deterioration of
ground/surface water quality. Another major challenge affecting the implementation of reuse projects is public opinion. Opinion influences intention to accept, and intention influences behaviour. It is for this reason that community opinion towards wastewater reuse has been identified as a key component of reuse project success (Okun 2002; Po et al 2004). The majority of opinion surveys on water reuse that have been published in the literature have emanated from the USA, Australia, Western Europe and the Middle East. For the purpose of forming appropriate policy and strategy, and due to the large variations in culture, climate, water availability, economy, etc, opinion studies need to be developed or adapted for national and sometimes sub-national contexts (Friedler et al 2006). In view therefore of the significant impact public opinion has on water reuse, it is important that opinion be investigated prior to the implementation of wastewater reuse.

Several studies (including Nancarrow et al 2008; Nancarrow et al 2009; Nancarrow et al 2010; and Po et al 2005) suggest that public acceptance of reuse is a product of attitude, emotion, control over source of water, subjective norms (influence of people around you), knowledge of the scheme, associated risks, trust in the implementing authority, physical quality satisfaction, choice, specific use, source(s) of recycled water, cost, water scarcity and socio-demographic factors. Individually, or in combination, these factors have been investigated in various places where water reuse schemes have been implemented or are planned. For example, Wilson and Pfaff (2008) carried out research in Durban, South Africa, (and compared their findings with international experiences) to determine if there were groups with specific religious or philosophical objections to potable reuse. They concluded that fundamental religious objections to potable wastewater reuse do not exist internationally and locally, but that people are generally not comfortable with the idea of potable reuse. Non-potable reuse of wastewater, on the other hand, is expected to be less uncomfortable, especially if this reuse involves minimum human contact (e.g. toilet flushing and irrigation). There is, however, currently no empirical research in South Africa confirming or debunking this supposition.

The research presented herein thus seeks to investigate the underlying factors that predict domestic and non-domestic respondents’ intention to accept wastewater reuse for non-potable purposes. Domestic and non-domestic respondents are distinct water users, and are therefore investigated as separate respondents in this study. The investigation is achieved by employing the Structural Equation Model (SEM) to analyse the factors that are considered to predict intention to accept wastewater reuse for non-potable water requirements in two arid South African communities in Limpopo (Seshego, Isisu and Ext 44 within the Capricorn District Municipality and the Vhembe District Municipality respectively) where wastewater reuse is being considered.

**BACKGROUND AND RESEARCH THEORY**

This paper investigates the empirical relationships between trust in the reuse implementing authority, knowledge of the advantages of reuse, physical quality satisfaction, perceived behavioural control, subjective norms and attitude, as proposed by Ajzen (1985) and Po et al (2005), and the intention to accept wastewater reuse amongst potential non-domestic and domestic respondents in South Africa.

According to the Theory of Planned Behaviour (TPB) proposed by Ajzen (1985), an individual’s behaviour is determined by the person’s intention to engage in the behaviour. Intentions are in turn predicated on three factors (also known as constructs, belief based measures or latent/unknown variables), i.e. attitudes, subjective norms, and perceived behavioural control. Definitions for these constructs are provided below:

- **Attitude** is a complex mental state involving feelings, values and disposition to act in certain ways. It measures overall positive or negative predisposition to behave in a certain way.
- **Subjective norms** are the beliefs about the normative expectations of others, and to comply with these expectations (also called normative beliefs). It measures the perception of how important people in the life of respondents would approve or disapprove of their performing a particular behaviour (i.e. social pressure).
- **Perceived behavioural control** is the belief about the presence of other factors that may facilitate or impede performance of the behaviour, and the perceived power of these factors (control beliefs). It measures the extent to which an individual has the capacity to perform the behaviour.

Based on the research conducted by Eiser et al (2002) which suggested the inclusion of the following constructs to Ajzen’s (1985) TPB, i.e. (i) perceived risks and benefits, (ii) knowledge of the advantages of reuse, and (iii) trust in implementing authorities, experts and technology, Po et al (2005) proposed a hypothesised model comprising various constructs that influence the acceptance or rejection of recycled water for various...
uses. A derivative of this hypothesised model is presented in Figure 1. The six constructs shown in Figure 1 are hypothetical and therefore cannot be directly observed, but must instead be inferred from respondents’ responses to questions/statements that statistically correlate with the constructs.

**Hypotheses used in this study**

In this study, the model predicting intention to accept or reject wastewater reuse was tied to a series of hypotheses discussed below:

**Knowledge of the advantages of reuse:** The knowledge of the advantages of reuse has not been tested in the context of wastewater reuse for non-potable water requirements. This hypothesis postulates that if respondents have good knowledge of the advantages of wastewater reuse, this knowledge would enhance their intention to accept wastewater reuse for non-potable uses. Hence, the following hypothesis:

**H1: Respondents’ knowledge of the advantages of wastewater reuse has a positive effect on intention to accept wastewater reuse for non-potable water requirements.**

**Trust in the implementing authority:** Prior research in Australia (Po et al 2005; Fielding et al 2009) identified trust in the Water Authority as a major determinant of the acceptance of recycled water. Also, a study conducted by Lin and Wang (2006) showed that trust had a positive effect on customers’ loyalty and consumers’ satisfaction. Eiser et al (2002), however, found that trust had a weak influence on consumers’ food satisfaction. The following hypothesis was therefore developed for this study:

**H2: Respondents’ trust in the treated wastewater service provider has a positive effect on intention to accept wastewater reuse for non-potable water requirements.**

**Subjective norms:** The term subjective norms is closely related to social pressure. It measures the perception of how important people in the life of respondents would approve or disapprove of their performing a particular behaviour. Subjective norms have been found to affect knowledge sharing intentions among groups (Ruy et al 2003) and among senior managers (Lin & Lee 2004). Fielding et al (2009) and Po et al (2005) also reported that subjective norms significantly affected intention to accept recycled water. In this study, subjective norms about recycled water refers to how social pressure affects the intention of respondents to accept recycled water for non-potable water requirements. Hence, the following hypothesis:

**H5: Higher subjective norms associated with wastewater reuse for non-potable water requirements have a positive effect on respondents’ intention to accept wastewater reuse.**

**Aesthetic appearance:** Hurlimann and McKay (2007) found out that the colour of recycled water was the most important attribute for consumers to accept recycled water for washing clothes. The following hypothesis was therefore formulated:

**H6: The aesthetically pleasing appearance of recycled wastewater will have a positive effect on respondents’ intention to accept wastewater reuse for non-potable water requirements.**

**RESEARCH METHODOLOGY**

**Questionnaire structure**

Two questionnaires were developed and administered to a random sample of potential non-domestic (i.e. agricultural, commercial, educational and parks) and domestic non-potable water consumers. Non-domestic respondents were individuals representing their various institutions, while domestic respondents were representatives of various households. The questionnaire was subdivided into three parts: introduction, perceptions and respondents’ personal data (domestic respondents only). The introductory part of the questionnaire clearly stated the aims of the project, which was to determine perceptions on the use of treated wastewater for non-potable purposes and the willingness to use dual water distribution systems. A concise definition of non-potable water was provided. The second section comprised statements (developed to test hypotheses H1–H6) aimed at measuring respondents’ positive or negative perceptions towards wastewater reuse. Justifications for the statements used

### Table 1: Constructs measuring intention to accept wastewater reuse for non-potable water requirements and their respective hypotheses

<table>
<thead>
<tr>
<th>Construct</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages of reuse (ADV)</td>
<td>H1: Respondents’ knowledge of the advantages of wastewater reuse has a positive effect on intention to accept wastewater reuse for non-potable water requirements.</td>
</tr>
<tr>
<td>Trust (TRU)</td>
<td>H2: Respondents’ trust in the treated wastewater service provider has a positive effect on intention to accept wastewater reuse for non-potable water requirements.</td>
</tr>
<tr>
<td>Attitude (ATT)</td>
<td>H3: Respondents’ positive attitude towards wastewater reuse will increase the intention to accept wastewater reuse for non-potable water requirements.</td>
</tr>
<tr>
<td>Control over source of water (CON)</td>
<td>H4: Respondents’ perceived control over the source of water and its application has a positive effect on the intention to accept wastewater reuse for non-potable water requirements.</td>
</tr>
<tr>
<td>Subjective norms (SNO)</td>
<td>H5: Higher subjective norms associated with wastewater for non-potable water requirements, has a positive effect on intention to accept wastewater reuse.</td>
</tr>
<tr>
<td>Physical quality satisfaction (PQS)</td>
<td>H6: The aesthetically pleasing appearance of recycled wastewater will have a positive effect on respondents’ intention to accept wastewater reuse for non-potable water requirements.</td>
</tr>
</tbody>
</table>
in this section are provided below. The last section of the domestic questionnaire consisted of questions requiring demographic information such as gender, age, racial group, marital status, and academic qualification. Type of house and approximate monthly income data were requested from only domestic respondents.

**Justification for the statements used to measure each construct**

**Advantages of reuse (ADV):** The sentences below attempt to determine respondents’ knowledge of the advantages of wastewater reuse and how this knowledge influences intention to accept reuse for non-potable water requirements:

i. ADV1 – “The use of non-drinking water will reduce the amount of wastewater discharged to the environment”: Treated wastewater is typically discharged into the environment (especially surface waters). By reusing treated wastewater therefore, the total volume of wastewater discharged into the environment is reduced.

ii. ADV2 – “Non-drinking water use will reduce the depletion of groundwater and surface water resources”: With increased wastewater reuse for non-potable water requirements, less fresh water is likely to be extracted from surface and ground waters. A reduction in the extraction of ground waters will reduce saline water intrusion in coastal areas, while a reduction in the extraction of surface waters will assist to maintain minimum environmental flows and mitigate negative and often irreversible ecosystem changes.

iii. ADV3 – “The use of non-drinking water can save many South African communities from drought”: The supply of treated wastewater for non-potable water requirements will mitigate the negative effects of a drought.

iv. ADV4 – “There can be considerable savings of fertiliser on farms irrigated with recycled wastewater”: Nutrients in wastewater, such as nitrogen and phosphorus, nourish irrigated soil and consequently provide nourishment for plants, and therefore can readily replace organic fertilisers. Reuse should not be contemplated where plants or crops to be irrigated have not been proven to be tolerant to the salts in the effluent, and there is no risk of salts from the effluent resulting in the deterioration of ground/surface water quality.

**Trust in implementing authorities (TRU):** A respondent’s trust in the provider of reclaimed water is measured by their responses to the statements below. The higher the level of trust, the more likely they will accept wastewater reuse. These statements assess trust in relation to the quality of the product or service rendered by the provider as perceived by the respondent. The statements include:

i. an overall statement TRU1 – “This institution (I) will use non-drinking water if the quality can be proven to be satisfactory”; and the following specific statements related to:
   - ii. disgust due to odour, colour, and suspended solids, TRU2 – “This institution (I) will use non-drinking water if it is not disgusting or irritating”;
   - iii. cloth-staining potential for domestic respondents, TRU3 – “This institution (I) will use non-drinking water if it does not stain washing”; and
   - iv. public health and safety, TRU4 – “This institution (I) trusts the municipality to provide non-drinking water that is safe and does not constitute a health risk.”

If reuse is implemented, reuse regulation and/or certification (such as the Green Drop Certification (DWAF 2009)) will likely improve respondents’ trust in a service provider.

**Attitude (ATT):** The statements below attempt to assess a respondent’s attitude towards wastewater reuse in the following ways:

i. The respondent’s social obligation towards water which is a renewable, albeit finite and often abused resource, ATT1 – “This institution is (I feel personally) obligated to do whatever I can to save water,” and ATT2 – “Water is a valuable resource that should be recycled.”

ii. The respondent’s choice/preference with regard to wastewater reuse, ATT3 – “This institution/I would prefer not to use non-drinking water,” ATT4 – “This institution/I would not use non-drinking water even in times of water shortages,” and ATT5 (for non-domestic respondents) – “This institution would only be prepared to use non-drinking water in times of water shortages.”

iii. The respondent’s willingness to be part of the solution and not easily apportion blame to government, ATT5 (for domestic respondents and ATT6 for non-domestic respondents) – “The government is responsible for water shortages.”

**Control over source of water (CON):** The sentences below assess how acceptance of reuse is influenced by a respondent’s perceived control over how water/wastewater is used/reused and how wastewater irrigated products are presented. The higher the perceived control, the more likely reuse will be accepted.

i. CON1 (for non-domestic respondents only) – “Every household should be free to choose its source of water supply (e.g. groundwater, surface water and recycled wastewater).”

ii. CON1 (for domestic respondents) – “I have the right to know if fruits or vegetables are irrigated with recycled wastewater.”

iii. CON2 – “Fruits and vegetables irrigated with non-drinking water (e.g. recycled wastewater) should be labelled in the supermarket.”

iv. CON3 (for domestic respondents) – “I have the right to adequate drinking water supply.”

**Subjective norms (SNO):** The influence of others on a respondent’s acceptance of wastewater reuse is measured by the following statements:

i. SNO1 – “This institution (I) will use non-drinking water if other institutions (others) are using it.”

ii. SNO2 – “Most institutions (people) who are close to our institution (me) will support the use of non-drinking water.”

iii. SNO3 (for domestic respondents only) – “Non-drinking water use is an option for the poor or the rich” measures if respondents perceive that reuse is for a certain class of people.

**Physical quality satisfaction:** Satisfaction with the physical quality of the reclaimed wastewater is most often the first determinant of a respondent’s willingness to accept reuse. These statements address this construct:

i. PQS1 – “This institution (I) will use non-drinking water if it is absolutely clear,” and

ii. PQS2 – “This institution (I) will use non-drinking water if it is colourless.”

**Identification of potential non-domestic and domestic respondents**

Two arid inland municipalities (Capricorn and Vhembe in the Province of Limpopo) were identified as suitable locations in South Africa to generate the data needed for the study. Limpopo is a water scarce province of South Africa while Capricorn and Vhembe contribute to South Africa’s agricultural production in the areas of field crops (e.g. cereals and oil seeds) and horticultural crops (e.g. potatoes, vegetables, citrus and deciduous). Wastewater reuse has therefore been proposed to many of the agricultural holdings within these two municipalities, as it shows promise of reducing the current dependence on fresh water for most activities, and reducing the total bill paid monthly on drinking water. Of the total water requirement within the area, agriculture was estimated to use about 85%. In terms of households, there were 1 243 167 people.
living in 285 565 households in Capricorn in 2007, while Vhembe houses 1 240 035 people in 287 190 households (Statistics South Africa 2008). Use of recycled wastewater for some household non-potable water requirements such as toilet flushing, is also promising when considering the arid climate within the municipalities.

**Sampling and data collection**

Potential non-domestic respondents within agricultural holdings, commerce, education and public parks were randomly approached to participate in this exercise, and several (especially within agricultural holdings) declined to participate. This may have been due to the fear that if the public knew they were willing or remotely considering wastewater reuse, the sale of their products may suffer. In contrast, however, the questionnaire response rate from a random sample of potential domestic consumers was higher (83% in comparison to 72% for non-domestic respondents). Table 2 summarises the questionnaires administered and returned.

The questionnaires were physically administered to respondents, i.e. participants were individually approached and encouraged to participate in the survey. Participation was voluntary. The demographics for potential domestic respondents were 52 males and 71 females aged 18 to 65 years, with a mean age of 25.2 years (SD = 7.2). The majority of respondents were black (99%). In terms of marital status, 60.1% were single, 12.2% were married, 25.2% were married with children and the remaining 1.6% were divorced or widowed. Most of the participants (69.5%) lived in Reconstruction and Development Programme houses, 19.8% in other houses, 4.9% in apartments, 4.1% in traditional houses and 1.7% in informal settlements. Household numbers varied from 2 to 10, with an average of 6 (SD = 5.2).

**Measurement validation and analysis**

As depicted in Figure 1, this study measured six constructs: attitudes, subjective norms, perceived control over the source of water and its application, physical quality satisfaction, knowledge of the advantages of reuse and trust in the service provider. The respondents were requested to rate how much they agreed or disagreed with each statement on a 5-point Likert scale from 1 (strongly agree) to 5 (strongly disagree).

Since the constructs were measured using multiple statements, it was necessary that the different statements used to assess the same construct should correlate with one another and exhibit high internal consistencies. This was achieved by determining the Cronbach’s alpha (α) value (which varies from 0 to 1.0) amongst multiple statements measuring a construct. It is generally accepted that a Cronbach’s alpha value above 0.70 is an indication of good internal consistency between items (Vicente & Reis 2008).

The analysis of the correlation between statements and their respective constructs was performed using SEM software called AMOSTM 6.0. The basis for the SEM approach is that the existence of a causal relationship between two variables does not imply the existence of a correlation between them (Iriondo et al 2003). Hence, AMOSTM 6.0 allows multiple relationships to be analysed simultaneously while maintaining statistical efficiency. AMOSTM 6.0 uses the maximum likelihood (ML) method to estimate parameters. SEM in its general form consists of a measurement model and a structural equation model. The measurement model specifies the relationship between statements and constructs, while the structural equation model specifies the relationships between constructs, describes their effects (either negative or positive) and assigns the explained and unexplained variance. The SEM therefore simultaneously estimates and tests a series of hypothesised inter-related relationships between a set of constructs, each measured by one or more statements. Intention to accept wastewater reuse for non-potable water requirements was measured as a second order construct, and thus predicted using the six constructs (e.g. trust, attitudes and control) in the hypothesised model shown in Figure 1.

**RESULTS AND DISCUSSION**

For potential non-domestic respondents, 20 statements in the questionnaire measuring the six constructs were subjected to item-to-total correlation and exploratory factor analysis. The item-to-total correlation is a correlation between a statement score and the sum of the remaining statements that form the scale. The test is performed to check whether any statement is inconsistent with the remaining statements. Once the number of correlated statements is determined,
exploratory factor analysis is performed to determine their factor loadings. Factor loadings are the correlation coefficients between the statements and the constructs. Factor loadings greater than 0.71 are typically regarded as excellent while less than 0.34 are regarded as very poor (Yongminga et al. 2006). Four of the 20 statements with factor loadings of less than 0.34 (statements used to measure the subjective norms and physical quality satisfaction constructs) were excluded from subsequent analysis. Details of factor loadings for each statement are shown in Appendix 1. The retained 16 statements, which are grouped according to their respective constructs (Appendix 1), explained 81.32% of the variance of the intention to accept wastewater reuse (Table 4) and were therefore reliable for further analysis.

For potential domestic respondents, 20 statements measuring the six constructs were subjected to item-to-total correlation and exploratory factor analysis. Two statements measuring the physical quality satisfaction construct generated a factor loading of less than 0.34 and were excluded from subsequent analysis. The retained 18 statements, which are grouped according to their respective constructs (Appendix 2), explained 87.02% of the variance of the intention to accept wastewater reuse (Table 5) and were therefore reliable for further analysis. For these respondents, the details of factor loadings for each statement are shown in Appendix 2.

Following the exclusion of statements with factor loadings less than 0.34, good fits were obtained for both domestic and non-domestic respondents (Table 3).

Table 3 Goodness of fit for revised model

<table>
<thead>
<tr>
<th>Fit index</th>
<th>Recommended value (Arbuckle 2005)</th>
<th>Non-domestic respondents</th>
<th>Domestic respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Structural model</td>
<td>Structural model</td>
</tr>
<tr>
<td>( \chi^2 ) / df</td>
<td>( \leq 3.00 )</td>
<td>2.60</td>
<td>2.30</td>
</tr>
<tr>
<td>AGFI</td>
<td>( \geq 0.80 )</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>NFI</td>
<td>( \geq 0.90 )</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>GFI</td>
<td>( \geq 0.90 )</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>CFI</td>
<td>( \geq 0.90 )</td>
<td>0.90</td>
<td>0.94</td>
</tr>
<tr>
<td>IFI</td>
<td>( \geq 0.90 )</td>
<td>0.90</td>
<td>0.92</td>
</tr>
<tr>
<td>TLI</td>
<td>( \geq 0.90 )</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>RMSEA</td>
<td>( \leq 0.10 )</td>
<td>0.08</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Table 4 Reliabilities and average variances extracted for potential non-domestic respondents

<table>
<thead>
<tr>
<th>Constructs</th>
<th>No of items</th>
<th>Composite reliability (( \alpha ))</th>
<th>Recommended value (Vicente &amp; Reis 2008)</th>
<th>Average variance extracted (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages of Reuse (ADV)</td>
<td>4</td>
<td>0.81</td>
<td></td>
<td>78</td>
</tr>
<tr>
<td>Trust (TRU)</td>
<td>4</td>
<td>0.82</td>
<td></td>
<td>80</td>
</tr>
<tr>
<td>Attitude (ATT)</td>
<td>6</td>
<td>0.78</td>
<td></td>
<td>86</td>
</tr>
<tr>
<td>Control over source of water (CON)</td>
<td>2</td>
<td>0.90</td>
<td>&gt; 0.70</td>
<td>71</td>
</tr>
<tr>
<td>Physical quality satisfaction (PQS)</td>
<td>2</td>
<td>0.31</td>
<td></td>
<td>35*</td>
</tr>
<tr>
<td>Subjective norm (SNO)</td>
<td>2</td>
<td>0.42</td>
<td></td>
<td>48*</td>
</tr>
<tr>
<td>Intention to accept</td>
<td></td>
<td>0.85</td>
<td></td>
<td>81</td>
</tr>
</tbody>
</table>

Table 5 Reliabilities and average variances extracted for potential domestic respondents

<table>
<thead>
<tr>
<th>Constructs</th>
<th>No of items</th>
<th>Composite reliability (( \alpha ))</th>
<th>Recommended value (Vicente &amp; Reis 2008)</th>
<th>Average variance extracted (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages of reuse (ADV)</td>
<td>3</td>
<td>0.82</td>
<td>&gt; 0.70</td>
<td>85</td>
</tr>
<tr>
<td>Trust (TRU)</td>
<td>4</td>
<td>0.73</td>
<td></td>
<td>77</td>
</tr>
<tr>
<td>Attitude (ATT)</td>
<td>5</td>
<td>0.68</td>
<td></td>
<td>84</td>
</tr>
<tr>
<td>Control over source of water (CON)</td>
<td>3</td>
<td>0.81</td>
<td>&gt; 0.70</td>
<td>71</td>
</tr>
<tr>
<td>Physical quality satisfaction (PQS)</td>
<td>2</td>
<td>0.43</td>
<td></td>
<td>46*</td>
</tr>
<tr>
<td>Subjective norm (SNO)</td>
<td>3</td>
<td>0.85</td>
<td></td>
<td>75</td>
</tr>
<tr>
<td>Intention to accept</td>
<td></td>
<td>0.80</td>
<td></td>
<td>87</td>
</tr>
</tbody>
</table>

* Constructs excluded from the computation of the average variance of the intention to accept wastewater reuse

Figure 3 shows the schematic of the hypothesised model for potential domestic respondents comprising five constructs. All the paths specified were statistically significant with \( p \)-values less than 0.01. Advantages of reuse (\( \beta = 0.39 \)) and trust in service provider (\( \beta = 0.21 \)) were found to have a moderate contribution to respondents’ intention to accept wastewater reuse. These constructs therefore moderately support hypotheses H1 and H2 respectively. Attitude (\( \beta = 0.60 \)) and control (\( \beta = 0.59 \)) have a strong contribution to respondents’ intention to accept wastewater reuse, and therefore strongly supported hypotheses H3 and H4. Hypothesis H6 could not be tested because a reliable measure of physical quality satisfaction was not obtained. The same applied to Hypothesis H5 for subjective norms.
Hypothesis H6 could not be tested because a reliable measure of physical quality satisfaction was not obtained.

**Implication of the results on planned wastewater reuse and future research**

i. In order of significance, the constructs that influenced the intention to accept wastewater reuse amongst the surveyed non-domestic consumers were their attitude towards wastewater reuse, the degree of control over the source of water and its application within their institution, the knowledge of the advantages of reuse and trust in the service provider. The influence of physical quality satisfaction could not be assessed because a reliable scale was not formed. The same applied to subjective norms. For subjective norms, this may be due to the fact that different institutions use water for different end uses, and hence the use of wastewater for a specific purpose in one institution would likely have a limited effect, if any, on another institution choosing to or not to reuse wastewater for a different end use. From the results of the item-to-total correlation and exploratory factor analysis, the two statements which measured physical quality satisfaction were either inconsistent with each other, uncorrelated to the construct they were to measure, or both. In future research in this regard, the use of more than two statements to measure each construct may result in a more reliable scale.

ii. For the domestic respondents surveyed, the constructs (in order of significance) that influenced intention to accept wastewater reuse were their knowledge of the advantages of reuse, the degree of control over the source of water and its application, attitude towards wastewater reuse, trust in the service provider and the subjective norms of the respondents. Similar to the results obtained for non-domestic respondents, the influence of physical quality satisfaction could not be assessed, because a reliable scale measuring this construct did not emerge. Future research with effective measurement of this construct is required, because households are typically concerned about the physical appearance of the reclaimed wastewater. Hence, similar to the argument above for non-domestic respondents, the two statements which measured physical quality satisfaction were likely either inconsistent with each other, uncorrelated to the construct they were to measure, or both.

iii. A limitation of this study was the small size of surveyed agricultural sector respondents – a very important potential user of treated wastewater. This limitation impacts negatively on the application of the results obtained above to agricultural holdings in the surveyed area, and hence future research which surveys a larger number of agricultural respondents is recommended.

iv. For a holistic assessment of respondents’ intention to accept reuse in locations contemplating wastewater reuse, it would be valuable to also survey and understand the religious and cultural perspectives of respondents (similar to the Wilson and Pfaff study (2008) conducted in Durban, South Africa.)

**CONCLUSIONS**

This paper presented a hypothesised model (adapted from Po et al 2005) which was used to predict intention to accept wastewater reuse for non-potable purposes amongst potential domestic and non-domestic consumers in two South African communities in the Limpopo Province. For the potential non-domestic consumers surveyed, intention to accept wastewater reuse was influenced by attitude towards wastewater reuse, the degree of control over the source of water and its application, the knowledge of the advantages of reuse and trust in the service provider. In order of significance for potential
domestic consumers surveyed, knowledge of the advantages of reuse, the degree of control over the source of water and its application, attitude towards wastewater reuse, trust in the service provider and the subjective norms of the respondents emerged as influences on the intention to accept wastewater reuse. Addressing these constructs would be valuable in determining whether wastewater reuse planning/implementation should proceed. Research to further interrogate this subject and the results presented herein was recommended in the previous section.

ACKNOWLEDGEMENTS

Dr Kelly Fielding is acknowledged for her invaluable contribution to this research. Funding from the Water Research Commission (project KS/1701) is gratefully acknowledged.

REFERENCES

Arbuckle, J J 2005. Amos v6.0 user’s guide. Spring House, PA, USA.

### Appendix 1 Factor loadings and internal consistency of statements for potential non-domestic respondents' questionnaire

<table>
<thead>
<tr>
<th>Construct</th>
<th>Statement</th>
<th>Factor loading</th>
<th>Composite reliability, α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages of reuse</td>
<td>1. The use of non-drinking water will reduce the amount of wastewater discharged to the environment, ADV1</td>
<td>0.80</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>2. Non-drinking water use will reduce the depletion of groundwater and surface water resources, ADV2</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. The use of non-drinking water can save many South African communities from drought, ADV3</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. There can be considerable savings of fertiliser on farms irrigated with recycled wastewater, ADV4</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Trust in implementing authorities</td>
<td>1. This institution will use non-drinking water if the quality can be proven to be satisfactory, TRU1</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2. This institution will use non-drinking water if it is not disgusting or irritating, TRU2</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. This institution will use non-drinking water if it does not stain or cause corrosion, TRU3</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. This institution trusts the municipality to provide non-drinking water that is safe and does not constitute a health risk, TRU4</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Attitude</td>
<td>1. This institution is obligated to do whatever it can to save water, ATT1</td>
<td>0.57</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>2. Water is a valuable resource that should be recycled, ATT2</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. This institution would prefer not to use non-drinking water, ATT3</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. This institution would never use non-drinking water, even in times of shortages, ATT4</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. This institution would only be prepared to use non-drinking water in times of water shortages, ATT5</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. The government is responsible for water shortages, ATT6</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Control over source of water</td>
<td>1. Every household should be free to choose their source of water supply (e.g. groundwater, surface water, recycled wastewater, etc), CON1</td>
<td>0.43</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>2. Fruits and vegetables irrigated with non-drinking water (e.g. recycled wastewater) should be labelled in the supermarket, CON2</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Subjective norms</td>
<td>1. This institution will use non-drinking water if other institutions are using it, SNO1</td>
<td>0.22*</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>2. Most institutions who are close to our institution will support the use of non-drinking water, SNO2</td>
<td>0.30*</td>
<td></td>
</tr>
<tr>
<td>Physical quality satisfaction</td>
<td>1. This institution will use non-drinking water if it is absolutely clear, PQS1</td>
<td>0.15*</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>2. This institution will use non-drinking water if it is colourless, PQS2</td>
<td>0.31*</td>
<td></td>
</tr>
</tbody>
</table>

* Items excluded from further analysis

### Appendix 2 Factor loadings and internal consistency of statements for potential domestic respondents' questionnaire

<table>
<thead>
<tr>
<th>Construct</th>
<th>Statement</th>
<th>Factor loading</th>
<th>Composite reliability, α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages of reuse</td>
<td>1. The use of non-drinking water will reduce the amount of wastewater discharged to the environment, ADV1</td>
<td>0.94</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>2. Non-drinking water will reduce the depletion of groundwater and surface water resources, ADV2</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. The use of non-drinking water can save many South African communities from drought, ADV3</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Trust in implementing authorities</td>
<td>1. I will use non-drinking water if the quality can be proven to be satisfactory, TRU1</td>
<td>0.70</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>2. I will use non-drinking water if it is not disgusting or irritating, TRU2</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. I will use non-drinking water if it does not stain washing, TRU3</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. I trust the municipality to provide non-drinking water that is safe and does not constitute a health risk, TRU4</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Attitude</td>
<td>1. I feel personally obligated to do whatever I can to save water, ATT1</td>
<td>0.51</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>2. Water is a valuable resource that should be recycled, ATT2</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. I would prefer not to use non-drinking water, ATT3</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4. I would only be prepared to use non-drinking water in times of water shortages, ATT4</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5. The government is responsible for water shortages, ATT5</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Control over source of water</td>
<td>1. I have the right to know if fruits or vegetables are irrigated with recycled wastewater, CON1</td>
<td>0.98</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>2. Fruits and vegetables irrigated with non-drinking water (e.g. recycled wastewater) should be labelled in the supermarket, CON2</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. I have the right to adequate drinking water supply, CON3</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Subjective norms</td>
<td>1. I will use non-drinking water if others are using it, SNO1</td>
<td>0.67</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>2. Most people who are close to me will support the use of non-drinking water, SNO2</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3. Non-drinking water use is an option for the poor or the rich, SNO3</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>Physical quality satisfaction</td>
<td>1. I will use non-drinking water if it is absolutely clear, PQS1</td>
<td>0.24*</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>2. I will use non-drinking water if it is colourless, PQS2</td>
<td>0.21*</td>
<td></td>
</tr>
</tbody>
</table>

* Items excluded from further analysis
The management of constructability knowledge in the building industry through lessons learnt programmes

V Kuo, J A Wium

In the 1980s the term "constructability" evolved in the USA. The proponents of this concept believe that constructability, which embraces both design and management functions, is comprehensive in facilitating construction operations and solving problems on site. Constructability problems are common on the construction site, due to the lack of construction experience in the design team and the absence of tools to assist designers in addressing constructability. Moreover, designs are predominantly done early in the project in the absence of contractor input, and there is yet no explicit means of defining or measuring constructability. This paper aims to address constructability problems in building construction, by understanding the nature of constructability knowledge and investigating how construction experience may be effectively disseminated amongst project participants, particularly through the use of lessons learnt programmes and inter-disciplinary knowledge sharing. It has been found that there is fundamental misalignment between consultants and contractors on the perceptions of criteria for a constructible design, implications of design decisions, and certain traits that may represent optimised vs poor constructability. The discrepancy in communication is the elemental cause of constructability problems and this research has demonstrated how lessons learnt programmes can be an effective tool in attaining better constructability knowledge management and collaboration.

INTRODUCTION

Constructability is a very common subject in the construction environment, often with controversial implications on the successful delivery of the project. Yet, little research has been dedicated to address such issues explicitly. Constructability issues arise from a design which does not sufficiently embody the expertise of construction processes, rendering the design “difficult to construct” on site. Subsequently this creates a variety of negative secondary effects during project execution, eventually manifesting as time-, budget- and quality risks to the project. This paper aims to investigate the nature of constructability problems in the building environment, and to understand how constructability in practice can be improved through better management of constructability knowledge, especially between consultant (designer) and contractor (constructor). Two key objectives are involved in this research:

1. Investigating explicit means to define constructability.
2. Investigating knowledge management activities, particularly through lessons learnt programmes, currently experienced in the industry.

Perceptions regarding constructability, similar to the implications of design and construction decisions, are mostly rooted in the experience of the industry professionals, and form the basis for this research. The primary methodology employed is thus a combination of questionnaire surveys and personal correspondences, in order to extract relevant knowledge from experienced professionals in the South African industry.

In this study, several traits have been identified as implications of optimised vs poor constructability, which are tested through the analyses of responses from South African industry practitioners. The lessons learnt programme (LLP) – its nature and implementation in organisations – is investigated in detail, and recommendations are given as to how LLPs can be a pertinent approach to improve the management of constructability knowledge, hence improving issues of constructability at large. The scope of the research is limited to the building industry, as most civil engineering organisations would have had substantial experience in building projects, where constructability issues are prevalent.

Some background on constructability and knowledge management related concepts are
Constructability codes
The implementation of constructability codes or guidelines poses a pertinent initiative in improving constructability as aid to the designer in the early stages. The challenge here lies in the explicity of constructability expertise—in other words, is it possible to explicitly represent constructability knowledge in a codified format, and to what extent?

In Asia, the Singapore government has put legislation in place as of 2001 to require minimum buildability scores of designs before approvals of building plans. The Buildable Design Scores are calculated based on the Buildable Design Appraisal System (BDAS), which was devised to measure buildability performance of designs in Singapore. The “3S” principles of Standardisation, Simplicity and Single Integrated Elements form the cornerstones of the BDAS. Good buildable designs will have to be complemented by the adoption of labour-efficient technologies and methods to improve productivity at the construction stage (BCA 2011). The Buildability Design Scores are thus determined primarily on this basis. In Singapore, the Buildable Design Score of any design must comply with the minimum allowable, before the project may commence.

Knowledge management
Different authors have presented different definitions of Knowledge Management (KM). Within construction, KM can be difficult to define precisely as there is not a general consensus on a single unified meaning of the concept (Egbu 2004). A combined generic definition for KM can be proposed and is used in this research (Davenport & Prusak 1998; Scarborough et al 1999; Robinson et al 2001):

“The process associated with the creation of new knowledge, the sharing and transfer of new and existing knowledge, the capture, storage, exploitation and measurement of the impact of knowledge, in such a way that it benefits the unit of adoption, which can be consulting or contracting organisations.”

The concept of constructability and related problems highlights the tacit nature of constructability knowledge. Constructability knowledge largely forms part of the experiences and expertise embedded within the minds of personnel in construction, and is difficult to standardise due to the diverse perceptions in the industry. Therefore, the sound management of knowledge resources is key to improving constructability overall. The sound management of knowledge within an organisation enables effective identification and dissemination of constructability problems and the subsequent efforts to mitigate or avoid them in future, by back-channelling constructability knowledge and expertise into the design and planning stages. This can be achieved through cross-organisational and cross-disciplinary knowledge exchange, particularly between consultants and contractors.

Tacit and explicit knowledge
Work by Polanyi (1958), and Nonaka and Takeuchi (1995) divided knowledge into tacit and explicit. Tacit knowledge represents knowledge based on the experience of individuals, expressed in human actions in the form of evaluation, attitudes, points of view, commitments and motivation (Nonaka et al 2000). Since tacit knowledge is linked to the individual, it is very difficult, if at all possible, to articulate. Researchers have argued that the diffusion of tacit knowledge is more difficult than sharing explicit knowledge (e.g. Nonaka & Konno 1998; Leonard & Sensiper 1998). Explicit knowledge, in contrast, is codified knowledge inherent in non-human storehouses, including organisational manuals, documents and databases.

Yet, it is difficult to find two entirely separated dichotomies of tacit and explicit knowledge; instead knowledge can fall within the spectrum of tacit knowledge to explicit knowledge. The management and sharing of tacit knowledge pose pertinent relevance to constructability, given its tacit nature. However, there is in existence also explicit, codified forms of constructability knowledge, such as Singapore’s Buildable Design Appraisal System. This research probes the extent to which tacit constructability knowledge can be explicated and used as codified knowledge by designers.

Lessons learnt programmes
Foy (1999) defines the concept of knowledge sharing as “facilitating learning, through sharing, into usable ideas, products and processes”. This naturally applies for both tacit and explicit forms of knowledge mentioned previously. A lessons learnt programme (LLP) consists of the people, processes and tools that support the (1) identification, (2) analysis and (3) implementation of new knowledge. This implies both the creation and sharing of validated lessons learnt.

Foy’s (1999) definition of knowledge sharing implies that “learning” is an artefact from the knowledge sharing process. An LLP therefore can be regarded as a vital tool in attaining effective knowledge management in the industry—especially of constructability knowledge. Harrison (2003) defines lessons learnt as “a good work practice or...
innovative approach that is captured and shared to promote repeat application, or an adverse work practice or experience that is captured and shared to avoid recurrence”. Organisations in the construction industry cannot afford to make repetitive mistakes on major projects. Conversely, there are great benefits to repeating positive experiences from past projects. An effective lessons learnt programme is a critical element in the management of constructability knowledge, in both explicit and tacit forms. The lessons learnt programme is investigated in detail in this study, along with the implementation of the project close-out meeting, which is a highly pertinent method for lessons learnt activities and cross-disciplinary knowledge sharing to be carried out.

**Methodology**

**Questionnaire and respondents**

For these investigations, the questionnaire survey was primarily used to extract tacit knowledge pertaining to constructability from experienced practitioners in the South African industry. Questionnaire surveys were distributed to a total of 50 industry respondents, and 28 completed responses were received – a response rate of 56%. Respondents of this study consist of both consultants and contractors – approximately the same number of respondent for each, so as to allow sensible comparison of the results. The 11 consultants and 17 contractors have varying years of experience, job positions and technical disciplines, as shown in Table 1.

It has not been the intention to focus particularly on a large-scale statistical or quantitative analysis of survey results. Rather, it was of greater significance to reach insightful practitioners to provide relevant and meaningful responses, hence the seemingly small number of respondents. Both consulting and contracting respondents have leading roles at their organisations, substantial amounts of professional experience, and a high level of familiarity with civil and building projects. Also, only contractors with a CIDB (Construction Industry Development Board) grading of 9CE and 9GB were chosen.

Due to the small number of the respondent group, regardless of respondents’ expertise, it is nevertheless worthy to note that the results from the surveys may or may not be representative of the industry at large, especially considering the fragmented nature of the construction industry. Furthermore, the quantitative analysis is done on qualitative data based on perceptions. The results were thus interpreted with a reasonable degree of scepticism and tolerance. The research takes care to consider all the limitations implicit in the research principles and questionnaire processes. However, the professional opinions of the practitioners offer valuable insight and experience, the credibility of which should not be ignored.

**Contents of questionnaires**

The questionnaire investigations were undertaken in two parts.

Part I consists of constructability-related investigations as follows:

- Labour efficiency principle of constructability
- Criteria of constructible design
- Constructability implications of design decisions.

Part II focuses on the knowledge management aspects of lessons learnt programmes (LLPs):

- Formality of current LLPs
- Methods where lessons learnt are carried out
- Project close-out meetings
- Perceived potential of LLPs.

**Part I Survey: Constructability Related Investigations**

The first survey essentially aims to define constructability more explicitly, based on the South African construction industry, and investigates the extent to which constructability can be codified. To do so, key aspects of constructability are identified to be investigated in detail, the discussions and interpretations of which are presented in the following sections:

- Labour efficiency principle
- Criteria of constructible design
- Constructability implications of design decisions.

**Labour efficiency principle**

The scoring system used in Singapore’s Buildable Design Appraisal System (BDAS) is based primarily on the labour efficiency principle. This aspect is tested in the South African context – whether or not designs and construction specifications promoting labour efficiency can be equated to good constructability.

In South Africa it can be said that the sizable industry opinion prefers the use of in situ concrete. Some reasons may be that it generates human labour and thus arguably increases employment, regardless of the efficiency of the labour; or that in situ concrete design typically has higher safety factors; or that in situ concrete construction processes do not require as much prudent coordination and planning as that of precast methods. Due to common usage of in situ concrete over the years, precast methods hold uncertainty that may be interpreted as potential project risks. South African industry personnel are speculatively more comfortable employing in situ concrete methods. This industry trend seems contradictory to the

<table>
<thead>
<tr>
<th>Table 1 Profile of survey respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Respondent</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Consultants</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Contractors</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2 Principle of work and choice of concrete system</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Work</strong></td>
</tr>
<tr>
<td>Equipment intensive</td>
</tr>
<tr>
<td>Consultant</td>
</tr>
<tr>
<td>Contractor</td>
</tr>
</tbody>
</table>
Table 3 Importance of criteria for constructible design

<table>
<thead>
<tr>
<th>Rank</th>
<th>Consultant</th>
<th>Contractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>Design requirements to be easily visualised and coordinated by site staff</td>
<td>Allow design to achieve safe construction sequence on site</td>
</tr>
<tr>
<td>2nd</td>
<td>Enable standardisation and repetition</td>
<td>Contractors to develop and adopt alternative construction details</td>
</tr>
<tr>
<td>3rd</td>
<td>Economic use of contractor’s resources</td>
<td>Enable standardisation and repetition</td>
</tr>
<tr>
<td>4th</td>
<td>Enable simplification of construction details in case of non-repetitive elements</td>
<td>Design requirements to be easily visualised and coordinated by site staff</td>
</tr>
<tr>
<td>5th</td>
<td>Allow design to achieve safe construction sequence on site</td>
<td>Economic use of contractor’s resources</td>
</tr>
<tr>
<td>6th</td>
<td>Contractors to develop and adopt alternative construction details</td>
<td>Enable simplification of construction details in case of non-repetitive elements</td>
</tr>
<tr>
<td>7th</td>
<td>Contractors to overcome restrictive site conditions</td>
<td>Contractors to overcome restrictive site conditions</td>
</tr>
<tr>
<td>8th</td>
<td>Freedom of choice between prefabricated and on-site works</td>
<td>Freedom of choice between prefabricated and on-site works</td>
</tr>
<tr>
<td>9th</td>
<td>Minimise the impact due to adverse weather by enabling a more flexible construction programme</td>
<td>Minimise the impact due to adverse weather by enabling a more flexible construction programme</td>
</tr>
</tbody>
</table>

labour efficiency principle of the Singaporean codes, where equipment-intensive work and precast concrete would be preferred, as they are of higher labour productivity (labour to output ratio).

In the survey, respondents were asked to indicate the preferred choice between equipment-intensive and labour-intensive construction, and between precast and in situ concrete. The results are shown in Table 2, as percentage of respondents for each choice.

The results show that both consultants and contractors prefer equipment-intensive over labour-intensive work. Consultants prefer in situ concrete, which in theory does not concur with the preference of equipment-intensive work concurrently indicated. On the other hand, contractors prefer precast concrete. A fundamental misalignment can be exhibited here regarding the preference of principle methods or approaches, implying different perceptions between consultants and contractors towards the inferred constructability of these approaches. This further emphasises the need for constructability issues to be studied to understand the nature of such misalignments. The reasons behind the differences in perceptions of the different parties are not explored in this paper. Nevertheless, sensible deductions can be made from these results.

Criteria of constructible design

Lam and Wong (2011) reviewed the Buildability Assessment Model (BAM), proposed for measuring buildability of designs and establishing benchmarks for the construction industry in Hong Kong. The BAM identifies nine “buildability factors” (as seen in Table 3), and appraisal is based on how well the designs embrace these nine factors as expressed by a large pool of experienced practitioners having hands-on construction expertise. The perceived importance of these “buildability factors” or “criteria for constructible design” within South Africa was investigated, and a comparison was drawn up between contractors and consultants. Survey respondents were asked to qualify the importance of these criteria and the results were analysed and ranked as shown in Table 3. Rankings of the consultant and contractor are placed alongside for comparison.

There are some similarities and differences in opinion between consultants and contractors regarding the criteria of constructible design. Good correlation can be seen in the three least important criteria ranked. Both consultant and contractor also agree on the importance of “standardisation and repetition”. However, contractors regard “allowance for adopting alternative construction details” (shown in blue in Table 3) and “safety of construction sequences” (shown in red in the Table 3) as important constructability criteria, while consultants rank these considerably lower. The consultants ranked “ease of visualisations of design requirements” (shown in green in Table 3) as the highest, while it is only ranked fourth by contractors. There is thus some misalignment between consultant and contractor regarding the importance of criteria of a constructible design.

Constructability implications of design decisions

Design decisions are predominantly made by the consultant in the absence of the contractor. Choices pertaining to the type of components to use in a building design, as well as the configurations of each component, have constructability implications during the execution of the project. This investigation essentially identifies the effect of different design choices on constructability.

In the Singaporean code, following the labour-efficiency principles, labour-saving indices are determined and calibrated with extensive industry input. These labour-saving indices essentially determine a constructability score, and is the crux of constructability quantification. Labour-saving indices are different for different variations of building components and subcomponents employed in the design. Based on a similar methodology of quantification, this study has identified some major building components and different configurations by which these components can be designed or constructed (components and configurations as shown in Table 4). The respondents were then asked to qualify each configuration in terms of “ease of construction”. This would ultimately indicate the preference of one configuration over another, which can in turn be equated to a constructability ranking. These perceptions (of consultants and contractors) regarding the favourability of different variations/configurations of the above building components can be used to attain a more explicit definition of constructability.

Instead of looking at the absolute constructability ratings, it is more relevant to regard the relative ranking of these configurations, as it essentially indicates the tendency for one configuration to be chosen over another from both consultant and contractor perspectives. Table 4 shows the overall ranking of constructability ratings for each configuration as indicated by the respondents. The top-ranked configuration for each building component is highlighted in each case. Note: Where constructability ratings are the same for two configurations, the standard competition ranking (so-called “1-2-2-4” ranking) is complied with. This implies that the tied items are assigned the equal ranks, directly below the preceding rank, and leaving a gap one less than the number of items tied (e.g. 1-1-3-4 or 1-2-2-4).

Regarding the top-ranked configurations of every component, good correlation can be seen between the consultant and contractor – all components correspond, except for the Roof Support. According to consultants and contractors, composite configurations (e.g. concrete-filled steel sections or steel encased in concrete), wherever applicable, are of low constructability. The in situ RC slab on beams is agreeably ranked last by both.
There are also some disparities between consultants and contractors. Consultants made it clear from the start that in situ concrete is preferred, which is reflected in the results here – consultants consistently rank precast configurations considerably lower. Contractors, on the other hand, rank the precast configurations higher than the consultants on all cases, indicating that contractors are more susceptible to using precast than what the consultants would perceive.

Despite contractors’ earlier indication that precast concrete is preferred over in situ concrete, some contradictions can be identified. For Structural Frame and Roof components, contractors ranked precast configurations lower than in situ configurations. Furthermore, despite the fact that both contractors and consultants indicated the preference of equipment-intensive over labour-intensive work, masonry wall configurations (which are more labour intensive) for the Façade Wall and Internal Wall components are ranked higher than both precast and in situ RC wall configurations. This apparent contradiction may indicate that there are other factors or reasons, besides the preferences of equipment or labour-intensive work, for favouring masonry configurations, or rejecting other configurations for that matter. These other factors may include direct costs (masonry being cheaper first-hand), uncertainty risks associated with seldom-used precast configurations, and availability of precast plants, but may also be a simple industrial habit, which is not clearly justifiable.

**PART II SURVEY: LESSONS LEARNT PROGRAMMES**

The sound management of constructability knowledge, from different phases of the project, can be seen as the fundamental solution to improving constructability in the industry. For one, the management of knowledge regarding different preferences (say from the Part I survey), and sharing this knowledge with relevant project parties, would invariably lead to better understanding and thus more informed decisions to avoid constructability problems.

Constructability problems are due to the poor integration of construction knowledge into the design phase. Knowledge management investigations give an indication of how constructability knowledge can be properly captured and disseminated amongst relevant project participants, particularly between consultants and contractors. Four investigations are done:

- Formality of lessons learnt programmes
- Methods where lessons learnt are carried out

**Formality of lessons learnt programmes**

The respondents were asked to choose whether the lessons learnt activities at their organisations are: formal (standardised protocol built into organisational process, with designated coordinator); informal (occurs haphazardly, no standard process, no designated process coordinator); or does not exist at all. Table 5 shows the results, indicating that most organisations undertake lessons learnt activities on an informal, unstructured basis.

**Methods where lessons learnt are carried out**

Three lessons learnt processes were identified in literature: lessons identification, analysis and implementation. The respondents were given the list of methods/occasions where lessons learnt can be carried out and asked to indicate which ones they use. The percentages of respondents for each method are presented in Table 6 (in no particular order). The highest ranked method/s is/are highlighted for each lessons learnt process.

The top methods that are shown here exhibit a people-orientated nature, where socialisation and human interaction are prevalent. The social nature seems appropriate for sharing constructability knowledge, considering its highly tacit characteristic. However, the social nature may also be associated with the lack of structure in such lessons learnt activities.

**Project close-out meetings**

As expected, and as shown in the previous section, the project close-out meeting is a very important and widely used method for
are is of key importance, as it roughly indi-
if so, who (by discipline or position) are the
out in design and/or construction firms, and
Are such close-out meetings indeed carried
industry. The essential question posed here is:
close-out meetings as implemented in the
is given to investigate the nature of project
of lessons. For this reason specific attention
carrying out lessons learnt processes – in

### Table 5 Formality of lessons learnt programmes

<table>
<thead>
<tr>
<th></th>
<th>Formal</th>
<th>Informal</th>
<th>Does not exist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultant</td>
<td>36%</td>
<td>64%</td>
<td>0%</td>
</tr>
<tr>
<td>Contractor</td>
<td>29%</td>
<td>59%</td>
<td>12%</td>
</tr>
</tbody>
</table>

### Table 6 Methods with which lessons learnt processes are carried out

<table>
<thead>
<tr>
<th>Lessons learnt processes</th>
<th>Methods to carry out lessons learnt processes</th>
<th>Consultants</th>
<th>Contractors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lessons identification</td>
<td>Project close-out meetings</td>
<td>73%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>Intermediate meetings</td>
<td>91%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
<td>27%</td>
<td>59%</td>
</tr>
<tr>
<td></td>
<td>Electronically</td>
<td>55%</td>
<td>59%</td>
</tr>
<tr>
<td></td>
<td>Paper forms</td>
<td>36%</td>
<td>53%</td>
</tr>
<tr>
<td></td>
<td>Informally (word of mouth)</td>
<td>91%</td>
<td>76%</td>
</tr>
<tr>
<td></td>
<td>Outside consultant</td>
<td>18%</td>
<td>59%</td>
</tr>
<tr>
<td>Lessons analysis</td>
<td>Project close-out meetings</td>
<td>64%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>Intermediate meetings</td>
<td>91%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>By a subject matter expert</td>
<td>45%</td>
<td>24%</td>
</tr>
<tr>
<td></td>
<td>Electronically</td>
<td>45%</td>
<td>47%</td>
</tr>
<tr>
<td></td>
<td>By an outside consultant</td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td>Lessons implementation</td>
<td>At meetings</td>
<td>82%</td>
<td>65%</td>
</tr>
<tr>
<td></td>
<td>As changes to a work process</td>
<td>73%</td>
<td>65%</td>
</tr>
<tr>
<td></td>
<td>At project kick-offs</td>
<td>45%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>Through electronic databases</td>
<td>18%</td>
<td>29%</td>
</tr>
<tr>
<td></td>
<td>Informally (word of mouth)</td>
<td>82%</td>
<td>71%</td>
</tr>
<tr>
<td></td>
<td>Training/mentorship programmes</td>
<td>36%</td>
<td>53%</td>
</tr>
</tbody>
</table>

### Table 7 Perceived effectiveness of lessons learnt programmes

<table>
<thead>
<tr>
<th></th>
<th>Not effective</th>
<th>Neutral</th>
<th>Somewhat effective</th>
<th>Very effective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>6.5%</td>
<td>8%</td>
<td>71%</td>
<td>14.5%</td>
</tr>
<tr>
<td>Potential</td>
<td>0%</td>
<td>0%</td>
<td>40.5%</td>
<td>59.5%</td>
</tr>
</tbody>
</table>

Therefore many issues (especially that of constructability) cannot be addressed with collaborative engagement from both consultant and contractor.

- Most common attendees of project close-out meetings are: contractor managers, project managers, site managers, directors, estimators and quantity surveyors.
- The dynamic exchange of knowledge and inter-disciplinary discussions of constructability issues are not substantial at project close-out meetings. This investigation reveals that there is very limited consultant presence at contractors’ project meetings. This implies that constructability problems experienced on site that may have been attributed to a poor design, cannot be discussed in full with engagement from design personnel. Even if the project close-out meetings are implemented prevalently, and lessons learnt activities are carried out, it may still not achieve the required benefits. Active engagement from both the consultant and the contractor simultaneously is the key to a relevant knowledge exchange practice.

**Perceived effectiveness of lessons learnt programmes**

The respondents were asked to choose the effectiveness of lessons learnt programmes (LLPs) currently implemented at their organisations, as well as their perceptions of the full potential. Table 7 shows the percentage of respondents (consultants and contractors combined) and the choices made.

Most respondents perceive current implemented LLPs to be “somewhat effective”, while the full achievable potential can be “very effective”. The results show that respondents perceive higher potential effectiveness than how it is currently implemented. There is thus premise for improvement in LLPs and implies a degree of industry susceptibility for such endeavours.

**CONCLUSIONS**

This study sets out to understand the nature of constructability knowledge and related problems, and how constructability may be defined more explicitly. In doing so, the study determines how knowledge management initiatives, such as the lessons learnt programme, may be effectively employed for better knowledge dissemination, the lack of which, between consultants and contractors, has been identified as the principal cause of constructability problems. The study succeeded in the following:

- Explored previous research on constructability and knowledge management related concepts...
Established preferences of certain design and construction choices and approaches that may be equated, or used to define, constructability more explicitly

Investigated the nature of project knowledge sharing and documentation in project close-out meetings

Understood how lessons learnt programmes are implemented in the industry context, and the types of methods where lessons learnt activities are facilitated. It has been found that constructability problems manifest in stages of project execution, due to the fragmentation in project design and construction. Consultants/designers typically do not embody enough knowledge about constructability to integrate into their designs. On the other hand, contractors share constructability knowledge mainly in a tacit form in the minds and experiences of personnel, making it very difficult to disseminate with designers. These challenges, together with traditional procurement standards, restrict the collaborative capacity between consultants and contractors. A more collaborative approach is key to improving constructability, where tacit constructability knowledge/lessons can be created, shared, learnt and actively integrated into the relevant stages of the project, particularly that of design. A summary of some conclusions arising from this study can be made:

■ Constructability knowledge exists predominantly in tacit format, forms complex networks of cause and effect, and related issues occur under unique circumstances.

■ There is some misalignment between consultants and contractors regarding certain design and/or construction decisions and their implications on the project. Thus, the definition of optimised vs. poor constructability is understood differently by consultants and contractors. This demonstrates a real need for more collaboration between the two disciplines.

■ Consultants prefer the use of in situ concrete, while most contractors prefer precast concrete. Concurrently, both agree on the preference of equipment-intensive methods, rather than labour-intensive.

■ It can be concluded from this study that a building design which best represents a constructible project, according to contractor respondents, would consist of:
  ■ an in situ reinforced concrete structural frame
  ■ precast slabs with in situ topping
  ■ concrete masonry façade walls
  ■ dry internal partition walls with applied finishes

■ timber roof trusses
■ precast concrete ring beams as roof support.

Lessons learnt programmes (LLPs) are carried out mostly informally on an unstructured basis, where the methods used are of a social, people-oriented nature. Respondents see higher potential effectiveness than what is currently experienced, indicating premise for improvement.

There is generally a lack of integration between construction and design, due to a poor collaborative mentality. The management of cross-disciplinary knowledge/lessons (particularly that of constructability) is insufficient.

RECOMMENDATIONS
As can be deducted, “constructability” in itself is not the problem per se, but it is the consequence of a complex relationship of preceding problems, both technical and social in nature. These are often rooted in the organisational culture and philosophies of the company, such as issues regarding collaboration, communication, or willingness to implement structured knowledge management, etc. Nevertheless, “constructability problems” manifest prevalently and cause not only tangible waste in time, budget and quality, but also an inherent over-exposure to a diverse range of project risks.

To motivate improvement measures through financial and economic analyses is impossible, due to the complex nature of this network of interactive problems, most of which are hardly quantifiable and lie within fields of psychology and social sciences. Attempts to monetarily justify certain decisions over others would be erroneous to improving constructability, since the attempts to quantify with respect to cost would be subjective to start off with, and misrepresentative of the problem at hand. Furthermore, due to the uniqueness of the reasons behind certain design or construction decisions, it is also very difficult (and probably not meaningful) to infer principles dictating which types of design/construction configurations or methods are definitely more constructible than another. Nevertheless, a fundamental certainty arises from this research – there is distinct misalignment between the perceptions and assumptions of different project participants (particularly consultant and contractor), due to poor communication or collaboration, and a lack of sound knowledge management practices.

In this research a multipronged approach is recommended as a strategic measure to improve constructability. The recommendations are on a broader, more holistic level in order to properly address the complex and diverse nature of constructability problems, given its non-explicit and often anecdotal property. The following recommendations should complement one another, and should not be seen as mutually exclusive initiatives:

■ Explicate constructability concepts further to develop codes and guidelines to aid designers.

■ Promote the consideration of precast methods as alternative to in situ, in order to understand and optimise the potential benefits of precast.

■ Establish an organisational culture, which promotes structured, cross-disciplinary knowledge sharing and a more collaborative mentality to project delivery.

■ Adopt and formalise lessons learnt programmes, integrated systematically into operational procedures.

■ Consultants should carry out designs with more consideration of safety on site, as well as increased flexibility for alternative construction details later in the project.

■ Increase research on projects in different procurement environments (such as design-build projects) to investigate whether more collaboration indeed occurs, or whether less constructability issues arise. A procurement model where design and construction considerations are integrated could potentially result in more optimal outcomes.

REFERENCES


Goddard/IEEE Software Engineering Workshop, Institute of Electrical and Electronics Engineers, California, Los Alamitos.
Evaluation of the strength behaviour of unpaved road material treated with electrochemical-based non-traditional soil stabilisation additives

R J Moloisane, A T Visser

Travel along unpaved roads is not always assured, because of their low standards, poor riding quality, impassability in wet weather, and the danger in the quantity of dust that is generated by moving vehicles and wind. Stabilisation with electrochemical-based non-traditional soil stabilisation additives (chemical additives) may offer a solution to this continual problem.

The objective of this paper is to report on the strength behaviour of a typical marginal quality weathered quartz gravel material treated with two electrochemical-based non-traditional soil stabilisation additives, enzyme and sulphonated oil to assess their potential value for unpaved road construction under wet and dry conditions. These treated panels were trafficked under 100 vehicles per day. The evaluation was done by means of laboratory tests and field investigations for three years. The characteristics of the natural material and the binding ability of the non-traditional soil stabilisation additives were established from the laboratory testing. Density and moisture, and the strength development of the treated material were determined from field investigations.

These two non-traditional soil stabilisation additives appear to have affected the particles and their water component, hence an increase in densities was achieved. The degree of formation and paste surrounding the particles appeared to have varied with time and differed between the additives. An increase in density in the sulphonated oil additive treated panel occurred three months after construction, and a further increase was again noticed eight months after construction (five months thereafter). Up to eight months after construction, the enzyme additive treated panel showed a significant decrease in density, but showed a slight increase thirty-one months after construction. This increase in densities might probably be because of further densification by traffic. The variations in density were attributed to testing variability.

In the in situ and soaked DCP-CBR strength measurements, the sulphonated oil additive treated panel reached its maximum in situ strength at two months after construction, while the enzyme additive treated panel reached its maximum in situ strength at five months after construction. Up to eight months after construction, both treated panels indicated a significant decrease in both the in situ and soaked DCP-CBR strength conditions. The decrease was attributed to rain. There was, however, little evidence to show that the additives had improved the material, with the control panel being consistently stronger in both the in situ and soaked DCP-CBR strength conditions.

The importance of considering the time factor in the strength development of non-traditional stabilisation test techniques, as well as the number of tests, was highlighted in the results of this study. The natural variability of the materials used in this type of study is generally high, and the precision of the test method is typically quite low. On this basis, it is usually difficult to draw definite conclusions.

INTRODUCTION

Unpaved roads make up a significant portion of the rural and urban road network, and by promoting access, this network is one of the key factors contributing to economic growth. Travel along unpaved roads is, however, not always assured because of their low standards, poor riding quality, impassability in wet weather, and the danger in the quantity of dust that is generated by moving vehicles and wind. This dustiness of unpaved roads reduces visibility and hinders traffic safety. Many unpaved roads do not have the ideal range and distribution of particle sizes to give a good load-bearing capacity when wet (coarser particles) or sufficient plastic capability (clay) when dry to prevent material from breaking loose. All unpaved roads, when dry, suffer from surface abrasion loss resulting in dust because the adhesion
between the particles is reduced. The loss of road construction material also leads to the formation of ruts and potholes that could collect/retain water, and this may decrease the natural durability and strength properties of the unpaved road system, making the road uncomfortable and dangerous to drive on.

The quality of most unpaved road materials can be improved with traditional chemical additives, such as cement and lime, to improve strength and durability properties, but stabilisation with electrochemical-based non-traditional soil stabilisation additives (chemical additives), such as enzyme and sulphonated oil (also referred to as ionic additive) may offer a more cost-effective and practical solution to address these ongoing problems. Chemical additives used for strength improvement are referred to as stabilisers or additives, and for dust suppression are referred to as dust palliatives (Jones & Ventura 2004).

The objective of this paper is to report on the strength behaviour of a typical marginal quality unpaved road material treated with two electrochemical-based non-traditional soil stabilisation additives, enzyme and sulphonated oil, to assess their potential value for unpaved road construction. This work is based on the laboratory and field testing of panels reported in Moloisane (2009). Although the field experiment included traditional soil stabilisation additives, the focus of this paper is the non-traditional soil stabilisation additives.

The Moloisane (2009) study was undertaken to evaluate the strength behaviour of unpaved road materials treated with non-traditional soil stabilisation additives and included short-term to medium-term investigations as part of the main study, as well as longer-term studies of a previous investigation by Van Veelen (2005). The experimental sections by Van Veelen (2005), built in the same area as those of the Moloisane (2009) study, were monitored for about eight months in 2005. Data collected in 2009, i.e. 48 months after construction, were used to investigate certain technical issues, including the non-traditional stabilisation additive properties and the in situ strength development behaviour. The Van Veelen (2005) study panels seldom received traffic, because they were constructed on the shoulder area of the existing surfaced road, whereas the panels in this study were constructed on an existing unpaved road and were trafficked at least under 100 vehicles per day. Since the Van Veelen (2005) study served as an effort in building sound guidelines on the long-term strength behaviour of the use of non-traditional soil stabilisation additives, it was used for comparison in the Moloisane (2009) study. The results of the tests are reported in this paper. Comparative strength of the treated unpaved road material and control (untreated) sections was also investigated during the testing. Both laboratory and field tests indicated that the application of enzyme and sulphonated oil increased the strength of the unpaved road material, provided that they are used properly.

**OVERVIEW OF CHEMICAL STABILISATION OF UNPAVED ROAD MATERIAL**

The main objective of chemical stabilisation on unpaved road material, as explained by among others Kézdi (1979) and Ballantine & Rossouw (1989), is to maintain the characteristics of the material, favourable from the aspects of the given engineering target, regardless of the moisture in its environment. As a result of chemical stabilisation, the bearing capacity and strength of the unpaved road material should be increased; and water-tightness, resistance to erosion and other properties should also be improved.

Stabilisation of unpaved road material with electrochemical-based non-traditional soil stabilisation additives (chemical additives) such as enzyme and sulphonated oil (ionic additive) has been investigated in a number of studies, and is on-going in the form of experimental, trial, control and demonstration sections and laboratory performance testing. Enzyme is defined as an organic catalyst that rapidly carries a chemical reaction to completion without becoming a part of the end product and being the reaction that would normally take place at a much slower rate (Burns 1978; Chang 1984; Tolleson et al 2003; Marasteanu et al 2005; Velasquez et al 2006; Pacific Enzymes Incorporated 2007). Enzyme catalyses very specific chemical reactions if conditions are conducive to the reaction (Tingle et al 2007). The manufacturers claim that this liquid additive, which is also non-hazardous and environmentally friendly, lowers the surface tension of water, hence aiding compaction, and that it creates the crystalline structures amongst individual soil particles through biologically catalysed reactions. Enzyme additives also help the soil bacteria to release hydrogen ions, resulting in pH gradients at the surfaces of the clay particles, which assist in breaking up the structure of the clay (Velasquez et al 2005). Sulphonated oil consists primarily of strong acidic sulphur-based organic mineral oils (Paige-Green & Groenewald 1993; Van Zyl et al 1993; Paige-Green & Coetser 1996; Savage 1998; Ali et al 1999; Tingle et al 2007; Van Veelen & Visser 2007). The manufacturer claims that this liquid anionic additive which is non-hazardous, non-flammable, non-corrosive when diluted, non-toxic, environmentally safe and user-friendly, is a synthetic compound with surface-active properties, which has been devised to change the hydrophilic (water adsorption) properties of clay materials to those of hydrophobic nature (water repellent) (Con-Aid International 1998). Performance of the electrochemical-based non-traditional soil stabilisation additives depends on the clay mineralogy, and sufficient time to react with the clay fraction (Austroads 1998; Bergmann 2000; Gauteng Provincial Government 2004; Jones & Ventura 2004).


Scholen (1992) presented positive results for soil-aggregate stabilisation with three commercial enzymes. Those enzymes included a bacterial culture with enzymes’ solution that multiplied rapidly when exposed to air, and produced the organic molecules which were necessary to attract to the clay minerals. Well-graded aggregates with high clay contents were found to have performed best by “locking” the larger aggregate particles within the matrix, hence the creation of a rigid surface and reduced ravelling. Stabilisation performance results from the study done by Wright-Fox et al (1993), using two enzyme additives on a highly plastic fat clay material, indicated that the unconfined compressive strengths (UCS) of the enzyme-treated products were 21% higher than the control (untreated) samples. The conclusion by Wright-Fox et al (1993) that enzymes might provide some additional shear strength for some soils, led...
to the recommendation that soil stabilisation with enzymes must be considered for various case-by-case applications. Funk (1993) reported that after seven years of using the enzyme in unpaved road stabilisation, blading was only required four times a year, compared with every one to three weeks prior to application, hence an enzyme was considered an extremely cost-effective product. Brownfield (1994) reported that roads treated with enzyme have maintained a hard and stable road surface.

In the study conducted by Bolander (1997), the laboratory tests met the expectations of the field performance. This also concurred with other laboratory studies conducted by Petry (1997), Santoni et al (2001) and Tingle & Santoni (2003) in which enzymes and sulphonated oils used in unpaved roads provided increased strength. The laboratory tests of expected field performance study of various additives, which included enzymes and sulphonates on dense-graded aggregate evaluated by Bolander (1999), included indirect tensile strength (ITS) and durability testing. Enzymes and sulphonates provided some tensile strength, but lost it with increasing moisture content. Based on the experiences of Bolander (1999) it is clear that thorough preparation, sufficient mixing and curing time all had a marked effect on the efficiency, durability and service life of the structures treated with enzymes and sulphonates.

Performance-based field testing using non-traditional soil stabilisation additives, which included an enzyme and ionic additive (sulphonated oil) conducted by Weedon & Croeser (1996), indicated that sections constructed with ionic additives did not reveal a marked difference compared with the control (untreated) section. The increased in situ strength was attributed to the reduction of in situ moisture content. In the laboratory experiment, using various non-traditional soil stabilisation additives which included enzymes and sulphonates conducted by Tingle & Santoni (2003), to evaluate the stabilisation of low- and high-plasticity clay soils, the focus was on increased load-bearing capacity as the basis of performance characterisation as indicated by the UCS. Sulphonates provided excellent UCS improvement under both dry and wet test conditions, and also provided the greatest strength improvements for the low-plasticity soil. Sulphonates further provided the best resistance to moisture, and this indicated excellent waterproofing characteristics. Enzymes showed a slight increase in the UCS for the low- and high-plasticity soils under both dry and wet test conditions. Consistent road enhancement and better performance from enzyme and sulphonated oil-treated roads based on records of the road performance were found by Brazetti & Murphy (2004).

Locally, the strength behaviour of four different soils treated with four different non-traditional soil stabilisation additives, which included an enzyme and sulphonated oil under dry and wet conditions, was evaluated by Van Veelen (2005). The enzyme-treated materials showed an improvement on the gravel and ferricrete materials, with a significant increase in soaked strength for the gravel material. Sulphonated oil treated materials showed an improvement on all materials with a significant increase in soaked strength for the clay material. The good performance of the sulphonated oil additive on the clay material was attributed to the fact that the reactive clay mineral in the clay material produced a permanent association between the additive and the clay particles. Both enzyme and sulphonated oil showed an increase in strength over the eight-month test period. The in situ California Bearing Ratio (CBR) strengths that were gained eight months after construction of the experiment indicated that the enzyme, which had no significant improvement in strength during the five months of testing, showed a significant improvement in strength on the ferricrete and gravel materials. This revealed that the additives needed a curing time of a few dry months to reach their maximum strength. Experimental field investigation study to evaluate stabilisation of road materials with non-traditional soil stabilisation additives, which included a sulphonated oil-based product, conducted by Visser (2007) revealed gain in strength in the material with active clay. It was also found that the product had no benefit when used on inert sand with low plasticity index (PI).

There is thus ample evidence that improvement in material characteristics occurred with the use of both non-traditional soil stabilisation additives, although the effect of increased strength caused by drying out of the treated materials is probably not always fully considered.

### EXPERIMENTAL PROGRAMME

#### Construction of the experimental panels

The chosen experimental site was an existing unpaved road called Larch Road, which is in the Benoni Agricultural Holdings in the Ekurhuleni Metropolitan Municipality, South Africa. This road is used by both local light and heavy-duty agricultural traffic. The test pavement construction was initiated on 8 May 2008. About 150 mm depth of the existing wearing course was ripped and mixed after spraying with potable water to about optimum moisture content and compacted to 93% Modified AASHTO density. A new 150 mm thick layer was imported on top of this reworked layer, which acted as subgrade. The pavement was then divided into panels 5 m long and 6 m wide. Panels that used untreated material were also constructed between the treated panels and used as controls. These control (untreated) panels were also used to prevent contamination between the treated panels. For consistency and good quality control, only one control (untreated) panel was used throughout the testing. Water was not applied to the road surface before mixing for fear of exceeding optimum moisture. A rotavator machine was used to pulverise the gravel-wearing course into workable material. The mixed treated wearing course material was compacted to 93% of Modified AASHTO density and then shaped by means of a motor grader to a crown-shape to allow for easier drainage.

#### Unpaved road material characterisation

The natural weathered gravel material used in the construction of the experimental panels was obtained from a commercial source, Benoni Sands Quarries. The following properties of this material were determined in the laboratory: sieve analysis, consistency limits, maximum dry density (MDD) and optimum moisture content (OMC), and California Bearing Ratio (CBR). Tests were conducted in accordance with TMH1 (CSRA 1986) standard methods and compared to the TRH14 (CSRA 1985) classification guidelines, and the COLTO (COLTO 1998) performance specifications. The material was also characterised according to TRH20 (CSRA 1990) shrinkage product (S<sub>p</sub>) and the grading coefficient (G<sub>c</sub>). S<sub>p</sub> is the product of linear shrinkage and the percentage passing the 0.425 mm sieve, while the G<sub>c</sub> is the difference between the percentage passing the 26.5 and 2.00 mm sieves multiplied by the percentage passing the 4.75 mm sieve, all divided by 100. The material was also classified according to the American Association of State Highway and Transportation Officials’ AASHTO’s M145 classification system as described by AASHTO (1998). X-ray diffraction (XRD) was used for the mineralogical analysis.

Sieve analysis test results indicated that the dark-reddish brown material was sandy gravel. The soil mortar test results showed that the material consisted of an average of 49% coarser particles, an average of 24% finer particles, and an average of 27% very fine particles. The average
grading modulus of the material was 1.70. Consistency limit test results showed that the material had an average plasticity index of 8%. The laboratory testing showed an average MDD of 2 100 kg/m³ at an average OMC of 8%, and average CBRs at 95% and 90% Modified AASHTO density of 15% and 11% respectively. The compacted density of the panels was determined after construction to assess future densification that might occur due to traffic and the stabilisation process. The panels were constructed on an existing unpaved road and it was deemed that there were no differences in the supporting layer for all the panels.

Performance-based material criteria of the untreated material as per COLTO (COLTO 1998), i.e. 

\[ S_p \] and \[ G_r \] results, indicated that the material was relatively good, but with some potential for ravelling and corrugating. The material was classified as A-2-4 according to the AASHTO M145 classification system (AASHTO 1998). The A-2-4 materials are deemed poorer because of their inferior binding characteristics, poor grading, or a combination of the two (AASHTO 1998; 2000).

From the XRD diffraction patterns of the untreated material, quartz was the main constituent (49%) in addition to traces of the other crystalline phases, namely, plagioclase (albite series) (26%), microcline (illite) (5%). In the clay mineral phases, montmorillonite was the main constituent (7%) in addition to a trace of kaolinite (3%). The natural material was classified as weathered quartz gravel.

The four unpaved road materials of the Van Veelen (2005) study were also classified according to the AASHTO M145 Classification System for comparison with the unpaved road material of the Moloisane (2009) study. A summary of the critical properties of the unpaved road materials of these two studies is presented in Table 1.

The Putfontein clay material of the Van Veelen (2005) study was also classified as A-2-4. This unpaved road material closely matched the dark-reddish brown gravel material of this study physically – it is not known whether the mineralogical compositions also match closely.

Non-traditional soil stabilisation additives used

The origins of the commercial non-traditional soil stabilisation additives used in the study are presented in Table 2.

Because of the proprietary nature of these commercial additives, their exact chemical compositions are not disclosed. The enzyme additive used in the study is a water-based additive (produced from fermenting sugar beet) which consists mainly of organic compounds (hence biodegradable) and surfactants. The sulphonated oil additive used in the study is a viscous, deep red or dark brown liquid anionic additive with no smell or taste, with the active agent being an organic acid, produced from a blend of synthetic-chemical products that originated from petroleum.

Methodology

The enzyme additive was applied at a rate of 0.005 ℓ/m² (0.033 ℓ/m³ or 0.15 ℓ in the 4.50 m³ panel) and the sulphonated oil additive at 0.01 ℓ/m² (that is 0.015 ℓ/m³ for the 4.50 m³ panel) to the newly imported 150 mm layer of imported natural quartz gravel material. The application rates were recommended by the suppliers/manufacturers, and thus adhered to and used as such. The long-term strength behaviour of enzyme and sulphonated oil additives on the unpaved road material (new wearing course) performance was observed after 14 days, one month, two months, three months, four months, five months, six months and eight months by means of the laboratory testing and field investigations, and after 31 months (almost three years) by field investigations only. The laboratory testing included scanning electron microscopy (SEM), while the field investigations assessed the in situ dry density and moisture content determined by a nuclear density meter, and the in situ and soaked strength was determined by a dynamic cone penetrometer (DCP).

The scanning electron microscope (SEM) is an electron microscope capable of producing high-resolution images of a sample surface morphology (Locquin & Langeron 1983). SEM uses electrons to illuminate and create an image of a specimen. The binding ability of the particles of the untreated natural material and the additives’ interactions in the panels were assessed using the SEM. The JEOL (JSM-840) model, which performs morphological and micro-structural assessment, was used for the analysis to obtain the micrographs or images that showed surface morphology information of the treated materials. The specimens were prepared by breaking off small pieces from the samples in order to look at a freshly exposed surface that broke along a natural plane of weakness.

The nuclear density measurement test was performed to determine the compaction of the treated material. This test compares the field density with the laboratory compaction density in order to determine the percent compaction achieved. The nuclear density meter measures the mass of the wet soil per volume and the mass of the water present in a unit volume of soil within a few minutes. This was used to compare and correlate the DCP results with the CBR at an appropriate Modified AASHTO density on treated material determined in the laboratory.

### Table 1 Summary of the comparison of the unpaved road materials of the two studies by Moloisane (2009) and Van Veelen (2005)

<table>
<thead>
<tr>
<th>Study</th>
<th>Material</th>
<th>Material property</th>
<th>Material group in AASHTO classification</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Van Veelen (2005)</td>
<td>Putfontein clay</td>
<td>33 20 27 10</td>
<td>A-2-4</td>
<td>Silty gravel/sand</td>
</tr>
<tr>
<td></td>
<td>Daveyton clay</td>
<td>81 36 15 4</td>
<td>A-4</td>
<td>Sandy clay</td>
</tr>
<tr>
<td></td>
<td>Benoni clay</td>
<td>46 30 28 11</td>
<td>A-2-6</td>
<td>Clayey sand/gravel</td>
</tr>
<tr>
<td></td>
<td>Quantum clay/sand</td>
<td>34 23 30 12</td>
<td>A-2-6</td>
<td>Clayey sand/gravel</td>
</tr>
</tbody>
</table>

### Table 2 Electrochemical-based non-traditional soil stabilisation additives used

<table>
<thead>
<tr>
<th>Additive name</th>
<th>Additive form</th>
<th>Country of origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>Liquid</td>
<td>United States of America</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>Liquid</td>
<td>South Africa</td>
</tr>
</tbody>
</table>
The DCP-CBR of the in situ material from the panels was compared with the CBR determined from the laboratory tests. This made it possible to compare the CBR values of the material from the tests conducted in a controlled environment (that is in the laboratory) with those obtained from the field investigations.

For the DCP testing in the soaked condition the procedure made use of a cleaned open-quarter of a 200 ℓ steel oil drum being placed on the panel surface and sealed with soil to reduce the water leaking out. This open-quarter drum was then filled with water to a depth of 50 mm, which was maintained for two hours. The amount of water added should soak the material to a depth of about 150 mm (assuming about 30% voids). In practice, a significant amount of water flows laterally and the material is probably not fully soaked to the full depth of 150 mm. The DCP test was then carried out in the middle of the wetted spot to test for the ‘soaked CBR’. The aim of the soaked DCP test was to determine the conditions comparable with the soaked CBR in the laboratory and to reflect the weakest road condition.

**RESULTS AND DISCUSSION**

**Scanning electron microscopy (SEM) analysis**

SEM analysis was carried out to identify the changes of the phase composition and microstructure. The images represent the general observed features.

**Untreated material**

Plate 1 depicts the SEM images of the microstructure characteristics of the weathered quartz gravel material. Plates l(A) and (B) are images of well-crystallised morphology of flaky arrangements of clay particles, which is montmorillonite and kaolinite as matrix between the fine grains. The concentration and the distribution of the clay particles

---

**Table 3** In situ dry density and moisture content results of the experimental panels

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Date constructed</th>
<th>Date tested</th>
<th>Maximum dry density (kg/m³)</th>
<th>Optimum moisture content (%)</th>
<th>In situ dry density (kg/m³)</th>
<th>Moisture content (%)</th>
<th>Difference in moisture content (%)</th>
<th>Relative compaction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>8 May 2008</td>
<td>20-05-08</td>
<td>2 107</td>
<td>7.4</td>
<td>2 028</td>
<td>4.4</td>
<td>+41</td>
<td>96.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>2 024</td>
<td>7.4</td>
<td>-18</td>
<td>91.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>1 975</td>
<td>3.6</td>
<td>-106</td>
<td>93.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>06-01-09</td>
<td>2 107</td>
<td>7.4</td>
<td>1 971</td>
<td>3.9</td>
<td>+8</td>
<td>93.5</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>8 May 2008</td>
<td>20-05-08</td>
<td>2 107</td>
<td>7.4</td>
<td>1 926</td>
<td>4.6</td>
<td>-18</td>
<td>91.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>2 067</td>
<td>3.9</td>
<td>-3</td>
<td>98.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>1 997</td>
<td>3.8</td>
<td>-3</td>
<td>94.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>06-01-09</td>
<td>2 107</td>
<td>7.4</td>
<td>1 959</td>
<td>3.7</td>
<td>-3</td>
<td>93.0</td>
</tr>
<tr>
<td>Control (untreated)</td>
<td>8 May 2008</td>
<td>20-05-08</td>
<td>2 107</td>
<td>7.4</td>
<td>1 842</td>
<td>3.9</td>
<td>+38</td>
<td>87.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>2 021</td>
<td>6.3</td>
<td>-152</td>
<td>95.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24-07-08</td>
<td>2 107</td>
<td>7.4</td>
<td>2 066</td>
<td>2.5</td>
<td>-108</td>
<td>98.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>06-01-09</td>
<td>2 107</td>
<td>7.4</td>
<td>2 046</td>
<td>1.2</td>
<td>-108</td>
<td>97.1</td>
</tr>
</tbody>
</table>

---

Plate 1 SEM images of untreated material

Plate 2 SEM images of enzyme additive treated material
(about 1 μm) are not regular; some parts contain more clay particles and less grains and other contain more grains and less clay particles. Individual quartz particles are more difficult to identify due to the amount of clay, but these quartz particles, which are free of clay coating, are clearly seen (indicated by ellipses, circles and arrows in images (B) and (C)). The quartz particles in the images seem more rounded. This means that the particles have been transported further, allowing more erosion of the particles. The image in (D) shows the individual clay particles against the background of a quartz particle. This matrix can be related to the plasticity characteristics of the quartz material. The SEM images show a clear crystallised surface morphology of the untreated material.

**Enzyme additive treated material**

Plates 2(A), (B) and (C) are images of the enzyme additive treated material at one, three and five months respectively; after construction and despite the different magnifications, they all appear generally similar. All these images show cornflake-like grains (indicated by arrows, circles, and ellipses) dispersed throughout the images. The enzyme additive seems to have coated the surface of the soil particles, though it is difficult to identify the inter-particle bonding. Plate 2(D) is the image at eight months after construction and shows bonding along the edges of clay flakes (indicated by ellipses). This confirmed the enzyme additive characteristics of forming inter-particle bonding. A homogeneous microstructure is developed which may indicate increase in strength. Micro cracks were not evident.

**Sulphonated oil additive treated material**

Plates 3(A) and (B) are the images of the sulphonated oil additive treated material at one and three months after construction respectively, and show the inter-particle bonding (indicated by arrows) forming a dense microstructure. Larger quartz particles seemed not to have bonded (indicated by ellipse), but smaller particles indicated the tiny forms of dense matrix (indicated by arrows), thus resulting in the occurrence of a homogeneous microstructure. Plate 3(C) is the image at five months after construction and shows a heterogeneous structure that consists of larger bonded particles (indicated by zones X and Y in the ellipses). This resulted in the occurrence of a homogeneous microstructure, thereby (an indication of) increasing strength. Plate 3(D) is the image eight months after construction and shows a complete bonded matrix to increase the strength (indicated by arrows). Micro cracks were also not evident.

During sample preparation for the SEM, a small specimen is broken from a larger mass of material. This can affect the surface morphology observed under the SEM. Surface morphology was predominant in the microstructure of the treated material of the study.

**Analysis of the in situ density according to relative compaction obtained**

The in situ dry density, or field density, and in situ moisture tests, where the oven-dried moisture was used, are presented in Table 3 and graphically depicted in Figure 1. Relative compaction results of the experimental panels are depicted graphically in Figure 2.
The purpose of the density tests was to ensure that the treated panels were compacted to the recommended target of 95% of Modified AASHTO density as per TRH20 (CSRA 1990) recommendation guidelines. The same compaction was applied to all the panels during construction. The Modified AASHTO density of the treated panels of the study ranged from 91% to 94%. The additives may have had an effect on the relative compaction obtained, and this may indicate that these additives may possibly be used as compaction aids.

The in situ dry density tested on 24 July 2008, almost three months after construction, showed a slight decrease of 0.2% in the enzyme additive treated panel, and an increase of 6.8% in the sulphonated oil additive treated panel. Further compaction in the sulphonated oil additive treated panel may be attributed to traffic. The average relative compaction density test results (Modified AASHTO density) eight months after construction revealed that enzyme additive and sulphonated oil additive treated panels had densities of 93.5% (a decrease of 2.91%) and 93.0% (a slight increase of 1.75%) respectively. The control (untreated) panel had a density of 97.1% (a significant increase of 11.1%).

Both studies of Moloisane (2009) and Van Veelen (2005) were compared for the treated in situ and soaked strength behaviour. The dark-reddish brown gravel material of the Moloisane (2009) study was compared with the Putfontein clay material of the Van Veelen (2005) study. These two unpaved road materials were characterised to be closely matching physically. A summary of the comparison of the in situ relative compaction of the treated panels is presented in Table 4.

Though the data was limited, the enzyme additive treated materials indicated that there might be an on-going improvement, and the sulphonated oil additive treated materials indicated that there would be no change. The best improvement was in the control (untreated) panel.

### Table 4 Summary of relative compaction per experimental panel

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Moloisane (2009) study</th>
<th>Van Veelen (2005) study</th>
<th>Comments on the bonding comparison</th>
<th>Rating¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>2009 95% 2005 90% 2009 93%</td>
<td>Improvement</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>2009 94% 2005 93% 2009 93%</td>
<td>No change</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Control (untreated)</td>
<td>2009 94% 2005 90% 2009 96%</td>
<td>Improvement</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>

¹ Strength development behaviour over time was the primary objective of the study:
– The additive was rated “good” if the relative compaction of that particular treated panel measured in 2009 was higher than (improvement) the relative compaction measured in 2005, or the 2005 results were the same as the 2009 results.
– The additive was rated “poor” if the relative compaction of that particular treated panel measured in 2009 was less than (reduction) the relative compaction measured in 2005.

The in situ dry density tested on 24 July 2008, almost three months after construction, showed a slight decrease of 0.2% in the enzyme additive treated panel, and an increase of 6.8% in the sulphonated oil additive treated panel. Further compaction in the sulphonated oil additive treated panel may be attributed to traffic. The average relative compaction density test results (Modified AASHTO density) eight months after construction revealed that enzyme additive and sulphonated oil additive treated panels had densities of 93.5% (a decrease of 2.91%) and 93.0% (a slight increase of 1.75%) respectively. The control (untreated) panel had a density of 97.1% (a significant increase of 11.1%). Both studies of Moloisane (2009) and Van Veelen (2005) were compared for the treated in situ and soaked strength behaviour. The dark-reddish brown gravel material of the Moloisane (2009) study was compared with the Putfontein clay material of the Van Veelen (2005) study. These two unpaved road materials were characterised to be closely matching physically. A summary of the comparison of the in situ relative compaction of the treated panels is presented in Table 4.

Though the data was limited, the enzyme additive treated materials indicated that there might be an on-going improvement, and the sulphonated oil additive treated materials indicated that there would be no change. The best improvement was in the control (untreated) panel.

### Table 5 In situ CBR of the experimental control (untreated) panel

<table>
<thead>
<tr>
<th>Date tested</th>
<th>Control (untreated) panel position</th>
<th>In situ relative compaction (%)</th>
<th>DCP-CBR</th>
<th>Soaked laboratory CBR Modified AASHTO density</th>
<th>Approximate TRH14 (1985) classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>In situ</td>
<td>Soaked</td>
<td>At 90%</td>
</tr>
<tr>
<td>21 May 2008</td>
<td>11</td>
<td>–</td>
<td>118</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>11 June 2008</td>
<td>11</td>
<td>–</td>
<td>107</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>18 July 2008</td>
<td>11</td>
<td>–</td>
<td>242</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>19 July 2008</td>
<td>11</td>
<td>94.6</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>24 July 2008</td>
<td>11</td>
<td>98.1</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>15 Aug 2008</td>
<td>11</td>
<td>–</td>
<td>261</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>12 Sept 2008</td>
<td>11</td>
<td>–</td>
<td>293</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>09 Oct 2008</td>
<td>11</td>
<td>–</td>
<td>235</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>06 Jan 2009</td>
<td>11</td>
<td>96.1</td>
<td>84</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>17 Nov 2010</td>
<td>11</td>
<td>–</td>
<td>139</td>
<td>105</td>
<td></td>
</tr>
<tr>
<td>24 Jan 2012</td>
<td>11</td>
<td>–</td>
<td>53</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>–</td>
<td>70</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>22 Feb 2012</td>
<td>Between 6 &amp; 7</td>
<td>–</td>
<td>42</td>
<td>102</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Between 7 &amp; 8</td>
<td>–</td>
<td>36</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Between 8 &amp; 9</td>
<td>–</td>
<td>47</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Between 9 &amp; 10</td>
<td>–</td>
<td>73</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Between 10 &amp; 11</td>
<td>–</td>
<td>64</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>–</td>
<td>53</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>
in the field took place over a period of two hours and the water depth was 50 mm in the cleaned open-quarter oil drum. The surface may have been well-sealed by traffic preventing ready water ingress. Furthermore the density was greater than 95% Modified AASHTO. In the laboratory test material larger than 19 mm is discarded, whereas in the field the coarser material provides strength. Two soaked CBR values of 51% and 84% obtained in February 2012 were higher than the in situ CBR values of the same month. This unlikely situation was attributed to the fact that the panels were tested under relatively wet climatic conditions, because February is in the rainy summer season (mid-October to mid-February). The DCP results, when converted to the CBR, only give an indication of the CBR of the material and not the exact CBR, hence there was poor correlation between the field and the laboratory testing.

**In situ strength development analysis**

The laboratory CBR-test quantitatively determines the bearing capacity of the material, where the shearing resistance of a material under carefully controlled conditions of density and moisture is measured. The strength of the treated panels in both in situ and soaked conditions was measured by means of the DCP to yield the correlated CBR values. The DCP has been increasingly used in many parts of the world in soil, granular material and lightly treated soils to assess the in situ pavement strength conditions (Scala 1956; Van Vuuren 1969; Kley 1975; Kley & Savage 1982; De Beer et al 1989; Burnham & Johnson 1993; Newcomb et al 1995; Salgado & Yoon 2003; Van Veelen 2005; Van Veleen & Visser 2007).

The theory behind the DCP is that the resistance to penetration of a steel 60° cone through the material is an indication of the in situ strength of the material, as measured by the CBR. The output of the DCP test is a penetration rate, expressed in millimetres per blow (DN), which is also called Dynamic Cone Penetration Index. Both the in situ and soaked CBRs were measured each month on the panels, with the exception of the two months before the eighth month (last month of the study) due to rainfall. The DCP tests were again carried out after three years. The in situ CBR was also tested next to the wetted area to provide the in situ DCP. The existing panels constructed by Van Veelen (2005) were again carried out after three years. The DCP results, when converted to the CBR, only give an indication of the CBR of the material and not the exact CBR, hence there was poor correlation between the field and the laboratory testing.

**Table 6 DCP-CBR strength development behaviour of the treated experimental panels over a five-month period**

<table>
<thead>
<tr>
<th>Additive used</th>
<th>11-06-08 In situ</th>
<th>11-06-08 Soaked</th>
<th>18-07-08 In situ</th>
<th>18-07-08 Soaked</th>
<th>15-08-08 In situ</th>
<th>15-08-08 Soaked</th>
<th>12-09-08 In situ</th>
<th>12-09-08 Soaked</th>
<th>9-10-08 In situ</th>
<th>9-10-08 Soaked</th>
<th>17-11-10 In situ</th>
<th>17-11-08 Soaked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>36</td>
<td>19</td>
<td>64</td>
<td>57</td>
<td>30</td>
<td>26</td>
<td>76</td>
<td>49</td>
<td>44</td>
<td>7</td>
<td>89</td>
<td>108</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>78</td>
<td>32</td>
<td>74</td>
<td>45</td>
<td>80</td>
<td>37</td>
<td>71</td>
<td>20</td>
<td>85</td>
<td>41</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td>Control (untreated)</td>
<td>107</td>
<td>72</td>
<td>242</td>
<td>92</td>
<td>261</td>
<td>142</td>
<td>293</td>
<td>70</td>
<td>325</td>
<td>67</td>
<td>139</td>
<td>105</td>
</tr>
</tbody>
</table>

**Figure 3 In situ DCP-CBR strength development behaviour of the treated panels over a five-month period**

**Figure 4 Soaked DCP-CBR strength development behaviour of the treated panels over a five-month period**

![Graph](https://via.placeholder.com/150)

The expression in Equation 1 as given by Kley *et al* (1989). The DCP represent multiple point measurements of the road-bearing strength, which are averaged.

\[
\text{CBR} = 410 \times (\text{DN})^{-1.27}
\]

(1)

Where:

- CBR = In situ CBR strength (%)
- DN = Penetration rate (mm/blow)

**Strength development behaviour analysis five months after construction**

The DCP-CBR results of the panels’ strength behaviour over five months after construction are presented in Table 6 and graphically depicted in Figures 3 and 4 for trend observation.
Table 7 Summary of strength behaviour of the treated experimental panels versus control (untreated) panel in the year 2008

<table>
<thead>
<tr>
<th>Additive used</th>
<th>2008 Maximum in situ strength</th>
<th>Comments on the strength</th>
<th>2008 Maximum soaked strength</th>
<th>Comments on the strength</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Month</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Month</td>
</tr>
<tr>
<td>Enzyme</td>
<td>September</td>
<td>293</td>
<td>76</td>
<td>Reduced</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>October</td>
<td>325</td>
<td>85</td>
<td>Reduced</td>
</tr>
</tbody>
</table>

Table 8 Summary of strength behaviour of the treated experimental panels versus control (untreated) panel in the year 2010

<table>
<thead>
<tr>
<th>Additive used</th>
<th>2010 Maximum in situ strength</th>
<th>Comments on the strength</th>
<th>2010 Maximum soaked strength</th>
<th>Comments on the strength</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Month</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Month</td>
</tr>
<tr>
<td>Enzyme</td>
<td>November</td>
<td>139</td>
<td>89</td>
<td>Reduced</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>November</td>
<td>139</td>
<td>22</td>
<td>Reduced</td>
</tr>
</tbody>
</table>

The enzyme additive treated panel had a higher in situ DCP-CBR value of 76% in September 2008 at four months after construction, while the sulphonated oil additive treated panel had higher in situ DCP-CBR value of 80% in August 2008 at three months after construction, and had approximately the same value in the fifth month. In July 2008 at two months after construction, the enzyme additive treated panel reached its maximum soaked DCP-CBR value of 57% compared to the sulphonated oil additive treated panel with a maximum soaked DCP-CBR value of 45%, and had approximately the same value in the fifth month. Although in July 2008 at two months after construction, the enzyme additive treated panel had the soaked DCP-CBR value of 57%, which was regarded as maximum in the five-month period, the in situ DCP-CBR value in the same month was 64%, higher than the soaked. The results are consistent as the in situ DCP-CBR is higher than the soaked value. In November 2010 at approximately 31 months after construction, the enzyme additive treated panel had a significant increase in both the in situ and soaked DCP-CBR strength, while the sulphonated oil additive treated panel had a significant decrease in both the in situ and soaked DCP-CBR strength.

The control (untreated) panel indicated the highest in situ DCP-CBR strength in October 2008, five months after construction, while the maximum soaked DCP-CBR strength was indicated in August 2008, three months after construction. In November 2010, approximately 31 months after construction, the in situ DCP-CBR strength was higher than the treated panels, while the soaked DCP-CBR strength was approximately the same as that of the enzyme additive treated panel, and much higher than that of the sulphonated oil additive treated panel. From the time series it does appear as if the additives have a negative influence on strength over time. From the time sequence the DCP-CBR strength values were variable; an individual result should not be used to draw conclusions. These ‘actually very similar’ values were attributed to the conditions, density and moisture variations, and the fact that the results are based on single tests.

The on-going strength developed during the first five months was ascribed to the fact that from June to September 2008 there was no rainfall and therefore the testing was conducted under dry conditions. It only rained about 17 mm in June 2008, with the most (about 14 mm) measured on 3 and 4 June; the remaining 3 mm fell on the 20th. During the eighth month the panels indicated a significant decrease in the in situ and soaked DCP-CBR strength (it should be noted that the road was probably nearing its wettest condition at that time of the year). That was attributed to the amount of rain that fell on the panels from October to December 2008, which prevented the panels from drying out to reach their maximum strength. During the period October to December 2008, about 285 mm of rain fell on the panels, which was high for the area. According to South African Explorer (2008), the area normally receives an average of about 116 mm of rain during this period. The experimental panels were constructed in May 2008, when about 185 mm of rain fell. According to South African Explorer (2008), the area normally receives about 579 mm of rain per year, but in 2008 it received about 785 mm of rain. The panels were therefore constructed and tested under relatively wet climatic conditions. The only testing that was conducted under dry conditions was during July to September 2008 (months two to four when no rain fell).

Strength development behaviour of the treated experimental panels versus the control (untreated) panel

The maximum in situ and soaked DCP-CBR of each treated panel were compared with the in situ and soaked DCP-CBR of the control (untreated) panel for the same months. The comparison procedure was such that even if the control (untreated) panel may have had a higher strength in another month, that value was not considered for the comparison. That was done to compare the true conditions of the same month, because conditions vary. With reference to the studies conducted by Van Veelen (2005) and Van Veelen & Visser (2007), an improvement was considered when the treated panels had a DCP-CBR of more than 10% greater than the control (untreated) panel. A summary of the maximum in situ and soaked DCP-CBR of the treated panels, compared with the maximum in situ and soaked DCP-CBR of the control (untreated) panel for the same month in the year 2008 is presented in Table 7.

In the 2008 results, both enzyme and sulphonated oil additive treated panels showed a decrease in strength compared with the control (untreated) panel in both in situ and soaked conditions. A summary of the maximum in situ and soaked DCP-CBR of the treated panels, compared with the maximum in situ and soaked DCP-CBR of the control (untreated) panel for the same month in the year 2010 is presented in Table 8.

In the 2010 results, the enzyme additive treated panel showed a decrease in strength compared with the control (untreated) panel in situ conditions only and a slight increase in soaked conditions. The sulphonated oil additive treated panel showed a decrease in strength compared with the control (untreated) panel in both in situ and soaked conditions. This decrease in strength is evidence that those additives did not perform well under wet conditions. The panels were constructed and tested during the rainy season. It is hypothesised that this poor performance may be attributed to the fact that none of the panels had time to dry out sufficiently for substantial strength.
improvement to occur after construction. It may also be concluded that the additives performed poorly because they are not totally compatible with the specific type of material, although the material did contain a small amount of the necessary smectite clays. It may also be considered that the low application rates (typically 0.03 ℓ/m² for stabilisation and 0.01 ℓ/m² as a compaction aid) were insufficient to treat the material.

Comparison of the strength development behaviour of the treated experimental panels

Both the studies of Moloisane (2009) and Van Veelen (2005) were compared for the treated in situ and soaked strength behaviour. The dark-reddish brown gravel material of the Moloisane (2009) study was compared with the Putfontein clay material of the Van Veelen (2005) study. These two unpaved road materials were characterised to be closely matching physically. A summary of the comparison of the in situ strength behaviour of the treated panels is presented in Table 9.

From the comparative information of the “in situ strength behaviour”, the enzyme and sulphonated oil additive treated panels indicated that the strength development would deteriorate. A summary of the comparison of the “soaked strength behaviour” of the treated panels is presented in Table 10.

From the comparative study of the “soaked strength behaviour”, the enzyme additive treated panel indicated that the strength would improve, while the sulphonated oil additive treated panel indicated that the strength development would deteriorate. This does not necessarily mean that the sulphonated oil additive performed poorly. The rating was just an indication of comparison, based on the conditions that existed during the monitoring.

Overall performance of the treated experimental panels

A summary of the overall behaviour and performance of the treated panels based on all the tests conducted over the 8 and 31 months period of the studies is provided in Table 11.

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Moloisane (2009) study</th>
<th>Van Veelen (2005) study</th>
<th>Comments on the strength behaviour comparison</th>
<th>Rating1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
</tbody>
</table>

1 Strength development behaviour over time was the primary objective of the study:

- The additive was rated “good” if the in situ DCP-CBR of that particular treated panel measured in 2009 (that is, in both studies of Moloisane (2009) and Van Veelen (2005)) was higher than (improvement) or the same as (no change) the DCP-CBR of the control (untreated) panel. The additive was also rated “good” if the treated in situ DCP-CBR measured in 2009 of the Van Veelen (2005) study was higher than (improvement) the DCP-CBR results of the Moloisane (2009) study.

- The additive was rated “poor” if the in situ DCP-CBR of that particular treated panel measured in 2009 (that is, in both studies of Moloisane (2009) and Van Veelen (2005)) was less than (reduction) the DCP-CBR of the control (untreated) panel. The additive was also rated “poor” if the treated in situ DCP-CBR measured in 2009 of the Van Veelen (2005) study was less than (reduction) the DCP-CBR results of the Moloisane (2009) study.

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Moloisane (2009) study</th>
<th>Van Veelen (2005) study</th>
<th>Comments on the strength behaviour comparison</th>
<th>Rating1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
</tbody>
</table>

1 Strength development behaviour over time was the primary objective of the study:

- The additive was rated “good” if the in situ DCP-CBR of that particular treated panel measured in 2009 (that is, in both studies of Moloisane (2009) and Van Veelen (2005)) was higher than (improvement) or the same as (no change) the DCP-CBR of the control (untreated) panel. The additive was also rated “good” if the treated in situ DCP-CBR measured in 2009 of the Van Veelen (2005) study was higher than (improvement) the DCP-CBR results of the Moloisane (2009) study.

- The additive was rated “poor” if the in situ DCP-CBR of that particular treated panel measured in 2009 (that is, in both studies of Moloisane (2009) and Van Veelen (2005)) was less than (reduction) the DCP-CBR of the control (untreated) panel. The additive was also rated “poor” if the treated in situ DCP-CBR measured in 2009 of the Van Veelen (2005) study was less than (reduction) the DCP-CBR results of the Moloisane (2009) study.

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Moloisane (2009) study</th>
<th>Van Veelen (2005) study</th>
<th>Comments on the strength behaviour comparison</th>
<th>Rating1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enzyme</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
<td>Control DCP-CBR</td>
<td>Treated DCP-CBR</td>
</tr>
</tbody>
</table>

1 Target was 95% of Modified AASHTO density.
2 Based on the effect the additive had on the control (untreated) material.
3 Strength development behaviour over time was the primary objective of the study:

- The additive was regarded “good” if it did improve both in situ and soaked conditions.

- The additive was regarded “poor” if it did reduce both in situ and soaked conditions.

TABLE 11 Summary of the overall results of the behaviour and performance of the treated experimental panels

<table>
<thead>
<tr>
<th>Additive used</th>
<th>Ranking based on the testing applied</th>
<th>Overall rating with focus on the strength behaviour3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average relative compaction1</td>
<td>Average maximum strength2</td>
</tr>
<tr>
<td></td>
<td>In situ</td>
<td>Soaked</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Above target</td>
<td>Reduced</td>
</tr>
<tr>
<td>Sulphonated oil</td>
<td>Below target</td>
<td>Reduced</td>
</tr>
</tbody>
</table>

1 Target was 95% of Modified AASHTO density.
2 Based on the effect the additive had on the control (untreated) material.
3 Strength development behaviour over time was the primary objective of the study:

- The additive was regarded “good” if it did improve both in situ and soaked conditions.

- The additive was regarded “poor” if it did reduce both in situ and soaked conditions.

Journal of the South African Institution of Civil Engineering • Volume 56 Number 1 April 2014

37
oil additive treated panels were rated as “poor”. The rating was just an indication of comparison, based on the conditions that existed during the monitoring. The variability and inconsistency of the results is typical of this type of study. It should be noted that limited testing is generally carried out (example, single DCP and density tests per panel) during this type of work. The natural variability of the materials used in this type of study is generally high and the precision of the test method is typically quite low. On this basis, it is usually difficult to draw definite conclusions, but general indications of trends emerged.

CONCLUSIONS
The degree of formation and paste surrounding the particles appeared to have varied with time and differed between the enzyme and sulphonated oil additives. It was hypothesised from the SEM images that the increase in strength was due to cementation and pore filling by the additives. An increase in density in the sulphonated oil additive treated panel occurred three months after construction, and a further increase was again noticed eight months after construction (five months thereafter). Up to eight months after construction, the enzyme additive treated panel showed a significant decrease in density, but showed a slight increase 31 months after construction. The variations in density were attributed to testing variability. In the comparison per additive, the enzyme and sulphonated oil additive treated panels did not indicate much difference in relative compaction. This is because both additives are effectively compaction aids. The overall conclusion was that an increase in the densities of the Van Veelen (2005) study over the eight and 40 months periods occurred, although the panels were not trafficked. The Moloiianse (2009) study of eight months duration, which was trafficked, also had an increase in the densities. This increase in densities in the latter case might probably be because of further densification by traffic.

In the in situ and soaked DCP-CBR strength measurements, the sulphonated oil additive treated panel reached its maximum in situ strength at two months after construction, while the enzyme additive treated panel reached its maximum in situ strength at two months after construction. The decrease was attributed to the rain. In the comparison of the treated versus control (untreated) of the panels, the control showed the maximum in situ DCP-CBR strength in the fourth month, while the maximum soaked DCP-CBR strength was shown in the second month. The fact that the “control (untreated) panel” varied in performance is indicative of problems arising from the limited amount of testing typically allocated to projects of this type. In the comparison per additive, the in situ long-term DCP-CBR strength development behaviour of both the enzyme and sulphonated oil additive treated panels rated deterioration. The rating was just used as an indication of comparison. There was, however, little evidence to show that the additives improved the material, with the control panel being consistently stronger in both the in situ and soaked DCP-CBR conditions.

The results highlighted the importance of considering the time factor in the strength development of non-traditional soil stabilisation. It is clear that the variability and inconsistency of the results are typical of this type of study. It should be noted that limited testing is generally carried out (example, single DCP, and density tests per panel) during this type of study. The natural variability of the materials used in this type of study is generally high and the precision of the test method is typically quite low. On this basis, it is usually difficult to draw definite conclusions.

ACKNOWLEDGEMENTS
This paper is based on the first author’s MSc in Applied Science project report submitted to the University of Pretoria. The National Department of Transport’s Northern Transportation Centre of Development and Tshwane University of Technology are fully thanked for financial support. Ekurhuleni Metropolitan Municipality and its personnel to the University of Pretoria. The National Institute for Transport and Road Research, CSIR, Technical Methods for Highways.

REFERENCES
Bergmann, R 2000. Soil stabilisers on universally accessible trails. Department of Transportation, Federal Highway Administration (FHWA), USA.
Birst, S & Hough, J 1999. Chemical additive usage on unpaved roads in the mountain plain states. Fargo, North Dakota: North Dakota State University, Upper Great Plains Transportation Institute, USA.
Bolander, P W 1997. Chemical additives for dust control: What we have used and what we have learned. Transportation Research Record: Journal of the Transportation Research Board (TRB), No 1589, pp 42–49.
Brazetti, R & Murphy, S 2004. Results and conclusions from a comprehensive World Bank study of the performance of roads rehabilitated with soil stabilisers in Paraguay, South America. Paper presented at the 83rd Annual Meeting of the Transportation Research Board (TRB), Washington D.C., USA.


The status of basic design ground motion provisions in seismic design codes of sub-Saharan African countries

A critical review

A. Worku

Basic provisions for design ground motions in seismic design codes of sub-Saharan African countries are critically reviewed. The seismic codes of Ethiopia, Kenya and Uganda are selected to represent the eastern region, Ghana to represent the west and South Africa to represent the south. The specific provisions considered are those pertaining to site effect and the recurrence period of the design earthquake. The codes are also compared with one another and with selected current international codes from the US and Europe, with respect to selected provisions. The provisions are further viewed from the perspective of the state of the art and the state of the practice. It has been concluded that these basic provisions in most of the sub-Saharan African codes considered are inadequate in guaranteeing safety of human life and limiting damage to property, suggesting a need for immediate updating, an exception being the South African code.

INTRODUCTION

According to the recent compilation of worldwide seismic regulations by the International Association of Earthquake Engineering (IAEE), 63 countries have issued seismic design codes as of October 2008. Five African countries are included in this list: Algeria, Egypt, Ethiopia, Ghana and Uganda. Copies of their seismic regulations are available on the IAEE website (http://www.iaee.or.jp/), although some are not necessarily up to date. This list is obviously not exhaustive, as countries like South Africa, Kenya and Ethiopia are known for a moderate degree of seismic activity and have visited frequently by strong earthquakes, including the magnitude 7.0 earthquake (moment magnitude, \(M_w\)) that rattled Mozambique as recently as 2006 (Sousa et al 2008). In 1990 an earthquake of a surface-wave magnitude of 7.2 shook the southern part of Sudan (currently South Sudan) accompanied by numerous aftershocks with magnitudes as large as 7.0 (Girdler & McConnell 1994). Such sizes of shallow (< 35 km focal depth) earthquakes have a destructive potential to the near field comparable to the 2010 Haiti earthquake that devastated the capital Port au Prince, killing over 300 000 people, displacing 1.3 million and destroying close to 100 000 houses (http://www.usgs.org/). Fortunately the earthquakes of Sudan and Mozambique occurred in uninhabited areas, resulting in only a few causalities. However, the human settlement and infrastructure development pattern in many places of Africa is changing fast. Cases in point are Juba, the capital of South Sudan; the Turkana area of northern Kenya, with recent developments related to oil discoveries; and southern Ethiopia, with on-going huge hydroelectric power, irrigation and oil prospecting projects.

The western half of the African continent is comparatively quieter. Some countries are known for a moderate degree of seismic vulnerability, including Guinea, Sierra Leone, Liberia, Ghana, Cameroon, Gabon and Congo. Ghana is the only nation in this region known to have seismic design regulations. Its first seismic code was issued in as early as 1977. This was revised in 1990, and again recently in 2010, even though the rather high peak ground accelerations (PGA) of up to 0.35 g specified by the latest editions (BRRI 1990, 2010) cast doubts on the reliability of the background seismic hazard assessment work.

An observation common to all sub-Saharan African countries is that awareness on imminent seismic risks in the region is surprisingly low. Countries having their own
seismic design regulations are quite few. Moreover, most of the codes are not regularly updated. Some basic provisions of the few codes available are obsolete and some are unrealistic. Cases in point include the Kenyan seismic code, which has never been updated since its issuance in 1973, whereas the knowledge of earthquake engineering has shown significant strides ever since. Also, most of the sub-Saharan African codes employ obsolete site-effect provisions that lead to unsafe design and that are no more in use in most parts of the world.

Quality control in the construction industry is in a dismal state. Building regulations, if available, are not often properly adhered to, even for static gravity loads. Reports of buildings collapsing during and after construction have become common news in Africa. In 2011 alone, the collapse of five buildings in different locations of Kenya was reported (http://www.a4architect.com/). In Kampala, the capital of Uganda, at least 60 people were reported dead from the collapse of 11 buildings within the recent few years (http://www.ugpulse.com/). In February this year, a 16-storey building under construction in Dar es Salaam, Tanzania, collapsed costing close to 40 lives. Equally tragic incidents have become a major concern in Nigeria (Oloyede et al 2010). Similar situations can be expected to exist in many other African countries where the building industry has been wrongly left as a business for a small elite of highly qualified people. The public control and enforcement of regulations is the major concern in Nigeria (Oloyede et al 2010). Similar situations can be expected to exist in many other African countries where the building industry has been wrongly left as an informal economic sector, with negligible public control and enforcement of regulations. Imagining how buildings of such quality would behave during strong earthquakes is extremely scary, to say the least.

A number of shortcomings can be cited of the rather few seismic codes available in some sub-Saharan African countries. Some of the fundamental pitfalls are related to the definition of the design ground motion. Primarily, there exists a lack of uniformity in the return period of the design earthquake specified by the codes. For example, the Ethiopian seismic code, EBCS 8, (MWUD 1995) provides for a return period of 100 years, whereas the South African code (SABS 2010) has adopted 475 years, while none of the other codes explicitly state the return period. An equally important issue is the fact that the approach adopted to account for the amplification potential of site soils in almost all these codes is obsolete, with the exception of the recently updated South African code (SABS 2010), which has adopted the recent approach introduced in the European code to account for site amplification.

This paper presents a study of these two fundamental issues related to input ground motion, namely site effect and recurrence period of design earthquake. It focuses on the pertinent provisions of the selected representative seismic codes. The corresponding provisions in these codes are compared with one another and with recent issues of American and European codes.

**BASIC GROUND MOTION PROVISIONS OF THE AFRICAN SEISMIC CODES**

The selected codes include the Ethiopian, Kenyan and Ugandan codes to represent the most seismic regions of the EARS, the Ghanaian code to represent the less seismic western Africa, and the South African code to represent the southern region. The provisions on the two basic issues are briefly presented in the following section, to be followed by comparisons among themselves and with selected recent American and European codes. The National Earthquake Hazard Reduction Program (NEHRP) is selected to represent the American codes (BSSC 2004), whereas the Eurocode (Eurocode 8 2004) automatically qualifies as representative of European codes.

**The Kenyan Code – 1973**

The Kenyan seismic code, probably one of the pioneers on the continent, was issued in 1973 by the Kenyan Ministry of Works and uses the Modified Mercalli intensity (MMI) scale to map the seismic hazard of the country, as shown in Figure 1 (MWK 1973). The map divides the country into four seismic zones: Zone V, VI, VII and VIII–IX, where the Roman numbers are in accordance with the MMI scale. This method of seismic hazard mapping using earthquake intensity is seldom in use nowadays for design purposes.

The code neither states the recurrence period of the design earthquake nor its probability of exceedance. Explicit numeric values of parameters like PGA, which are more relevant to engineering design, are not provided. However, these can be inferred from the relation provided by the code for estimating the seismic coefficient (see Equation (1)), and the accompanying explanation that the PGA corresponding to Zones VIII–IX, VII and VI are 0.05 g, 0.025 g and 0.0125 g respectively. According to the map, populous Kenyan cities within the eastern branch of the EARS, like Nakuru, belong to Zone VIII–IX; the
The code provides the following expression for the computation of the seismic coefficient in the equivalent static force (ESF) approach for a case it defines as the “basic” case:

\[ S_{e_{\text{basic}}} = \frac{0.05}{\sqrt{T}} \quad (1) \]

where \( T \) is the fundamental period of the building.

The “basic” case is defined as flexible-frame buildings built on a hard ground condition in Zone VIII–IX. The seismic coefficient of the “basic” case is halved for Zone VII and quartered for Zone VI. Though not explicitly stated in the code, it can be understood that the numerator in Equation (1) is the seismic factor, \( C \), or the PGA normalised with respect to the gravitational acceleration, \( g \), for the seismic zone under consideration.

With regard to site effect, the code crudely classifies sites into just two classes: hard and soft ground, without further satisfactory descriptions. The code suggests that \( S_e \) for hard ground be raised by 30% to account for site amplification due to a soft ground. By doing so, this approach disregards the inherent wide variation in the dynamic behaviour of natural soil deposits, and is not in agreement with even the earliest site-dependent response spectra devised in the 1970s, as will also be discussed later. The seismic coefficient, \( S_e \), in Equation (1) normalised with respect to the seismic factor, \( C \), is given in Figure 2 for the two site conditions.

Obviously, the curves do not have resemblance to design spectral curves specified even in old versions of known seismic design codes. The code does not have provisions for dynamic analysis of structures. Also, no account is made for inelastic response of structures. Despite being one of the pioneering seismic codes in Africa, the code has not been updated since it had first been issued about four decades ago. Ironically, Kenya is located well within the eastern branch of the active seismic region of the EARS which is prone to strong earthquakes.

**The Ethiopian Code – EBCS 8, 1995**

The building design code of Ethiopia was first introduced in 1978. Its seismic provisions have been revised twice since then. The first revision took place in 1983. The current version – the Ethiopian Building Code Standard, EBCS 8: 1995 – dedicates a separate volume for seismic provisions (MWUD 1995). A committee, of which this author is a member, has been formed very recently and entrusted with the task to revise the code for a third time.

EBCS 8:1995 provides the seismic hazard map of the country given in Figure 3, which is based on a 100-year return period. According to this map, each of Zones 1, 2, 3 and 4 is assigned a constant bedrock acceleration ratio, \( \alpha \), of 0.03, 0.05, 0.07 or 0.1 respectively, whereas Zone 0 is considered seismic free. The capital, Addis Ababa, home to the headquarters of many international bodies, including the African Union, belongs to Zone 2, with \( \alpha_0 = 0.05 \). The boundary with the more seismic neighbouring region of Zone 3, with \( \alpha_0 = 0.07 \), is only 20 km away from the city centre and is already within the city due to the recent rapid urban expansion. The factor \( \alpha_0 \) is the PGA at bedrock level normalised with respect to the gravitational acceleration \( g \). It is used to scale down the normalised design spectra provided by the code. Several large towns, including capitals of federal states, belong to the most seismic region of Zone 4 with \( \alpha_0 = 0.1 \).

EBCS 8 specifies the seismic coefficient \( S_{e_{\text{d}}} \) for the equivalent static force (ESF) method in the form of Equation (2) for the design base shear \( F_d \):

\[ F_d = S_{e_{\text{d}}} W = (\alpha_0 \beta_0) W \quad (2) \]
where $I$ is the importance factor of the building that is assigned values of 0.8 to 1.4; $W$ is the weight of the building; $\beta$ is the design response factor of the structure that accounts for site effect; and $\gamma$ is the behaviour factor which accounts for ductility, non-linear response and the influence of damping other than 5%.

The influence of the site soil is embedded in the design response factor (or elastic design spectrum) given by

$$\beta = \frac{1.2 S}{T^{2/3}} \leq 2.5 \quad (3)$$

$S$ is the site coefficient, which can assume the values of 1.0, 1.2 or 1.5 for Subsoil Class A, B or C, respectively. Subsoil Class A includes rock or similarly competent formations and dense deposits of sand, gravel or over-consolidated clay characterised by $v_s \geq 800$ m/s; Subsoil Class B consists of deep deposits of medium dense sand, gravel or medium stiff clays with thicknesses of at least several tens of metres and $v_s \geq 200$ m/s; Subsoil Class C comprises loose cohesionless soil deposits and soft to medium stiff cohesive soils with $v_s < 200$ m/s; and $v_s$ is the average shear-wave velocity of the geological formation of the site. Plots of Equation (3) are given in Figure 4(a).

EBCS 8 also provides the site-dependent elastic design spectra of Figure 4(b) for dynamic analysis which are not identical to the spectra of Figure 4(a). Whereas the right-hand segment varies according to $1/T^{2/3}$ in Figure 4(a), it varies in accordance with $1/T$ in the spectra of Figure 4(b). Similarly the amplification factors in Figure 4(b) for Classes B and C are 1.5 and 2.25 instead of 1.2 and 1.5. The curves in Figure 4(b) are almost identical to the ATC-3 Spectra, which are based on the early findings of the empirical studies of Seed et al (1976) presented in Figure (11) and will be discussed later.


The seismic code of Uganda, US 319:2003, issued by the Ugandan National Bureau of Standards in 2003, specifies the following design spectrum, $C_d$, for both the ESF and modal analysis (UNBS 2003):

$$C_d(T) = C(T)ZI$$

where $C(T)$ is the ordinate of the “basic” response spectrum given for three different soil groups as shown in Figure 5; $Z$ is the seismic zone factor given for the three seismic zones of the country shown in Figure 6 assuming the values of 1, 0.8 and 0.7 for Zones 1, 2 and 3 respectively; $I$ is the structural importance factor that takes values of 1 up to 2; and $K$ is the structural performance factor which is dependent on the type of the structural system and materials. The recommended minimum values of $K$ vary from 1 for the most carefully designed ductile structural systems to 4 for structural systems of minimal ductility.

The design spectrum in Figure 5, which is defined as the “basic” design spectrum by the code, is the minimum design inelastic
spectrum specified for the most seismic zone, Zone 1, and for the most ductile structural system, for which $K$ is assigned the value of unity. In contrast to this, the basic design spectrum in most seismic design codes is conventionally defined as the elastic design spectrum before any factor is applied to it to account for inelastic response and ductility. This corresponds to the spectrum obtained from Figure 5 by applying the maximum value of the structural performance factor of $K = 4$ specified for the least ductile system in the same seismic zone.

Thus, the maximum ordinate of the elastic design spectrum corresponding to the flat part of the spectral curves in Figure 5 obtained in this manner will have the value of 0.32. This ordinate is as a rule 2.5 times the rock-level normalised PGA (Newmark & Hall 1969). Therefore, the corresponding minimum rock-level PGA for Zone 1 can easily be back-calculated through dividing 0.32 by 2.5 to obtain 0.13 g. The corresponding minimum PGAs for Zone 2 and Zone 3 are obtained as 0.10 g and 0.09 g by applying the zone factors of 0.8 and 0.7 respectively, given by the code.

Comparison of these values with those in the recent continental GSHAP map of Africa presented in Figure 10 suggest that these PGA values tally with a 475-year return period of the design earthquake, even though the code does not mention the length of the recurrence period.

The code also fails to clearly specify the values of amplification factors associated with Subsoil Types II and III of Figure 5. Nor does it indicate how the descending segments on the right-hand side of the spectral curves vary with period. Nonetheless, using ordinates directly read from the curves at selected points, the amplification factors can be inferred to be 1.25 and 2.5 respectively. Even though these site-dependent design spectral curves look similar in shape to those of ATC-3, they also exhibit some differences in the values of the amplification factors, especially in the case of the softest soil class, for which the Ugandan code assigns a value of 2.5 instead of 2.25.

**The Ghanaian Code – 1990**

As in most sub-Saharan African countries, recorded strong earthquake ground motions in Ghana are nonexistent. A few studies conducted since recently indicate that the Accra metropolitan area is one of the most seismic-prone areas. Earthquake magnitudes of up to 6.5 on the Richter scale have been recorded (Amponsah 2004; Amponsah et al 2009; Allotey 2010; Oteng-Ababio 2012). Using a hybrid deterministic seismic hazard assessment technique, Amponsah et al (2009) simulated the 6.5 magnitude earthquake to predict peak horizontal and vertical ground accelerations in the ranges of 0.14 g to 0.57 g, and 0.2 g to 0.34 g respectively, with a maximum amplification factor of 4 for the horizontal motion in the period range of 0.2 s to 0.5 s. The highest amplifications are estimated at sites of unconsolidated or poorly consolidated deposits. Such magnitudes of shaking are quite large in the context of the known history of seismic activity in the African continent.

The latest Ghanaian code is the Code for the Seismic Design of Concrete Structures issued in 2010 by the Ghanaian Building and Road Research Institute (BRRI 2010). It provides the seismic hazard map shown in Figure 7(a), which divides the country into three seismic zones of 1, 2 and 3. Each zone is assigned a constant peak rock-surface acceleration of 0.15 g, 0.25 g and 0.35 g respectively. This hazard map seems to have been influenced a lot by the deterministic study of Amponsah et al (2009) that is based on a single event and appears to have already accounted for site amplification. Also, the recurrence period of the design earthquake is not stated, indicating the influence of the deterministic approach employed in the work of Amponsah et al (2009).

The rather large values of PGA specified by the code are comparable to values specified for the highly seismic western coastal region of the US (including California) proposed for a return period of 475 years. It is considered unlikely that these values apply to the comparatively less seismic region of West Africa, unless corroborated by more detailed probabilistic studies that account for the range of earthquake magnitudes expected to occur in the region.

As shown in Figure 10(a) of the GSHAP map for Africa, the maximum PGA in the entire continent for a return period of 475 years is 0.24 g, and the maximum possible PGA specifically for Ghana is 0.16 g, which is applicable to the zone containing Accra and its environs. This is based on a probabilistic seismic hazard assessment study conducted by Grünthal et al (1999).

The design spectra of the Ghanaian code are given in Figure 7(b). The spectra for PGAs less than 0.3 g and the corresponding amplification factors are similar to those of the Ethiopian seismic code, EBCS 8, given in Figure 4(b) or to the ATC 3, 1978 spectrum given in Figure 11(b), though the right-hand segment in the Ghanaian code...
varies according to $1/T^{2/3}$. Apparently the code uses the same spectra for both the ESF and dynamic methods of analysis. The code reduces the ordinate of the constant part from 2.5 to 2 for soil group S3 when the PGA is larger than 0.3 g, as shown in Figure 7(b). This measure seems to have been introduced in order to account for the increased damping in soils during high-intensity shaking that tends to reduce soil amplification, even though such a level of shaking is a very remote possibility in Ghana, as explained above.

The provisions for design ground motions in the 2010 edition (BRRI 2010) described above are practically identical to those in the 1990 edition (BRRI 1990) and do not account for recent findings on site amplification effects, which are discussed in a later section.

**The South African Code – SANS 10160-4:2010**

The recent revision of the South African standard for seismic actions, SANS 10160-4:2010 (SABS 2010), was issued in June 2010 and makes up one of the eight parts of the South African National Standard SANS 10160:2010 (Wium 2010), which replaces the older version of SABS 0160:1989.

Seismic hazard in SANS 10160-4:2010 is presented in the form of the seismic hazard map of Figure 8 in terms of the reference peak ground acceleration $a_g$ for rock site and a return period of 475 years. Notably, this return period has been introduced in 1989, when the earlier edition, SABS 0160, was issued. Two major zones are distinguishable in the map: Zone I of natural seismic activities and Zone II of predominantly mining-induced seismic activities. Zone I is assigned $a_g = 0.1$ g, whereas Zone II can assume larger values.

With regard to site effect, the code has directly adopted Type 1 Spectrum of the recent European code (Eurocode 8 2004) together with the four Ground Types A to D, omitting Ground Type E and softer sites. The corresponding elastic design spectra normalised with respect to the PGA are provided in Figure 9. The amplification factor $S$ varies from 1 to 1.35 depending on the ground type. The amplification factors in all cases are constant over the entire period range and are much smaller than those introduced in current US codes. Similarly, the transition periods, $T_B$ and $T_C$, are dependent on the ground type, whereas $T_D$, the transition to

---

**Figure 8** The seismic hazard map of South Africa according to SANS 10160-4:2010 (SABS 2010)

**Figure 9** Normalised elastic design spectra of SANS 10160-4:2010 (SABS 2010)

**Figure 10(a)** Seismic hazard map of Africa for a return period of 475 years according to GSHAP (data and plotting tool from http://gmo.gfz-Potsdam.de/)

**Figure 10(b)** Seismic hazard map of EARS region for a return period of 475 years according to GSHAP (data and plotting tool from http://gmo.gfz-Potsdam.de/)
the long-period range, is constant, taken as 2 seconds.

Given the relatively stable seismic nature of South Africa compared to other seismic-prone regions, especially those in the EARS, the stringent requirements of this code are quite exemplary for future revisions of seismic design codes of other sub-Saharan African countries.

**Recent seismic hazard map of sub-Saharan Africa**

The Global Earthquake Model (GEM) initiative is an on-going collaborative worldwide effort launched in 2009 with the aim of building a heightened public understanding and awareness of seismic risk, leading to increased earthquake resilience worldwide through sharing of earthquake data, models and knowledge; through application of GEM tools and software to inform decision-making for risk mitigation and management; and through expansion of the science and understanding of earthquakes (http://www.globalquakemodel.org). However, tangible results from this programme, like updated seismic hazard maps, will only be available in a couple of years’ time.

In contrast, the seismicity of the entire globe was compiled in the 1990s for the first time as part of the Global Seismic Hazard Assessment Program (GSHAP), which was launched by the International Lithosphere Program (ILP) and endorsed as a demonstration programme in the framework of the United Nations International Decade for Natural Disaster Reduction (UN/IDNDR). The project was active from 1992 to 1999 (http://www.seismo.ethz.ch/).

The GSHAP global seismic hazard map is given in terms of rock-surface PGA for a 10% exceedance in 50 years, which is equivalent to a return period of 475 years for the design earthquake. This level of hazard has been introduced in US codes since more than three decades ago, and subsequently adopted as a standard design-level hazard worldwide.

The data base of GSHAP is accessible to users (http://www.seismo.ethz.ch/). A seismic hazard map of Africa prepared using an online tool (http://gmo.gfz-Potsdam.de/) is given in Figure 10(a). A similar map of the EARS region – the most earthquake-prone region of Africa – is also given in Figure 10(b). Compared with these regional maps, the current local seismic hazard maps of some East African countries like Ethiopia and Kenya, given in Figures 1 and 3, underestimate the seismicity of their cities and towns by an average of about a half. PGAs of up to 0.24 g are assigned by GSHAP to some localities in East Africa, particularly to the Afar region of Ethiopia and to an area around Arusha in Tanzania. A number of populous towns and cities are also assigned PGAs up to 0.16 g (coloured yellow).

Kisumu, located on the northeastern shore of Lake Victoria and the third most populous city in Kenya, has probably seen one of the extreme discrepancies – from 0.0125 g in the local Kenyan code to 0.12 g in the GSHAP map – with a nearly tenfold difference. This would imply that, whereas seismic loading is practically negligible for the design of structures in Kisumu according to the existing Kenyan code, it is an important consideration if the GSHAP map is to be followed.

One of the reasons for such discrepancies is obviously the difference in the adopted return period of design earthquakes. Even though no return period is stated in the Kenyan code, it can be inferred to be not more than 100 years. In the Ethiopian code the return period is explicitly stated as 100 years. In contrast, the local seismic hazard map of Uganda in US 319:200 provides PGAs comparable with GSHAP hinting that the return period in this code is most likely 475 years, though not explicitly stated in the code.

According to the GSHAP map, the western side of the sub-Saharan Africa is not as seismic as the eastern side. The maximum PGA in the western region is 0.16 g, which is assigned to limited areas in countries like Gabon, Congo, Cameroon and Ghana (Figure 10(a)). Thus, the PGA of 0.15 g, 0.25 g and 0.35 g specified in the Ghanaian local code for its three seismic zones are obviously too large for a return period of 475 years. This indicates that the present local seismic map of Ghana significantly overestimates the seismicity and would result in unnecessarily conservative and expensive design.

Obviously, the seismic hazard map of Ghana needs reassessment, even though it was updated very recently, in 2010.

The return period in the South African code of SANS 10160–4:2010 is explicitly stated as 475 years. In fact, the seismic hazard map given in Figure 9 specifies PGA values a little in excess of those recommended by the GSHAP map (on the average by 25%).

**SITE EFFECT ON DESIGN SPECTRA**

**Site-dependent design spectra: early studies**

Figure 11(a) shows some of the earliest site-dependent average spectra from the pioneering work of Seed et al (1976). Figure 11(b) gives the ATC-3 simplified spectra introduced in 1978 for routine design purposes based on the curves of Seed et al (1976), whereby the fourth site class is omitted, as the associated data were not of satisfactory quality, as cautioned by the authors themselves.

The spectral curves for Soil Types S2 and S3 of Figure 11(b) in the descending branch are obtained by simply raising the spectral curve of Soil Type S1 (rock) by a single factor...
in each case. This factor, referred to as the ratio of response spectra (RRS), takes the values of 1.5 and 2.2 for Soil Types S2 and S3 respectively. The descending segments corresponding to the velocity-sensitive period range decline according to $T^{-1}$ (Dobry et al 2000). (The velocity-sensitive period range is the period range in which the velocity of the single-mass oscillator model, due to the ground motion, is amplified most, compared to its acceleration and displacement.)

The ATC-3 spectra were later integrated in the series of editions of US seismic codes including the National Earthquake Hazard Reduction Program (NEHRP) up to 1994, and in the Uniform Building Code (UBC) series up until 1997. In 1988 a fourth soil type, S4 for deep soft clays, was included with an increased amplification factor with the aim to address the rather high amplification potential of soft soils as observed in the 1985 Mexico City earthquake (Dobry et al 2000; Dobry & Susumu 2000; Ghosh 2001). Almost all of the current sub-Saharan African seismic codes are still using these spectra, which have meanwhile been replaced by new ones in the US and Europe, an exception being the South African code.

**Site-dependent design spectra: recent studies**

During the 7.1 magnitude Loma-Prieta earthquake of 1989, most of the damages linked to site-soil amplification and liquefaction took place in the Bay Area of San Francisco and Oakland located about 100 km NW of the epicentre. A great deal of recorded evidence was obtained from this area (Borcherdt 1994; Dobry et al 2000; Dobry & Susumu 2000).

One of the most important outcomes of post-Loma-Prieta empirical site-effect studies is the more pronounced amplification of response spectra by soft soil sites than observed ever before. Average spectral accelerations of numerous ground motion records from thick soil sites near the San Francisco Bay Area and Oakland are reproduced in Figure 12 for a damping of 5%, together with the average spectra from adjoining rock sites after the works of Dobry et al (2000) and Dobry & Susumu (2000).

Figure 12 shows that the rock-surface spectra are at least doubled by the soil in the short-period range of up to 0.2 s. In the velocity-sensitive period range of 0.2 to 1.5 s, the spectra are amplified to a much larger degree. Similar trends were also observed at stiffer soil sites, though to a lesser degree (Dobry et al 2000; Dobry & Susumu 2000).

Comparison of the spectral curves in Figure 12 with those in Figure 11(a) shows that the short-period amplifications were not revealed in the pre-Loma-Prieta studies. Also, the amplifications in the velocity-sensitive region were not as large as those in Figure 12. This may be attributable to the limited data base available at the time of the earlier studies. For this obvious reason, the older single-factor approach is no longer found adequate to account for site-soil effects. This fact led to the introduction of improved site-dependent design spectra in US seismic codes since 1994, and in other design codes worldwide afterwards.

**Systematic evaluation of improved amplification factors**

A number of systematic empirical studies conducted after the Loma-Prieta earthquake suggested that the soil amplification is
proportional to the mean shear-wave velocity, \( v_S \), of the upper 30 m thickness raised to a certain negative exponent, which is dependent on the period band and the intensity of the rock acceleration (Borcherdt 1994; Rodriguez-Marek et al; Borcherdt & Fumal, 2000; Dobry et al 2000; Dobry & Susumu 2000; Borcherdt 2002; Crouse & McGuire 2002; Stewart et al 2003). It was thus found important that site soils are classified on the basis of this important parameter.

The empirical studies of Borcherdt (1994; 2002) in particular showed that, for the low-amplitude rock accelerations not exceeding 0.1 g recorded in California Bay Area during the 1989 Loma-Prieta earthquake, the amplification factors for the acceleration-sensitive and velocity-sensitive period ranges, denoted by \( F_a \) and \( F_v \), are approximately proportional to \( v_S^{-0.4} \) and \( v_S^{-0.6} \) respectively. The statistically established values of \( F_a \) and \( F_v \) for low-intensity ground motions were used to calibrate one-dimensional analytical site response analysis programmes, which in turn were used to extrapolate the values of \( F_a \) and \( F_v \) for the higher range of rock accelerations of up to 0.4 to 0.5 g through analytical parametric studies (Borcherdt 1994; Dobry & Susumu 2000; Dobry et al 2000).

By combining the results of the empirical and analytic studies, Borcherdt (1994) arrived at the following best-fit generic relations for the two amplification factors that are applicable to a wide range of shaking intensity:

\[
F_a = \left( \frac{1050}{v_S} \right)^{m_a} \quad F_v = \left( \frac{1050}{v_S} \right)^{m_v}
\]  

(5)

The factor \( F_a \) is proposed for the period range of about 0.1 to 0.5 s and \( F_v \) for the range of about 0.4 to 2 s. The values of the exponents \( m_a \) and \( m_v \) are also provided by Borcherdt (1994) as functions of the intensity of the rock acceleration. The plots of Equation (5) are given in Figure 13, showing that \( F_a \) is consistently larger than \( F_v \) for \( v_S \) up to around 1 000 m/s–v_S of the reference rock site. Both factors tend to unity, with \( v_S \) approaching 1 000 m/s, and decrease with increasing intensity of rock shaking, as this is associated with increased damping.

### Current system of soil classification

For a generally stratified formation of \( n \) layers, each having a thickness of \( h_i \) and a shear-wave velocity of \( v_{S,i} \) within the upper 30 m thickness, the representative \( v_S \) can be established using the following relationship (BSSC 1995, 1998, 2004; EN 1998-1 2004; SANS 2010; Dobry & Susumu 2000; Dobry et al 2000):

\[
v_S = \frac{30}{t_{50}} = \frac{30}{\sum_{i=1}^{n} h_i v_{S,i}}
\]

(6)

Each term in the summation represents the time taken for the shear wave to travel through the respective individual layer. The shear-wave velocity computed in this manner is based on the time, \( t_{50} \), taken by the shear wave to travel from a depth of 30 m to the ground surface, and is thus not computed as the arithmetic average of the shear-wave velocities of the individual layers.

This approach also allows for the use of more readily measurable quantities in the field such as the standard penetration test blow count, \( N \), for granular deposits, or the undrained shear strength, \( S_u \), for saturated cohesive soils, though they are less reliable due apparently to the inherent double correlations.

Based on a landmark consensus reached by geotechnical engineers and earth scientists in the USA in the early 1990s, five distinct soil and rock classes, A to E, are introduced in accordance with this approach, as provided in Table 1. Corresponding approximate soil classes as per the older method (pre-1994) are also provided in the first column for comparison. A sixth much softer site class, F, is also defined that requires site-specific studies. This important subject is explained in more details in the works of Dobry et al (2000) and in the commentary volumes of NEHRP (BSSC 1995, 1998, 2004).

### New site amplification factors

Using the representative \( v_S \) of each soil class given in Table 1, one can establish the site amplification factors from Figure 13 for the appropriate value of rock-motion intensity considered. The discrete values so obtained according to Borcherdt (1994) and adopted by NEHRP (BSSC 1995, 1998, 2004) are given in Table 2. The effective peak acceleration, \( A_v \), and the effective velocity-related acceleration, \( A_v \), in the table are rock-level seismic hazard parameters employed to characterise site seismicity in the USA for a 475-year return period (BSSC 1995). The two effective ground accelerations have in the meantime been replaced by response-spectral parameters at two selected periods – 0.2 s and 1 s – in the more recent issues.

The tabular values show that soil classes C to E amplify the rock motion significantly, especially when the rock-shaking intensity is small, which is associated with reduced soil damping. The amplification is much larger in the velocity-sensitive period range than in the acceleration-sensitive period range for the non-rock soil classes, i.e. \( F_a \) is larger than \( F_v \). The short-period amplification factor, \( F_v \), is insignificant for rock-motion intensity larger than about 0.25 g, but is very significant for smaller-shaking intensity. This fact was not revealed in the earlier studies of Seed et al (1976).

In summary, the new amplification factors exhibit the following salient features (Ghosh 2001, 2004; Worku 2001):
1. The original three (later four) site categories are replaced by six new categories A to F.
2. The older qualitative site classification method is replaced by a new unambiguous and more rational classification method using a unique value of $v_s$ of the upper 30 m geological formation. Alternatively, average SPT blow counts and/or undrained shear strength can be used.
3. Two seismicity-dependent site coefficients, $F_a$ and $F_v$, replace the single site coefficient, $S$, of the old system. Both factors increase with decreasing shaking level due to the associated decreased damping. This leads to larger seismic design forces for many classes of structures in low-seismic regions, like in Africa, founded on soft formations.

It is important to point out that results of more recent studies based on an enlarged database including records from more recent earthquakes, like Northridge 1994, have not indicated significant changes to the values of the above site amplification factors (Borcherdt & Fumal 2000; Borcherdt 2002; BSSC 2004). For this reason, no major changes have been made so far to these values since their first introduction in 1994.

**SITE-DEPENDENT DESIGN SPECTRA IN RECENT US AND EUROPEAN CODES**

**Site effect provisions in the recent NEHRP series**

The recent site amplification factors described in the foregoing sections were first introduced in conjunction with the basic design spectrum of NEHRP 1994 given by the following relationship (BSSC 2004);

$$C_{Se} = \frac{2C_a}{T^{2/3}} \leq 2.5C_a; C_v = F_v A_v; C_a = F_a A_a$$ (7)

The short-period amplification factor $F_a$ is applied on the constant part of the spectra, whereas the intermediate-period amplification factor $F_v$ is applied on the descending segment. In order to show the basic shape of the elastic design spectrum, a plot of Equation (7) against period normalised with respect to $C_a$ is given in Figure 14(a) for $C/C_a = 1$.

Equation (7) is plotted in Figure 14(b) for the five soil classes A to E for a seismic zone characterised by $A_a = A_v = 0.1$. As pointed out earlier, the amplification occurs over the entire period range. Similar sets of spectral curves can be prepared for other seismic zones. These spectral curves for various soil classes are entirely different from those proposed by ATC-3, as given in Figure 11(b), especially in the short-period region.

In more recent versions of NEHRP, the amplification factors remained basically the same, whereas some changes have been introduced related to the seismic hazard parameters, namely $A_a$ and $A_v$, which have meanwhile been replaced by spectral accelerations at short period and at 1-second period, $S_a$ and $S_1$ respectively. Subsequent changes made to the basic shape of the design spectrum shown in Figure 14(a) include the reintroduction of the linearly rising left part and the change made to the descending right side from $T^{-2/3}$ to $T^{-1}$.
Site effect provisions in the recent Eurocode

The first edition of the European seismic code (Eurocode 8 1994) employed only three site classes, A, B and C, similar to those in ATC-3, 1978, which were also adapted by almost all African codes as presented above. EC 8 1994 strangely specifies a smaller maximum value for the softest site class C than for the stiffer sites A and B, as shown in Figure 15(a), and an underestimated amplification potential in the entire velocity-sensitive range. In light of the foregoing discussion, such a representation of the dynamic behaviour of the softest class of formations casts some doubt on its reliability. Similar doubts have also been expressed by Rey et al. (2002), who attributed this pitfall to lack of sufficient ad hoc studies prior to the publication of the document at the time. These spectra are now obsolete and are presented here for comparison purposes only.

The recent edition of Eurocode 8 (2004), issued in 2004, introduced five soil classes A to E with an additional class requiring site-specific studies. The system has many features in common with the recent NEHRP editions and other American codes, but also exhibits important differences. According to the new system, all rock and rock-like geological formations with $v_s > 800 \text{ m/s}$ are categorised under the single group of Ground Type A. Unlike in the former edition of Eurocode 8 (1994), each soil class in Eurocode 8 (2004) is assigned an amplification factor larger than unity, applied uniformly over the entire period range, though the code does not employ two different amplification factors for the acceleration and velocity-sensitive regions.

The values of the amplification factors vary in the range of 1 to 1.4, and 1 to 1.6 for Type 1 and Type 2 spectra, which are specified for regions of earthquakes of surface-wave magnitudes larger than 5.5 and less than 5.5 respectively. Type 1 spectral curves are shown in Figure 15(b). To be noted is that the new amplification factors are significantly smaller than the new NEHRP factors, $F_a$ and $F_r$, which can assume values up to 2.5 and 3.5 respectively. Note that in Figure 15(b) a segment descending according to $T^{-2}$ is included for periods longer than 2 seconds.

The number of independent research works that led to the curves of Figure 15(b) do not seem to be as many as those that led to the NEHRP spectra. This can partly be due to the difference in the size of the database available to the researchers. Type 2 spectra are similar to Type 1 spectra, but with consistently increased amplification factors and reduced control periods.

Note that Type 1 spectra are almost identical to those of SANS (SABS 2010).

**COMPARISON OF SITE-DEPENDENT DESIGN SPECTRA**

In this section, a comparison of the design spectra of the sub-Saharan African codes with those in selected recent codes of the US and Europe is presented. NEHRP 2003 (BSSC 2004) is selected as representative of the NEHRP series, which serves as the main resource document to many US seismic codes. Eurocode 8 (2004) automatically qualifies as the current seismic code in Europe, which replaced the older version (Eurocode 8 1994).

**African versus NEHRP and Eurocode 8 design spectra: rock sites**

Figure 16 compares the basic normalised design spectra for rock or firm ground condition of all codes considered. The normalisation is important in order to exclude the influence of seismicity and the return-period of the design earthquake.

With the exception of the spectra of the Kenyan code and NEHRP’s Class A, the plots show that the spectra for rock sites in the other codes are almost identical, with minor differences. This is not unexpected, as most of them are based on the findings of the pioneering works of Newmark & Hall (1969, 1982). The Kenyan spectrum is different in shape from the rest and smaller by up to 54% in the important period region of up to 1 second, and its background is unknown. As a large class of buildings in the big cities like Nairobi falls in this period range, the deficiency of the code must be of serious concern. Since Class A of NEHRP represents a stronger class of rock ($v_s > 1500 \text{ m/s}$) than the others, the corresponding spectral curve exhibits the lowest spectral ordinates for the most part. This class of rock is included to account for the strong crustal rock formation prevalent in the central and eastern part of the US. Many codes worldwide, including...
the European code, do not include this class of rock.

**African versus NEHRP spectra: soil sites**

The design spectra of the NEHRP 2003 (BSSC 2004) for soil classes C to E are plotted in Figure 17 for a selected PGA of 0.05 g together with the design spectra of the African codes for soil sites. This particular value of PGA is selected for consistency, with the comparison made in the next section to study the influence of return period. To avoid further congestion, and because of their similarities, the South African and the Eurocode spectra are plotted separately in the next section.

The plots in Figure 17 show that the spectra of most sub-Saharan African codes can be smaller by up to 60% of those of NEHRP. The spectra of the Kenyan code are even much smaller, being only 25% of NEHRP spectra in an important range of periods typical of commonly built types of buildings.

In general, the spectra of the sub-Saharan African codes are deficient over a wide range of period up to at least 1 second. This range encompasses most common-purpose structures built in African cities, including individual residential houses, schools, apartments, office flats, public offices, hotels, hospitals, etc, with heights reaching up to around 15 storeys. In the case of Soil Class E, buildings of almost any height would be under-designed by the provisions of sub-Saharan African codes.

The spectral discrepancies in Figure 17 clearly show the inadequacy of the old approach, on which most African codes are based, to account for site effect. Comparisons for other seismic zones can be made analogously, but the trend remains similar.

**African versus Eurocode 8 spectra: soil sites**

Comparison of the design spectra in the sub-Saharan African codes with the spectra of the European code is more direct forward, as all of them use PGA to characterise seismicity. This fact enables the comparison of normalised spectra regardless of the seismicity of regions. Comparison with Type 1 spectra of Eurocode 8 (2004) are given in Figure 18.

The plots show that the spectra of most sub-Saharan African codes can be smaller by up to 30% of Type 1 spectra of Eurocode 8, whereas in the Kenyan case, this can reach 58%. These discrepancies are still substantial, but comparatively smaller than the discrepancies observed with respect to the NEHRP spectra, whose amplification factors are consistently larger than those of Eurocode 8.

It should be emphasised that, compared to the Eurocode 8 amplification factors, the amplification factors of NEHRP are supported by evidence from a significant number of independent research works on a larger database enriched by ground motion records from the 1989 Loma Prieta and the 1994 Northridge strong earthquakes (Borchert 1994; Rodriguez-Marek et al 1999, Borchert & Fumal 2000; Dobry et al 2000; Dobry & Susumu 2000; Borchert 2002; Crouse & McGuire 2002; Stewart et al 2003).

**The combined influence of return period and site soil**

Of the five seismic codes of sub-Saharan African nations considered in this study, the South African and Ugandan codes adopted a return period of 475 years for the design earthquake, though the latter does not state this explicitly. On the other hand, the Ethiopian code unambiguously states a return period of 100 years only. The return period employed by the Kenyan code is not clearly stated, but one can infer from its seismic map and the associated zone factors that its return period cannot be more than 100 years. Similarly, the Ghanaian code does not explicitly state its return period. However, when compared to the GSHAP map of Figure 10(a), the PGA reaching up to 0.35 g in this local code suggest that the corresponding return period should be much larger than 475 years.

In order to study the combined effect of return period and site effect, we shall focus on the Ethiopian, Kenyan and Ghanaian codes, which adopt a return period different from 475 years. For this purpose, a site is selected in each of these countries within the same zone of the GSHAP map of Figures 10(a) and 10(b). This zone is selected as the one shaded in yellow with PGAs of 0.08 to 0.16 g. Addis Ababa of Ethiopia, Nakuru of Kenya and Accra of Ghana belonging to this zone are selected for comparison purposes of the combined influence of site soil and return period.

According to the GSHAP map of Figures 11(a) and 11(b), a representative PGA of 0.1 g may be assumed for the design earthquake in all of these places. In contrast, according to the respective local codes, Addis Ababa belongs to Zone 2 with a PGA of 0.05 g, Nakuru belongs to Zone VIII–IX with a PGA of 0.05 g, and Accra belongs to Zone 3 with a PGA of 0.35 g.

The NEHRP spectra corresponding to a PGA of 0.1 g for soil sites are presented in Figure 19 in comparison with the spectra for soil sites according to the African local codes. The comparisons are made for rock and soil sites separately.

The plots in Figure 19(a) for rock sites show that the design spectra of the East African codes of Ethiopia and Kenya can result in less than 50% of the spectra of NEHRP adapted to the GSHAP seismic hazard map. Since the site effect is zero for rock sites, the differences in these spectral curves reflect the influence of the return period of the design earthquake only. On the other hand, the spectra of the Ghanaian code can be close to threefold of the NEHRP spectra and five to tenfold of the Ethiopian and Kenyan spectra respectively.

Similarly, the plots in Figure 19(b) for soil sites exhibit a more pronounced difference between the NEHRP spectra and those of the East African countries of Ethiopia and Kenya. The NEHRP spectra could be as large as fivefold of the East African spectra. This suggests that the design ground motion provisions of the East African codes are highly inadequate to ensure safety. The Ghanaian spectra are still larger than the...
NEHRP spectra. This is despite the small site amplification factors of a maximum of 1.5 employed by this code, compared to up to 3.5 proposed by NEHRP. The results are indicative of flaws likely to have occurred in the background work of seismic hazard assessment that led to the local seismic hazard map of Ghana.

According to the plots in Figure (19), nearly all ranges of buildings (T of up to 3 s) built on any soil formation would be severely under-designed by the inadequate provisions of the East African codes of Ethiopia and Kenya. On the extreme opposite end, the Ghanaian code demands unjustifiably too large design forces for structures of all types on all categories of soil sites.

CONCLUSION
The status of the design ground motion provisions of representative sub-Saharan African codes has been assessed in terms of two basic issues of site effect and return period. The provisions are compared among the five selected codes and against NEHRP and the current European code. With the exception of the Ethiopian and the South African codes, the return periods are not clearly stated in the codes. The Ethiopian code clearly states 100 years as its return period for the design earthquake, whereas the South African code explicitly states 475 years. Through inference, the Kenyan and Ugandan codes seem to employ 100 years and 475 years, respectively. The Ghanaian code seems to have a major flaw in its seismic hazard map, because its zone factors are too large to suit the seismicity of the region.

In terms of site effect provisions, all except for the recent South African code, use site-dependent design spectra that are based on the outdated average spectra proposed by Seed et al (1976). The South African code directly adopted the design spectra specified by the current European code by simply omitting the softest site group. It is interesting to note that the recently revised Ghanaian and Ugandan codes have not incorporated the state of the art in site amplification potential. For this reason, most of the sub-Saharan African site-dependent design spectra can result in design seismic forces that are 30% to 75% smaller than would be obtained in accordance with the provisions of American and European codes. This is without considering the effect of the return period.

A comparison of the design spectra for rock sites reveals that the use of a return period of the design earthquake less than 475 years results in further underestimation of the design seismic forces. For example, the use of 100-years return period, which is the case in the Ethiopian code, would cause a further reduction of around 50% of the design force.

In most of the sub-Saharan codes considered, combined deficiencies from both site effect and return period are prevalent. When both shortcomings are considered, design seismic forces could be smaller than 20% of what would be calculated based on provisions of recent codes in Europe and USA.

The South African code has properly incorporated the state of the art of both site effect and return period. However, it is not clear why the European site-dependent spectra are favoured over the NEHRP spectra, while the NEHRP spectra and the corresponding site classification and amplification factors are supported by a larger number of independent studies on a large database, compared to the European spectra.

Apparently both the 1990 and 2010 editions of the Ghanaian code equally suffer from flaws in the background seismic hazard assessment study that led to the unrealistic seismic zone factors equivalent to PGA values ranging from 0.15 g to 0.35 g. In addition, even the 2010 edition failed to incorporate the current state of knowledge of site effect.

Based on the observations made above from the perspective of design ground motion alone, it may be concluded that most of the sub-Saharan seismic codes need careful revision, and in some cases a thorough overhaul, to provide adequate safety to human life and to the badly needed infrastructure and building structures. It is hoped that the GEM initiative will come up soon with important tools to rectify most of the identified shortcomings observed in the local codes and to provide a common platform that can lead to a continental harmonisation of approaches.

REFERENCES


http://www.globalquakemodel.org/

http://gmo.gfz-Potsdam.de/pub/GSHAP_Map_Online/gshap_map_online_frame.html

http://www.iae.or.jp/

http://www.seismo.ethz.ch/static/gshap/Gshap98-stc.html


http://www.usgs.org/


INTRODUCTION
The behaviour of site soils is one of the three major factors that can significantly influence the intensity of ground shaking due to an earthquake at any given site, the other two factors being the earthquake source mechanism and the geology of the seismic-wave path. The influence of site geotechnical conditions on ground-shaking intensity is studied following one of two approaches. The first is an empirical approach based on comparison of ensembles of recorded ground motions at nearby rock and soil sites of known geotechnical characteristics whenever these are available. The results of such studies are presented in the form of smoothed, statistically averaged site-dependent design spectra. These spectra are factored forms of the basic design spectrum for the corresponding rock site. The amplification factors are in general dependent on the nature of the site and the seismicity of the region. In the absence or scarcity of recorded ground motions for a given seismic region, it is common practice to adapt design spectra from regions of similar geologic and tectonic setup.

The second approach is appropriate for site-specific studies which involve modeling of the site soil as any other dynamic system subjected to the ground motion at the rock level. The soil can be modeled as a continuous or discrete mass system. The end results could be ground motion time histories, peak ground motions or response spectra at the ground surface.

This effect of site soils to amplify the rock-level ground motion is generally detrimental to the integrity of structures built on them.

Another important influence of site soils on structures is related to soil-structure interaction (SSI), which is rendered unfairly less attention, especially in routine building design. When the ground motion, amplified by the site soil in the manner described above, strikes the foundation, two forms of SSI take place. The first is attributed to the difference in rigidity between the foundation unit and the soil, which causes, among others, reflection and refraction of the seismic waves back into the soil mass. As a consequence, the motion of the foundation and the free ground become different, with the foundation motion usually being smaller. This aspect of SSI is known as kinematic SSI. Ideally, the foundation motion should be used as input motion in the analysis of the structure. However, studies have shown that the difference between the two motions can be regarded as negligible. For this reason, the free-ground motion is used as the input ground motion in practice (Fenves & Serino 1992; Stewart et al 1999; Stewart et al 2003).
The second, and more important, form of SSI is manifested when the superstructure starts to vibrate as a result of inertial forces triggered by the excitation at the foundation level. The inertial forces distributed over the height of the structure cause a resultant base shear and an overturning moment at the foundation, which in turn cause deformation of the soil. This deformation initiates new waves propagating into the soil mass. These waves carry away part of the energy imparted on the structure by the incoming earthquake waves and act as a means of energy dissipation in addition to the material/hysteretic damping inherent in the system. This form of SSI is known as inertial SSI. Its effect in most structures is to increase total displacement due to the additional soil deformation, and to decrease the base shear demand due to the associated reduced structural inertia forces as a result of the additional energy dissipation into the soil (Fenves & Serino 1992; Worku 1996; Stewart et al 1999; Stewart et al 2003; Tileyioglou et al 2011).

In the sense of the reduced base shear, the consideration of SSI effect is beneficial for most building structures. Unfortunately this important effect is mostly ignored by engineers, with the notion that the design is on the safe side without the additional computational effort needed to account for SSI effects (Stewart et al 1999; Stewart et al 2003). This tendency lacks scientific rationality and is happening despite the fact that provisions for this phenomenon have been made available in some design codes since the 1980s. The original versions of these provisions have meanwhile been updated through calibrations with actual records from relatively recent strong earthquakes, including the 1989 Loma-Prieta and the 1994 Northridge earthquakes (Fenves & Serino 1992; Stewart et al 2003; BSSC 2004; BSSC 2010). Results of such calibration works and recent experimental verifications are encouraging the use of the recent versions of code-based SSI provisions (Stewart et al 1999; Stewart et al 2003; Tileyioglou 2011). However, it is also worth pointing out that, in certain seismic and soil environments, an increase in the fundamental natural period of a moderately flexible structure due to SSI may have detrimental effects on the imposed seismic demand (Mylonakis & Gazetas 2000; Ziotopoulou & Gazetas 2010). In both cases it is becoming more evident that neglecting seismic SSI is not sustainable.

The recently revised South African seismic code (SANS 10160-4) adapted the site-dependent design spectra of EC8 (2004) with some modifications. These spectra are in general more demanding than those of the previous versions (SABS 0160 1989; EC8 1994). Some South African engineers have expressed concern during the preparation period leading to the issue of the recent design code regarding the potential escalation of material and construction costs associated with such stringent requirements (Wium 2010).

With due account for this concern, this paper attempts to demonstrate that a good potential exists for some of the costs associated with site amplification to be partially offset by the beneficial effects of inertial SSI on many structures. This happens if engineers are allowed to exercise some degree of freedom to employ SSI provisions available in some international codes until these make their way to the South African seismic code in due course.

**INERTIAL SOIL-STRUCTURE INTERACTION AND IMPEDANCE FUNCTIONS**

In order to understand the influence of inertial SSI on the response of building structures subjected to seismic ground motions, it is helpful to briefly introduce the basic principles and concepts of dynamics of foundations supported by flexible media like soils. For this purpose, we consider the vibrations of the rigid circular foundation of radius $R_0$ resting on the surface of the soil idealised as a homogeneous elastic half space shown in Figure 1 (a) and excited by the vertical harmonic load. Let the half space have an elastic modulus of $E$ and a mass density of $\rho$. For purposes of mathematical expediency and better insight, let us further represent the half space by the rudimentary model of the truncated solid cone of cross-sectional area of $A_0$ at the ground level which is the same as the contact area of the foundation. The cone defines the angle $\alpha$ with the horizontal and the height $h_0$ up to its apex above the ground (Worku 1996; Wolf & Deeks 2004).

After formulating the equation of motion of the conical soil beam based on the equilibrium of the differential soil element shown in Figure 1(c), it can be shown, without resorting to the details, that the differential equation for the capping rigid circular foundation of Figure 1(b) becomes (Worku 1996):

$$m_f \ddot{w}_0(t) + \frac{EA_0}{h_0} \ddot{w}_0(t) + \frac{EA_0}{h_0} w_0(t) = P_s \sin \omega t$$

where $m_f$ is the mass of the foundation, $c_L = \sqrt{E/\rho}$ is the velocity of the longitudinal elastic wave travelling away from the foundation through the conical soil column, and $w_0(t)$ is the vertical displacement of the foundation.

This equation is similar to the conventional equation of motion of the replacement single-degree-of-freedom (SDOF) oscillator shown in Figure 1(d) and given by:

$$m_f \ddot{w}_0(t) + C \dot{w}_0(t) + K w_0(t) = P_s \sin \omega t$$

where $K$ and $C$ are the spring and dashpot coefficients of the mechanical model respectively. Comparison of the two equations results in the following expressions for the parameters of the SDOF model of Figure 1(d) in terms of the geometry of the foundation and the elastic properties of the soil:

$$K = \frac{EA_0}{h_0}, \quad C = \frac{EA_0}{c_L}$$

This result obtained on the basis of a rudimentary idealisation of the soil-foundation system as a truncated conical column capped by the rigid foundation (Figure 1(a)) demonstrates the following fundamental facts:

- The semi-infinite continuum can be replaced by a simple SDOF mechanical massless model supported by a spring and a dashpot of coefficients, $K$ and $C$, respectively, arranged in parallel, and these parameters can be expressed in terms of the foundation geometry, the elastic parameters of the continuum and a pertinent wave velocity.

![Figure 1](https://example.com/figure1.png)
Unlike in conventional dynamic models of structures, the damping term – the second term in Equation (1) – is not an assumed addition of viscous damping; it is a mathematical outcome showing that the damping is an intrinsic behaviour of the system. This term represents an additional equilibrium force due to energy dissipation through waves propagating away from the foundation as represented by the wave velocity in the coefficient. It is in addition to the material damping of the continuum not considered in this discussion.

The truncated-cone approach was first devised and the above important outcomes observed about eight decades ago (Reissner 1936; Ehlers 1942). Interestingly, this seemingly primitive approach is extensively used in the recent book authored by Wolf and Deeks (2004) in a systematic manner. The new simplified approach has the potential of enabling engineers to easily solve a range of practical problems in structural dynamics involving SSI without reverting to complex finite-element techniques to model the site soil.

In a more rigorous treatment of the soil-foundation system, the spring and dashpot coefficients of Equation (3) are dependent on the frequency of excitation among many other factors (Luco & Westman 1971; Veletsos & Wei 1971). These coefficients, commonly termed as impedance functions, are now available in the literature for a wide range of conditions after several decades of intensive research works. They have meanwhile been well compiled, and have already made their ways into design codes starting from around 1980 (Gazetas 1983, 1991; Worku 1996; BSSC 2004; BSSC 2010).

Reverting to the mechanical model of Figure 1(d), its equation of motion given by Equation (2) for zero mass takes, for any degree of freedom considered, the form:

$$ C \ddot{w}(t) + K w(t) = P \delta(t) $$

The subscript of the deformation is dropped for brevity reasons, and the harmonic load is represented in its complex form for purposes of generality. The trial solution to this differential equation should also be complex. After substituting a complex function for $w(t)$ and solving for the complex-valued impedance function, which by definition is the ratio of the load to the response, yields:

$$ \frac{P(t)}{w(t)} = \tilde{K} = K + iaC $$

On the other hand, the complex-valued impedance functions obtained from rigorous mathematical treatments of the semi-infinite continuum are often presented in the literature in the following form:

$$ \tilde{K} = K_0[\alpha(\omega) + ia_0\beta(\omega)] $$

where $K_0$ is the static spring stiffness, $a_0$ is a dimensionless frequency parameter given by $a_0 = \omega R/V_s$, $V_s$ is the shear wave velocity of the continuum, $\alpha(\omega)$ and $\beta(\omega)$ are frequency-dependent dynamic impedance coefficients (also known as dynamic modifiers). By equating Equations (5) and (6) one obtains the following important relationships for the real-valued, frequency-dependent parameters of the massless spring-dashpot model in Figure 1(d):

$$ K = K_0 \alpha(\omega); \quad C = K_0 \omega R/V_s \beta(\omega) $$

As indicated above, the impedance coefficients, $\alpha(\omega)$ and $\beta(\omega)$, are available for various foundation conditions, soil conditions and vibration modes.

A circular foundation on the surface of the homogenous viscoelastic half space is the most basic and most important case. Studies have shown that use of an equivalent circular foundation gives satisfactory results for foundations of other shapes, provided that the aspect ratio of the encompassing rectangle of the foundation plan does not exceed 4:1. For other cases, suggested modifications are available (Gazetas 1991). The subsequent discussion will thus focus on circular foundations. The same discussion can easily be expanded to other shapes and soil-foundation conditions.

The static spring coefficients in Equation (7) for a circular foundation are given by the expressions in Equation (8) for the horizontal translation and rocking degrees of freedom respectively that are important in seismic design (Gazetas 1991; Worku 1996):

$$ K_{sh} = \frac{8GR_\theta}{2 - v}; \quad K_{\theta} = \frac{8GR_\theta^3}{3(1 - v)} $$

Note that the radii in the two cases are different for non-circular foundations and are determined by equating the area $A$ and moment of inertia $I_\theta$ for rocking motion of the actual foundation to those of the equivalent circular foundation. Thus,

$$ R_h = \sqrt{A/\pi}; \quad R_\theta = \sqrt{4I_\theta/\pi} $$

The corresponding dynamic modifiers for a surface circular foundation were originally provided by Veletsos and his co-workers (Veletsos & Wei 1971; Veletsos & Verbic 1973) and Luco and Westmann (1971), independently of one another, as functions of the frequency parameter, $a_0$. For other cases, appropriate impedance functions are available and should be used in order to determine the dynamic spring and dashpot coefficients as per Equation (7). Important factors to be further accounted for when using impedance functions include foundation embedment depth, foundation depth, foundation flexibility, soil layering and increase in stiffness of soil with depth. Relevant literature should be consulted for this purpose (Gazetas 1991; Pais & Kausel 1988; Worku 1996; Stewart et al. 1999).

**FLEXIBLE-BASE MODEL PARAMETERS**

In the most general three-dimensional case, a single mass oscillator fixed at its base acquires six additional degrees of freedom (DOF) when the base is released. The additional DOFs consist of a translational DOF in each direction of the Cartesian coordinate axes and a rotational DOF around each of them.

For an excitation due to upward propagating seismic shear waves, inclusion of the horizontal and rocking DOFs alone is sufficient in planar analysis. This condition is depicted in Figure 2 for a structure represented by an SDOF model, in which the complex-valued springs are lumped at the base in each of the horizontal and rotational
DOFs. Accordingly, the system now has three degrees of freedom. This representation is equivalent to a real-valued spring and dashpot arranged in parallel for each DOF. The height $h$ refers to the height of the roof in the case of a single-storey building and to the centroid of the inertial forces associated with the fundamental mode in the case of a multi-storey building which is commonly taken as $0.7h$ assuming a linear fundamental mode of vibration (Stewart et al. 1999; BSSC 2004).

In time-history analysis (THA), the frequency dependence of the foundation parameters and the nature of the system damping renders flexible-base models more difficult to analyse than fixed-base models. Such systems are termed as *non-classically damped systems* and can be solved using specially tailored closed-form or iterative analysis methods (Worku 1996, 2005, 2012).

In contrast to THA, in response-spectrum and pseudo-static analyses, SSI is accounted for by dealing with an equivalent SDOF system as shown in Figure 2(b) with modified parameters to account for the foundation flexibility. This was proposed by Veletsos and Meek (1974), who drew a parallel between the two models and found that the maximum displacement of the mass in Figure 2(a) can be accurately predicted using the replacement SDOF system in Figure 2(b) with a modified natural period of $T\tilde{}$ and a modified damping ratio of $\tilde{\zeta}$. These modified parameters are called *flexible-base parameters* and have the convenience of enabling the engineer to use the conventional code-specified seismic design spectra as usual.

Veletsos and Meek (1974) found out that the flexible-base period may be determined from:

$$\frac{T\tilde{}}{T} = \sqrt{1 + \frac{k}{kh} + \frac{k\theta}{k\theta}}$$

The fixed-base period is given by the well-known relationship of $T = \frac{2\pi}{\sqrt{k/m}}$, where $k$ is the stiffness of the structure and $m$ is its mass. According to Equation (10), the flexible-base period $T\tilde{}$ is always larger than the fixed-base period and increases with decreasing stiffness of the foundation. Measured period lengthening of more than 50% are reported by researchers (Stewart et al. 2003). Note that the period ratio is dependent on frequency (or period) because of the frequency-dependent foundation stiffnesses. It is, however, sufficient to establish the stiffnesses for the fundamental frequency/period of the fixed-base system (Stewart et al. 2003; BSSC 2004).

The effective flexible-base damping $\tilde{\zeta}$ is contributed from both the structural viscous damping $\zeta$ and the foundation damping $\zeta_0$, consisting generally of radiation and material damping components. Veletsos & Nair (1975) established the following relationship for the system damping based on equivalence of maximum deformations of the two oscillators in Figure 2:

$$\tilde{\zeta} = \zeta_0 + \frac{\zeta_0}{(T/T\tilde{})^3}$$

The plots of Equation (11) against the period ratio are given in Figure 3 for the commonly assumed fixed-base structural damping (FBSD) of 5% and a number of foundation damping (FD) values ranging from 3% to 20%. Such ranges of foundation damping ratios have been reported in the past (Stewart et al. 2003). The plots show that the overall effective damping of the flexible-base system is larger than the fixed-base damping (FBSD = 5%) with the exception of the rare case of the foundation damping itself being very low (smaller than 5%), and the period ratio being large. For any given foundation damping, the system damping gradually decreases with increasing period ratio due to the decreasing
The contribution of the structural damping with increasing period ratio. It should, however, be noted that the effective damping may not generally be taken less than the structural damping of 5% (BSSC 2004, BSSC 2010).

INFLUENCE OF INERTIAL SSI ON DESIGN SPECTRA

The influence of the lengthened period and the modified damping on a smoothened response spectrum is shown schematically in Figure 4. The figure shows that, for a fixed-base period of up to around 0.3 seconds, SSI has the effect of increasing the spectral response of the structure. However, for the most common case of building structures having a fundamental natural period larger than about 0.3 seconds, SSI has the effect of reducing the spectral response and thereby reducing the design base shear force (compare ordinates of the two curves corresponding to the pairs of \( T \) and \( \tilde{T} \) on either sides of \( T \approx 0.3 \) s).

A more direct insight into the influence of SSI on code-specified design spectra can be obtained by considering the EC8 (2004) Type 1 design spectra specified for five different site soil classes shown in Figure 5 for a structural damping ratio of 5%. The various site soil classes are defined in the code (EC8 2004). The amplification potential of the site soils is evident from the spectral curves. These spectra are incorporated into the provisions of the recently revised South African seismic code, with the exception of the spectrum for Site Class E (SANS 10160-4).

Let us consider the two soft site soil classes of C and D characterised by an average shear-wave velocity of 180 to 360 m/s and less than 180 m/s, respectively, over the upper 30 m depth in accordance with EC 8 (2004). The corresponding design spectra for the two site classes are presented separately in Figures 6(a) and 6(b) together with the spectrum for Site Class A – rock site.

Based on the definition of Site Class C, the maximum attainable foundation damping ratio including both material/hysteretic and geometric damping is estimated at 10% for the purpose of this study, even though larger damping ratios are reported for such a class of soil in the literature (Stewart et al 2003). The corresponding period lengthening of the SDOF system due to SSI is also conservatively estimated at 10% so that \( \frac{T}{\tilde{T}} \) reaches up to 1.10. With the effective system damping calculated from Equation (11) or read from Figure 3 as 13.76%, the corresponding design spectral curve is determined as per the provisions of EC 8 (2004) by scaling down the site-dependent spectral curve using a scaling factor to account for the modified damping. The factor is given by:

\[
\eta_{(\text{EC8})} = \frac{10}{5 + \tilde{\zeta}} \geq 0.55 \quad (12)
\]

In this equation \( \tilde{\zeta} \) is the effective system damping in percentile that accounts for both structural and foundation damping. The plot is given in Figure 6(a) by the dashed curve, which indicates that a significant reduction in the design base shear of up to 30% could be achieved for structures with a fundamental period larger than 0.2 seconds. Many classes of buildings belong to this period range. Most actual cases are expected to plot on or above the dashed curve.

Similarly, a little larger maximum limit for the foundation damping of 15% is assumed for the much softer Site Class D with corresponding period lengthening of up to 15%. The effective damping calculated as 18.3% resulted in the dashed curve shown in Figure 6(b). A larger reduction in the design base shear than in Site Class C seems attainable in this case. It is, however, important to point out that current code provisions for SSI cap the maximum permissible base-shear reduction to 30% (BSSC 2004).

The modified spectral curves for the two site classes are compared in Figure 7 against

![Figure 5 EC 8 2004 design spectra for different site conditions for a damping ratio of 5% (after EC8 2004)](image)

![Figure 6 Comparison of design spectra of EC8 2004 with those modified for SSI for (a) Site Class C, and (b) Site Class D)](image)
the corresponding site-dependent design spectra specified by the older version of EC8 (1994). In Figure 7(a), the design spectra for Site classes A and B of EC8-1994 are compared with the EC8-2004 design spectrum for Site Class C modified for SSI. Similarly, the spectra for site classes A and C of EC8 (1994) are compared in Figure 7(b) against the EC8 (2004) design spectrum for Site Class D modified for SSI.

It is interesting to note from the plots that the design spectra, and thus the design base shear, as per EC8 (2004) modified for inertial SSI effects can even be significantly lower than the values specified by the older EC8-1994 spectra for the corresponding soil classes over a significant range of fundamental period. The reduction is particularly significant in long-period structures.

The factor in the National Earthquake Hazard Reduction Program (NEHRP) document – a resource document for most seismic codes in the USA – for scaling down the site-dependent spectral curves corresponding to Equation (12) is given by (BSSC 2004):

\[ n_{(\text{NEHRP})} = (5/\zeta)^{0.4} \]  

The plots of Equations (12) and (13) are compared in Figure 8, which shows that the reduction proposed by the NEHRP document (BSSC 2004) is slightly larger than that of EC 8 (2004).

Finally, to be emphasised is the fact that the foundation damping and the period lengthening are key factors that affect the amount of spectral reduction due to SSI. It is, however, important to note that the reductions demonstrated in the above plots are based on assumed ranges of foundation damping and period lengthening for the purpose of this study, even though these are based on reasonable engineering judgment and reported cases (Stewart et al 2003).

Hence, the actual gains must be established by the design engineer on a case by case basis, and no generalisation is warranted on the basis of the presented material alone.

Nevertheless, the plots in Figures 6 to 8 demonstrate that the magnitude of spectral amplification by soil sites could be substantially offset by inertial SSI effects. If properly employed, SSI provisions could have the potential of leading to a significant financial saving in many cases, as is evident from the plots.

However, to be remembered is also the other important effect of inertial SSI that increases the lateral displacement of the building. This effect must be taken into account when considering ductility issues, secondary effects like P–Δ and possibilities of pounding with contiguous structures – considerations that are important in the design of tall buildings regardless of whether reduction in base-shear is achievable or not.

EXAMPLES

In order to illustrate the use of code provisions of SSI in seismic design of buildings

<table>
<thead>
<tr>
<th>Soil type</th>
<th>( v_{so} ) (m/s)</th>
<th>( f ) (kN/m²)</th>
<th>( v )</th>
<th>( G_0 + \frac{\gamma v^2}{g} ) (kN/m²)</th>
<th>( \frac{v}{v_{so}} )</th>
<th>( \frac{G}{G_0} )</th>
<th>( v_s ) (m/s)</th>
<th>( G ) (kN/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>220</td>
<td>18</td>
<td>0.40</td>
<td>88 807</td>
<td>0.95</td>
<td>0.90</td>
<td>209</td>
<td>79 926</td>
</tr>
<tr>
<td>E</td>
<td>150</td>
<td>18</td>
<td>0.45</td>
<td>41 284</td>
<td>0.64</td>
<td>0.47</td>
<td>96</td>
<td>19 403</td>
</tr>
</tbody>
</table>
and of the potential benefits, the site-soil classification and SSI procedures proposed in NEHRP are employed (BSSC 2004, 2010). EC8 (2004) does not have provisions for SSI.

Four different idealised reinforced-concrete buildings of height ranging from 5 to 25 storeys founded on the site soils types of D and E, according to the NEHRP classification system, are considered. Thus, the influence of SSI on eight different cases of dynamic system is studied simultaneously. The buildings are assumed to be located at sites characterised by a design peak ground acceleration (PGA) of 0.1 g, where g is the gravitational acceleration. Most sites in South Africa, where seismic design is required, are assigned a PGA of 0.1 g. According to the current response-spectra based on NEHRP seismic hazard mapping (BSSC 2004, 2010), a site of such seismicity can be represented by a short-period normalised spectrum $S_s$ of about 0.25 and an intermediate-period (1 second) spectrum $S_s$ of 0.1. Note that the US codes are no longer using PGA for seismic hazard characterisation.

A small-strain shear-wave velocity of 220 m/sec and 150 m/sec, a Poisson’s ratio of 0.4 and 0.45 are assigned to the two site classes D and E, respectively, whereas an effective soil unit weight of 18 kN/m$^3$ is assumed for both. The characteristics of the site soils are summarised in Table 1, in which the small-strain shear modulus is also computed from the direct relationship with the small-strain shear-wave velocity.

The actual shear-wave velocity and shear modulus corresponding to the large strains sustained during strong earthquakes at any given site are smaller and depend on the actual strain level, which in turn depends on the intensity of the anticipated earthquake shaking as represented by the seismicity of the site. The pertinent NEHRP provisions specify the ratios of $v_s/v_sB$ and $G/G_sB$ as per the seismicity of sites. These recommended ratios are given in columns 6 and 7 of Table 1, and the reduced values of the two dynamic properties are provided in the last columns of Table 1. It can be noted that the reduction is larger in the softer soil E due to the expected larger strains.

The data pertaining to the building structures are given in Table 2. All four building types considered are supported by a 20 m by 30 m rectangular raft foundation, have an additional basement storey and have a uniform story height of 3 m including the basement floor. The radius of the equivalent circular foundation for the horizontal and rocking (around the longer side of the rectangular foundation) degrees of freedom are computed from Equation (9) as 13.82 m and 12.63 m respectively.

The fundamental period is estimated using the relationship provided in the code: $T_a = C_r h_n^x$ (14)

The constants $C_r$ and $x$, which depend on the structural system, are also provided in Table 2 as proposed by the provisions of the code. The height $h_n$ is the total height of the building measured from the foundation level.

The periods computed using Equation (14) are presented in Table 3. A uniformly distributed permanent gravity load of 10 kN/m$^2$ is assumed on each floor for the subsequent computation of the building mass. The structural stiffness is computed using the natural period and the effective mass $m$ obtained by reducing the total mass by a factor of 0.7 as recommended by the code using the relationship in Equation (15).

$$k = 4 \pi^2 m/T^2$$ (15)

The coefficients $a_0$ and $a_\theta$ are generally frequency-dependent dynamic modifiers applied on the respective static stiffness given by Equation (8) for the horizontal and rocking motion, respectively. Whereas the modifier $a_0$ may be taken as unity for all practical purposes, the modifier $a_\theta$ must be established depending on the ratio $R_{m\theta}/T_a$ (BSSC 2004, 2010). Both the modifiers and the stiffnesses are computed and provided in Table 4.

Once the stiffnesses are established, the system (effective) period is calculated from Equation (10). The computed values are given in Table 5. The foundation damping $\beta_0$ is dependent on the aspect ratio $h/R$ of the building, the period ratio $T/T_a$ and the seismicity of the site, where $h$ is the effective height taken equal to $0.7h_n$. It is determined in accordance with graphs provided in the code document. Then the effective system damping is determined as per Equation (11), in which a structural damping of 5% is assumed for concrete structures as usual. The foundation damping and the system damping are given in the last columns of Table 5. Note that for computed values of the effective damping that are less than 5%, a minimum damping of 5% is taken according to the recommendations of NEHRP.

The seismic response coefficient corresponding to the fixed-base period $T$ is given

### Table 2 Building data

<table>
<thead>
<tr>
<th>Building type</th>
<th>Structural system</th>
<th>Storeys</th>
<th>$h_n$ (m)</th>
<th>Total mass (ton)</th>
<th>$C_r$</th>
<th>$x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>framed</td>
<td>6</td>
<td>18</td>
<td>3 600</td>
<td>0.0466</td>
<td>0.90</td>
</tr>
<tr>
<td>2</td>
<td>dual</td>
<td>11</td>
<td>33</td>
<td>6 600</td>
<td>0.0488</td>
<td>0.75</td>
</tr>
<tr>
<td>3</td>
<td>dual</td>
<td>16</td>
<td>48</td>
<td>9 600</td>
<td>0.0488</td>
<td>0.75</td>
</tr>
<tr>
<td>4</td>
<td>dual</td>
<td>26</td>
<td>78</td>
<td>15 600</td>
<td>0.0488</td>
<td>0.75</td>
</tr>
</tbody>
</table>

### Table 3 Computed system data

<table>
<thead>
<tr>
<th>Building type</th>
<th>Fixed base period $T_a$ (sec)</th>
<th>Natural frequency $\omega$ (sec$^{-1}$)</th>
<th>Structural stiffness $k$ (kN/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.63</td>
<td>9.973</td>
<td>250 657</td>
</tr>
<tr>
<td>2</td>
<td>0.67</td>
<td>9.378</td>
<td>406 305</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>7.060</td>
<td>334 926</td>
</tr>
<tr>
<td>4</td>
<td>1.28</td>
<td>4.909</td>
<td>263 125</td>
</tr>
</tbody>
</table>

### Table 4 Computed system data

<table>
<thead>
<tr>
<th>Building type</th>
<th>$a_h$</th>
<th>$k_m/T^2$</th>
<th>$a_\theta$</th>
<th>$k_h/(\pi m^2 \times 10^6)$</th>
<th>$k_\theta/(\pi m^2 \times 10^6)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>D E D E</td>
<td>0.096</td>
<td>0.209</td>
<td>0.93</td>
<td>0.81</td>
<td>5.523</td>
</tr>
<tr>
<td>D E D E</td>
<td>0.096</td>
<td>0.196</td>
<td>0.94</td>
<td>0.82</td>
<td>5.523</td>
</tr>
<tr>
<td>D E D E</td>
<td>0.068</td>
<td>0.148</td>
<td>0.97</td>
<td>0.85</td>
<td>5.523</td>
</tr>
<tr>
<td>D E D E</td>
<td>0.047</td>
<td>0.103</td>
<td>1.00</td>
<td>0.93</td>
<td>5.523</td>
</tr>
</tbody>
</table>

The pertinent NEHRP provisions specify the ratios of $v_s/v_sB$ and $G/G_sB$ as per the seismicity of sites. These recommended ratios are given in columns 6 and 7 of Table 1. It can be noted that the reduction is larger in the softer soil E due to the expected larger strains.

The data pertaining to the building structures are given in Table 2. All four building types considered are supported by a 20 m by 30 m rectangular raft foundation, have an additional basement storey and have a uniform story height of 3 m including the basement floor. The radius of the equivalent circular foundation for the horizontal and rocking (around the longer side of the rectangular foundation) degrees of freedom are computed from Equation (9) as 13.82 m and 12.63 m respectively.
by (for an elastic response and a normal-occupancy building):

\[ C_s = \frac{2}{3} F_s S_v \times \frac{2}{T_f} \]  

(17)

The coefficients \( F_s \) and \( F_v \) are the site amplification factors for the short-period and the intermediate-period regions of the design spectrum, respectively. They are determined from tables provided in the design code. \( F_s \) assumes the values of 1.6 and 2.5 for soils D and E, whereas \( F_v \) takes the values 2 and 3.2, respectively. Similarly, the seismic response coefficient \( C_s \), corresponding to the flexible-base period \( T_f \) is determined from Equation (17).

Finally, the effective system damping and the modified seismic response coefficient are employed together with the structural damping and the fixed-base seismic coefficient to calculate the reduction in base shear due to SSI using Equation (18). The results are presented in the last column of Table 6 as percentages of the base shear of the fixed-base system.

\[ \Delta \frac{V}{V} = 0.7 \left[ 1 - \frac{C_s}{C_s} \left( \frac{\beta}{\beta_0} \right)^{0.4} \right] \times 100\% \]  

(18)

The results obtained demonstrate that a significant amount of reduction in design base shear can be achieved if SSI provisions of design codes are properly used. In these particular examples a reduction in base shear of 7% to 39% is achieved. However, it is important to point out that the series of NEHRP documents (BSSC 2004, 2010) limit the maximum base-shear reduction due to SSI to a maximum of 30% as shown in brackets in the last column of Table 6. Obviously, the resulting cost saving in general could be of significant proportion, especially in medium-height buildings. The percentage saving increases with decreasing stiffness of the soil. However, the increasing trend of reduction in base shear with increasing building height seen in Table 6 is not expected to continue with further increase in the number of storeys outside the range considered, as the influence of SSI generally decreases with increasing slenderness of the building in taller buildings.

CONCLUSIONS AND RECOMMENDATIONS

The material presented in this paper demonstrated the importance of inertial SSI, which has the beneficial effect of reducing design spectral values or base shear in most building structures, but also increasing their lateral deformation. It was observed that effects of SSI increase with decreasing stiffness of the site soil. This effect of soils is in addition to their amplification potential and tends to compensate for part of the seismic base shear demand associated with response amplification.

According to the state of the art, the actual amplification potential of site soils is much more than stipulated in older design codes like EC8 (1994) and SABS (1989). It has been shown in this paper that the cost implications due to site amplifications, which in some cases could be prohibitive, may be significantly offset if SSI provisions are introduced in design codes. The necessary procedures are available in recent code provisions such as the NEHRP series (BSSC 2004, 2010).

Recent research has shown that code-specified relationships in design codes for computing the period lengthening, the effective damping and the reduction in base shear are meanwhile calibrated using recorded and measured data from the near past such that these provisions can give reliable results. In fact, it can be said that the state of current knowledge and confidence attained with regard to seismic SSI is comparable with that of the amplification potential of site soils. The examples considered in the paper demonstrated that substantial savings could indeed be achieved by accounting for seismic SSI effects. It is thus suggested that engineers are encouraged to use such provisions for a potentially economical structural design until these provisions make their way into the local code.

As a final note, it should be recalled that SSI has also the effect of increasing the total lateral deformation of buildings, which will in turn have an impact on the ductility requirements of the structure and on secondary effects like P-Δ. This aspect of SSI should also be duly accounted for in the design process, especially for tall structures.

REFERENCES


Proposed guideline for modelling water demand by suburb

M L Griffioen, JE van Zyl

This study investigated factors affecting the average domestic water demand of a large number of suburbs in South Africa. Suburbs form an ideal demand grouping since they tend to have similar stand areas, climatic conditions and user characteristics. In addition, since properties within a suburb are close to one another, it may be reasonably assumed that differences in user demands will cancel one another out so that a designer only has to cater for the average demand of the suburb. A database on measured domestic water demands was used to determine the average of the Annual Average Daily Demand (AADD) for a large number of suburbs in South Africa (i.e. the average AADD per suburb), and this data was linked to census and climate data. The combined data set was then subjected to various regression analyses to identify the most important influencing factors. Stand area was found to be the most important influencing factor, validating the approach followed by the current South African design guidelines. However, the current guidelines were found to exclude a large number of measured data points, and thus a new, more comprehensive design envelope is proposed.

INTRODUCTION

Domestic water demand is stochastic in nature, and is influenced by a large number of factors. These factors may be categorised as socio-economic (household size, income, stand area, social status, number of household appliances, social patterns, public and school holidays, tourism and water price), climatic (temperature, rainfall, humidity, time since last rainfall and the number of preceding hot days) and structural (number of users, water metering, plumbing fitting properties, pressure and network capacity) (Van Zyl et al 2008b).

The most basic descriptor of domestic demand is the Average Annual Daily Demand or AADD. When designing or analysing water distribution systems, parameters such as the design peak demand and reservoir capacity are based on the AADD. Thus it is important for engineers to have accurate estimates for the AADD of a system they are designing.

Figure 1 Current South African guideline for the annual average daily demand of domestic consumers; the designer selects a value between the upper and lower limit based on economic and climatic factors.
The AADD may be estimated from measured demands of similar areas, or from design guidelines. The South African design guideline for AADD (the “Red Book”) provides an upper from design guidelines. The AADD may be estimated from measured demands of similar areas, or working on. The AADD may be estimated from measured demands of similar areas, or from design guidelines.

The South African design guideline for AADD (the “Red Book”) provides an upper bound on stand area (CSIR 2003) as shown in Figure 1. The design engineer estimates the AADD for a particular area between these bounds by taking factors such as climate, income and local conditions into account. Municipal data management tools have become common in South African municipalities, and this has given researchers access to large amounts of municipal consumer water meter data. A number of previous studies have been done on domestic water consumption based on this data.

Garlipp (1979) investigated a number of parameters influencing water consumption in Pretoria, Bloemfontein, Cape Town, Port Elizabeth and Durban. Parameters influencing the domestic consumption the most were household size, prolonged high temperatures, stand area and income. In 1996 Stephenson and Turner investigated different income users in 14 areas in Gauteng, and concluded that the stand area exerted the most influence on domestic consumption. This study also found that income, population density and the level of service of water supply, as well as the housing type, were parameters influencing the domestic water consumption. Van Vuuren and Van Beek (1997) investigated both domestic and non-domestic water consumption of areas in Pretoria. The study found that high-income households consumed significantly more water than low-income consumers, and that their demands showed higher sensitivities to climatic variations due to large outdoor consumption. The study also found that the South African design guidelines (CSIR 2003) over-estimated domestic water consumption for Pretoria. A study by Van Zyl et al. (2003) found that the price of water, water pressure, household income and stand area influenced residential water consumption. The data consisted of residential areas in Gauteng only.

Van Zyl and Geusteyn (2007) performed an extensive study to develop a national water consumption archive for all South African municipalities that had implemented the SWIFT software. A study by Van Zyl et al. (2008a) used the archived data to analyse the demands of individual domestic consumers in South Africa. The study found that inland consumers tend to use more water than those in coastal regions. The study also concluded that stand area was the best parameter to use when predicting domestic water demand.

Jacobs et al. (2004) proposed a new guideline for domestic water demand estimation by analysing over 580 000 individual domestic users throughout South Africa. Stand area was used as influencing parameter, and the study found the current South African design guideline to be conservative. Husselmann and Van Zyl (2006) investigated the effect of stand area and stand value (as surrogate for income) on water demand for residential stands in Gauteng. The income (stand value) varied a lot, and stand area was found to be most influential. It was found that the South African design guideline underestimates the water demand for stand areas between 300 and 700 m², and that it overestimates the water demand for stand areas larger than 700 m². Finally, Jacobs et al. (2012) investigated residential water demand in nine service delivery areas in Ekurhuleni. A design guideline for water demand of 10.4 kl/d/ha was proposed after considering the average residential stand area of 70 suburbs in Ekurhuleni.

Studies of individual water demands are useful, but have limited application to the design of water distribution systems, since supply areas are not uniform, but consist of a mix of higher and lower consumption users. It is necessary to include the typical mix of consumers when estimating the AADD, which was done in this study by considering the AADD of suburbs rather than individual consumers.

Suburbs are seen as a suitable unit for analysing AADD, since they tend to group consumers of similar socio-economic conditions and are similar in many parameters, such as stand area, climate and pressure. Importantly, since consumers in a suburb are close together and share the same pipe network, variations in demand between individual consumers will tend to cancel one another out.

The aim of this study was to analyse the water demand of residential consumption in a large number of suburbs throughout South Africa, in order to determine the most important factors influencing demand and to propose a new design guideline based on these results. The next section describes the data used, linking of consumption data to climatic and socio-economic parameters, and the measures taken to ensure the integrity of the data. This is followed by a regression analysis of the data and the ranking of the parameters that most affect water demand. Finally, the average water demands by suburb are plotted against the average stand area and used as the basis for a new proposed design guideline for AADD.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Source of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand area, stand value with improvements and monthly meter readings</td>
<td>Municipal treasury data</td>
</tr>
<tr>
<td>Level of unemployment, annual household income, level of service for sanitation, level of service for water, household size, floor area and dwelling type</td>
<td>South African Municipal Demarcation Board</td>
</tr>
<tr>
<td>Rainfall data and daily temperatures (average maximum and minimum)</td>
<td>South African Weather Service</td>
</tr>
<tr>
<td>Mean annual precipitation and mean annual evaporation</td>
<td>Surface Water Resources of South Africa (Midgley et al. 1994)</td>
</tr>
</tbody>
</table>

DATA

Municipal water bills are based on consumer water meter readings taken at regular intervals. The past decade has seen significant software developments that now give researchers access to water consumption data held in treasury databases (Jacobs et al. 2004). One such tool is SWIFT a software package that has been implemented by a number of municipalities throughout South Africa, including most of the metropolitan municipalities (GLS 2007; Jacobs et al. 2012).

The data set used as the basis for this study is the same as that used in two previous studies (Van Zyl et al. 2008a; Van Zyl et al. under review) on domestic and non-domestic water demand respectively. It includes consumption data of forty-eight municipal treasury databases totalling more than 2.5 million records in four metropolitan areas (Johannesburg, Tshwane, Ekurhuleni and Cape Town) and 151 other towns. The data is described in detail in Van Zyl et al. (2008a). Only suburban residential stands were considered, and thus non-domestic and high-density developments, such as blocks of flats, were excluded.

Information on possible factors influencing domestic demand was collected from various sources and linked to the demand data to allow correlations to be analysed. The climatic data, i.e. mean annual precipitation (MAP), mean annual temperature (MAE) and evaporation (MAD), were linked to the water demand data on the basis of municipality, and, where higher resolution was possible, on suburb or group-of-suburbs level. Table 1 shows the factors collected in this study and their sources.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Source of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stand area, stand value with improvements and monthly meter readings</td>
<td>Municipal treasury data</td>
</tr>
<tr>
<td>Level of unemployment, annual household income, level of service for sanitation, level of service for water, household size, floor area and dwelling type</td>
<td>South African Municipal Demarcation Board</td>
</tr>
<tr>
<td>Rainfall data and daily temperatures (average maximum and minimum)</td>
<td>South African Weather Service</td>
</tr>
<tr>
<td>Mean annual precipitation and mean annual evaporation</td>
<td>Surface Water Resources of South Africa (Midgley et al. 1994)</td>
</tr>
</tbody>
</table>
The AADD for each consumer in the database was first determined based on twelve consecutive months’ demand values. This means that seasonal demand variations do not influence the values, and errors due to differences in reading dates or erroneous reading would generally be small. The meter readings exclude water supplied from other sources, such as boreholes or grey water, as well as leakage from the distribution system.

To ensure the integrity of the data, three data cleaning phases were implemented. In the first phase, SWIFT adjustment codes (assigned where SWIFT identifies certain anomalies or errors in the data) were used to exclude records with inconsistent readings or dates due to meter replacements or clock-overs, and records with less than 12 consecutive months of data. In addition, records or stands identified as unmetered, vacant, prepaid, duplicate (including more than one type of land use) or non-domestic were excluded. After this phase the database consisted of 1 353 740 records grouped into 1 753 suburbs.

In the second data cleaning phase, stands with unrealistically high or low stand areas/values were excluded, as described in Table 2. Note that the stand valuations were based on historic values and do not reflect the actual market value of the properties, which is likely to be significantly higher.

After the second phase, the database consisted of 1 218 053 records grouped into 1 497 suburbs.

In the third data cleaning phase, suburbs that could not be linked to census data, or that contained less than 100 consumers, were removed, leaving 744 852 records in 739 suburbs. This data formed the basis for the statistical analysis of water demand by suburb.

Finally, in order to allow a regression analysis to be conducted, suburbs with incomplete data records were removed, leaving 467 026 residential users in 459 suburbs. For each of these suburbs, the following data was available:

- Location (municipality and suburb)
- Average AADD
- Average stand area
- Average stand value
- Type of water supply (piped water, borehole, spring, etc)
- Type of sanitation (sewer line, septic tank, chemical toilet, etc)
- Type of building
- Type of area (urban, smallholding, farm, etc)
- Average income
- Racial distribution (fractions of consumers that were black, white, Indian, etc)
- Average number of rooms per house
- Mean annual precipitation

DEMAND CHARACTERISTICS

The demand data resulting from the third cleaning phase was analysed to determine the characteristics of domestic water demand in South Africa when grouped by suburb. Table 3 summarises the main statistical descriptors for both the AADD per property, and AADD per unit area. The latter is important, since outdoor consumption is strongly linked to stand area and forms an important component of domestic consumption.

The cumulative distribution of the AADD by suburb is shown in Figure 2. The average AADD is 1 100 ℓ/property/day, which is reasonably close to the median value of 1 000 ℓ/property per day. The AADD data shows a large variation with a coefficient of variation of 55% and ranging between 200 and 3 800 ℓ/property/day. The suburbs with the highest AADD were found to be characterised by high-income consumers on large stands, typically larger than 2 000 m². Consumption per unit area has an average value of 1.33 L/m²/day and a median value of 1.19. While it has a smaller coefficient of variation (44%) than AADD per property, the AADD per unit area displays a large range

![Figure 2 Cumulative distribution of AADD by suburb](image-url)
of values, varying between a minimum and maximum of 0.12 and 4.42 L/m²/day respectively. Areas with high AADD per unit area were found to be characterised by low-income consumers on stands typically smaller than 500 m². The cumulative distribution of the AADD per unit area is shown in Figure 3.

FACTORS AFFECTING DEMAND

To determine the relative importance of different parameters on domestic water demand grouped by suburb, a regression analysis was conducted on the 459 suburbs with complete data records. The regression coefficients found are given in Table 4 ranked by importance of the parameter (or absolute value of the standardised regression coefficients). The results confirm that stand area is the most important explanatory variable for AADD, and thus supports the continued use of stand area as the basis for estimating the AADD.

The regression equation was applied to the individual suburbs in the data set, and the predicted AADDs are compared to the measured values in Figure 4. While the regression equation clearly shows the right trend, it is far from a perfect correlation, with some calculated values containing substantial errors. This observation, combined with the nearly impossible task of accurately estimating all the model parameters listed in Table 4, means that the regression equation is unlikely to find practical application.

PROPOSED DESIGN GUIDELINE

Suburbs are residential areas that form part of a city or town, typically consisting of a few hundred properties. A given suburb would typically consist of stands of similar size, service levels and value relative to the city as a whole. In addition, the residents of a suburb tend to be similar in their socio-economic status. A basic premise of this paper is that the water consumption behaviour of users in a given suburb is as close to uniform as it is practically possible. In addition, since consumers in a suburb are physically situated close to one another and share the same pipe network, differences between higher- and lower-demand consumers will be cancelled out locally. As a result, it is only necessary to design for the average and not the highest demand in a suburb.

The average AADDs of 739 suburbs throughout South Africa are plotted against their respective average stand areas in Figure 5, also showing the current South African design envelope. The figure shows that many of the measured data points fall outside the current guideline – 4.5% fall above the upper line and 34% below the

### Table 4 Results of the regression analysis on AADD by suburb

<table>
<thead>
<tr>
<th>No</th>
<th>Parameter</th>
<th>Regression coefficient</th>
<th>Standardised regression coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Area (m²)</td>
<td>0.0004737</td>
<td>0.44432</td>
</tr>
<tr>
<td>2</td>
<td>Stand value (R)</td>
<td>1.116E-06</td>
<td>0.35825</td>
</tr>
<tr>
<td>3</td>
<td>Min temp (°C)</td>
<td>0.0641537</td>
<td>0.21135</td>
</tr>
<tr>
<td>4</td>
<td>Fraction of Black African consumers (% points)</td>
<td>0.0049177</td>
<td>0.19903</td>
</tr>
<tr>
<td>5</td>
<td>Mean Annual Evaporation (mm)</td>
<td>0.0003934</td>
<td>0.17746</td>
</tr>
<tr>
<td>6</td>
<td>Fraction of properties with formal houses or brick structures (% points)</td>
<td>-0.0035367</td>
<td>-0.1213</td>
</tr>
<tr>
<td>7</td>
<td>Average number of rooms per house</td>
<td>0.0560035</td>
<td>0.10107</td>
</tr>
<tr>
<td>8</td>
<td>Fraction of stands with piped water inside the dwelling (% points)</td>
<td>0.0021721</td>
<td>0.07046</td>
</tr>
<tr>
<td>9</td>
<td>Fraction of properties with flush toilet connected to a centralised sewerage system (% points)</td>
<td>0.0041268</td>
<td>0.06017</td>
</tr>
<tr>
<td>10</td>
<td>Fraction of properties that have town, cluster or semi-detached houses (% points)</td>
<td>-0.0020448</td>
<td>-0.0488</td>
</tr>
<tr>
<td>11</td>
<td>Mean annual precipitation (mm)</td>
<td>0.0001203</td>
<td>0.02322</td>
</tr>
<tr>
<td>12</td>
<td>Intercept</td>
<td>-1.8018469</td>
<td>n/a</td>
</tr>
</tbody>
</table>
lower line of the guideline envelope. Thus it is clear that the current guideline should be updated to reflect the true distribution of AADD found in the field.

New proposed envelope curves for the design of domestic AADD were developed based on the data. Since the data shows a very large range, envelope curves that include all points would be too wide to make practical sense. An optimisation approach was therefore followed to find envelope lines that would best trade off the following two competing objectives:

1. Minimising the number of data points falling outside the envelope curves.
2. Minimising the sum of the differences between the envelope line and the data points. An envelope line will achieve this objective best by following the shape of the data as closely as possible.

An objective function was developed by first weighting and then adding the results of these two measures. The envelope lines were adjusted to include more or less of the extreme points by modifying the relative weights of the two objectives.

Different line types were tried for the envelope, and linear lines were eventually found to describe the most inclusive and consistent envelope. In addition, a trend line (the blue line) was fitted to all the data included in the envelope. The trend and proposed upper (purple) and lower envelope lines are shown with the data in Figure 6.

Visual inspection of the upper line indicated that a lower upper line (the grey line, also called the alternative upper line) can be applied without changing the operation of the new design guideline by much. This upper boundary includes 96.6% of the points between the upper (purple) and lower (red) lines and narrows down the design guideline to a more user-friendly design tool.

Designers should use local factors such as climate, income and population density to determine an appropriate design AADD between the lower and upper envelope curves. It should also be noted that the measured AADD values per suburb tend to be closer to the mean value trend line for larger stands.

The equations for the trend and upper and lower envelope lines are given by the following equations:

\[
\begin{align*}
AADD_{\text{mean}} &= 0.000089A + 0.3 \\
AADD_{\text{upper}} &= 0.0011A + 1.1 \\
AADD_{\text{alternative upper}} &= 0.001025A + 0.85 \\
AADD_{\text{lower}} &= 0.00054A
\end{align*}
\]

where AADD is the annual average daily demand per property in $\text{kℓ/property/day}$ and $A$ is the stand area in $\text{m}^2$.

Finally, it is necessary to analyse the points that fall outside the proposed envelope curve to identify the types of properties that are not included in the proposed design envelope. Suburbs for which the AADD is underestimated are most critical, since this may lead to design failure. An analysis of these suburbs identified two categories:

- Very high-income consumers, mostly in gated developments, such as Kyralami Estate, Benmore Gardens and Dainfern in Johannesburg.
- Suburbs surrounding Hillbrow in Johannesburg, which is one of the highest population density areas in the country. These suburbs originally developed as higher-income residential areas, and thus consist of relatively large houses. It is suspected that these areas have a much higher population density than similarly sized suburbs, due to their proximity to Hillbrow, and that houses are shared by several families. Suburbs included in this category are Troyeville, Yeoville, Bellevue East and Berea.

A third category is introduced by the inclusion of the 96.6% boundary line (grey line). The areas that fall between the original upper boundary and the 96.6% boundary include:

- Two areas in Breede Valley, i.e. Langerug, Van Riebeeck/Panorama.
- Four areas in northern Johannesburg, i.e. Parkview, Maroelalad, Mayfair, Newclare.
- Six areas in eastern Johannesburg, i.e. Fourways, Parktown North, Parkwood, Bruma, Sydenham, Malvern.
- One area in southern Johannesburg, i.e. La Rochelle.

CONCLUSION

This study analysed residential water demand based on the premise that a suburb provides the best basis for analysing AADD. Not only are stands in a suburb of similar size and users of similar socio-economic status, but variations in demands cancel one another out due to the close proximity of the users to one another.

An analysis of the average AADD of 739 suburbs spread throughout South Africa showed that the average AADD is 1 100 $\text{ℓ/property/day}$, but a large range of AADDs was observed.

A regression analysis of the AADDs of 340 suburbs resulted in an equation to

![Figure 5 The average demand and stand area of 739 suburbs throughout South Africa and the current South African design guidelines](image-url)
predict the AADD of a suburb based on various stand, climatic and socio-economic parameters. However, this equation produces significant errors for some suburbs, and is too complex to have practical application. Significantly, the regression analysis showed that the stand area is the most important descriptive parameter of AADD.

Finally, the average AADDs of 739 suburbs were plotted against their average stand area, showing that the current guidelines do not provide a suitable description of the range of AADDs for design purposes. A new envelope curve is proposed that includes the vast majority of suburbs. Suburbs with consumption above the upper envelope curve include very high-income gated communities and specific high-density suburbs bordering Hillbrow in Johannesburg. Suburbs with AADDs below the lower envelope curve consist mainly of holiday destinations and certain low-income rural settlements.

REFERENCES


Accounting for moment-rotation behaviour of connections in portal frames

Portal frames are steel structures used to construct industrial buildings. Conventional analysis techniques used by practising engineering professionals assume that the eave, ridge and base connections are either infinitely rigid or perfectly pinned. This approach leads to less accurate analysis of the displacement behaviour of portal frames when subjected to external loading. Portal frames must therefore be analysed with rotational springs at all connections to yield accurate displacement behaviour. This investigation focused on determining the accuracy and economic feasibility of modelling portal frame connections with rotational springs. The rotational spring stiffnesses of all connections required before the portal frame could be analysed in a second-order two-dimensional non-linear analysis. The rotational spring stiffnesses unique to each connection were determined from the moment-rotation behaviour obtained from a series of finite element analysis simulations of each connection. Thereafter these stiffnesses were used to determine the vertical and horizontal displacements of the portal frame. These displacements were compared with experimental test results. The reasons for the discrepancies between the numerical and experimental results were investigated through a sensitivity analysis. The findings suggest that it is not computationally feasible to analyse portal frames with rotational springs, even though the model’s predicted results are more accurate than those of conventional analysis using rigid and pin connections.

INTRODUCTION
Portal frames are steel structures composed of columns and rafters with various types of connection between the structural elements. Figure 1 shows a two-dimensional view of a portal frame with a ridge connection without a haunch and eave connections with a haunch. Column bases can be designed either as pinned (hinged) or moment-fixed (infinitely rigid) connections. Most column bases in portal frames are designed as pinned connections. This approach leads to a more economical design than portal frames with rigid column bases. Pinned bases are less expensive to manufacture and foundations are smaller since no moment resistance is required. Eave and ridge connections are usually designed as moment-fixed connections, i.e. a transfer of bending moment takes place between the connecting members. However, tension bolts within the rafter height are inadequate for developing sufficient moment capacity at the connections (Narayanan & Kalyanraman 2003). Therefore, haunched elements using tapered I-sections are introduced on the bottom flange of the rafters at the connection to increase the moment capacity of the element, instead of increasing the rafter size (Moore & I-Acronym 1984).

Figure 1 Two-dimensional view of a portal frame
The current practice of design engineering professionals in South Africa is to model the portal frame as a two-dimensional structure during the analysis procedure. It is the norm for column base connections to be modelled as perfectly pinned, while the eave and ridge connections are modelled as infinitely rigid. These assumptions are flawed, resulting in incorrect displacement behaviour of the portal frame which leads to incorrect steel sections being used (Kruger et al 1995).

In addition, higher grades of steel have been rolled in South Africa over the past few decades. The yield strength of structural steel has increased from 300 MPa for 300WA to 355 MPa for 355J (SAISC Handbook 1995). This results in members having greater axial and bending resistances due to the increase in yield strength for the same member size. These stronger elements are now used in design, and result in lighter sections being used. The elastic design of steel structures for normal use is mainly governed by serviceability limit state requirements, which are controlled by the displacements of the structure (Narayanan & Kalyanraman 2003). Members are therefore initially sized in analyses according to the serviceability limit state criteria. Final design of the structure is then conducted according to the ultimate limit state requirements, which are controlled by the strength capacity of the elements. The displacement of the structure is controlled by the flexural stiffness of the member and not by the yield strength of the material. Thus, the use of higher grade steel has no effect on the allowable, informative, codified guidelines for such structures as recommended by SANS 10162-1:2005 (SANS 2005). The increase in the yield strength of structural steel does not result in a codified reduction in the displacement behaviour of the portal frame.

It has thus become necessary to determine the displacement behaviour of portal frames accurately by using proper modelling techniques and by taking the increased yield strength of structural steel into account. This paper focuses on determining the real displacement behaviour of a portal frame by modelling the real behaviour of connections in portal frames with appropriate material properties.

**LITERATURE BACKGROUND**

**Moment-rotation connection behaviour**

Infinitely rigid connections transfer axial forces, shear forces and bending moments between structural members. Under loading, the connections will undergo deformations as forces and bending moments are transferred between the members. Deformations caused by axial and shear forces are negligible in comparison with the rotational deformations that will occur and will therefore not be considered in this study (Kruger et al 1995). In-plane rotation of the connection is the most prominent type of deformation in portal frames, and is caused by the bending moment acting at the connection. It can thus be stated that the rotation of a connection is a function of the moment applied to it. The rotation (θ) of a connection is defined as the change in angle of the structural components connected to it, that is, the change in angle between the centre lines of the column and the beam due to the loading of the portal frame resulting in a moment being generated at the connection. This is illustrated in Figure 2 (Shi et al 2007).

**The effect of moment-rotation behaviour of connections on portal frames**

Modelling of the real behaviour of the connections will reveal the distribution of forces, the bending moments and the displacements of the structure (Gerstle 1988).

**Joint behaviour**

In practice a rigid connection under loading exhibits rotation as a moment develops. Before the moment is transferred between the members, some rotational deformation occurs as a result of the elasticity and real behaviour of the connection, which delays the transfer of
springs (Simitses et al. 1984; Chan et al. 2005).

Past research found that different types of structural analysis rotational springs in pinned supports results in greater displacements occurring in the ultimate moment (Lui & Chen 1987). This leads to greater displacements in the structure before the ultimate capacity of the connection is achieved.

**Column base behaviour**

Modelling column base connections using pinned supports results in greater displacements than the actual (as-built) displacements. This is due to the rotational restraint provided by the column base plate/concrete interaction of the actual connection. The rotational restraint caused by the concrete results in a moment developing at the base of the column, which is contradictory to what is assumed in the design of the base connection (Jaspart et al. 2008).

Bending moments will develop at the column base if the connection is assumed to be rigid, thus resulting in greater displacements being recorded in practice compared with the theoretical analysis due to the flexibility exhibited by the actual base.

**Rotational springs in structural analysis**

Past research found that different types of connection can be modelled as rotational springs (Simitses et al. 1984; Chan et al. 2005). As stated previously, deformations caused by shear and axial forces are negligibly small compared with rotational deformations. A rotational spring permitting in-plane rotational deformations can be incorporated between various members of the portal frame to simulate the joint stiffness. The reader is referred to Chan and Chui (2000) for the mathematical formulation of rotational springs.

Rotational springs can be assigned to individual nodes in most structural analysis software. The spring stiffness is usually provided in terms of the relationship of the bending moment to the rotation of the connection. The stiffness of each connection is obtained by taking the derivative of the moment-rotation behaviour, i.e. the slope of the initial curve. The slope of the initial linear elastic region of a moment-rotation curve is referred to as the “initial stiffness” of the connection. Serviceability limit state design guidelines refer to “limiting the elastic deflections of the structure” (SANS 2011). The structure is therefore analysed with all components remaining within their elastic response regions. In this investigation the connection stiffnesses are modelled with their “initial” stiffness, which is sometimes referred to as their “elastic” stiffness.

**RESEARCH METHODOLOGY**

The investigation was conducted on a 5 m portal frame. It was divided into the following tasks:

1. An experimental investigation was conducted to determine the actual displacement behaviour of the portal frame with hinged supports (idealised conditions) and grouted supports (construction conditions) for three different loading conditions.
2. A two-dimensional numerical second-order analysis of the portal frame was conducted for the same conditions as the experimental configuration. In the analysis various methods of modelling the connections of the portal frame were considered. These included pinned, fixed and rotational spring connections. The rotational spring stiffnesses were determined from the moment-rotation curves which were obtained from a finite element analysis.
3. The experimental displacement results were compared with the numerical analysis displacement results for various types of connections to assess their accuracy.

**EXPERIMENTAL SET-UP**

The purpose of the experimental investigation was to determine the actual displacement behaviour of the portal frame when subjected to different loading and support conditions. The results were used as the benchmark to determine the accuracy of the displacement behaviour of the portal frame obtained from the numerical analysis.

**Experimental configuration**

Figure 4 shows a two-dimensional view of the experimental 5 m span portal frame with a column height of 1.5 m. All the sections are manufactured from IPE AA 100 sections. The portal frame was subjected to the load cases LC1, LC2 and LC3 shown in Figure 4, which were individually applied to the portal frame. The individual load cases shown in Figure 4 are now described.

- LC 1: a vertical downward load of 4.96 kN applied at the ridge which simulates the permanent load of the rafters, sheeting and purlins.
- LC 2: a vertical upward load of 12.8 kN applied at the ridge which simulates the wind load.
- LC 3: a horizontal load of 7.85 kN applied at the apex of the column perpendicularly to the portal frame which simulates the wind load.

Column bases are usually numerically analysed using pin supports as shown in Figure 5a, whereas the actual (grouted) support
used in practice is shown in Figure 5b. From Figure 5b it is clear that the actual support does provide some degree of rotational restraint, which is due to the base plate/ concrete interaction. Both cases were experimentally investigated to determine their effect on the displacement behaviour of the portal frame.

**NUMERICAL MODEL**

A numerical model of the full-scale experimental test configuration was developed in Strand 7, a commercially available structural analysis software. The 1.5 m columns and the 2.54 m rafters were meshed with 0.1 m quadratic shear flexible (Timoshenko) beam elements. A series of different connections types were modelled, namely:

- **Rigid connection**: These connections transfer moments between the members and are modelled as fixed. This type of connection was initially used to model the interaction at the ridge and eaves.
- **Pin connection**: These connections only transfer shear and axial forces between the members and are modelled as hinged connections. This type of connection was initially used to model the column bases.
- **Rotational spring connection**: These connections are neither fixed nor pinned and thus they allow the transfer of shear and axial forces, as well as a percentage of the moment, depending on the degree of fixity of the joint. This type of connection was subsequently used to accurately model the semi-fixed connections at the eaves, ridge and column bases.

A pinned support prevents vertical and horizontal displacements of the node, thus inducing vertical and horizontal forces. The same applies to the fixed support, except that the rotation of the node is restricted, thus inducing an additional bending moment. With a rotational spring, both vertical and horizontal displacements are restricted, while providing some resistance to rotation. Therefore the stiffness of the rotational spring must be determined before rotational springs can be implemented in the numerical analysis.

**Rotational spring stiffness**

Rotational spring stiffness is unique to each connection that is affected by the following attributes, among others: the size of the steel profiles, the size of the haunch, the number and size of the bolts, the position of the bolts, the torque of the bolts and the size of the connection end-plate. The rotational spring stiffness in the numerical simulation can also be affected by the element type, element mesh density, contact formulation, material properties and the type of analysis. The rotational spring stiFFnesses were obtained by modelling a part of each connection with all members connected to the joint, using ABAQUS version 6.10-2, a general finite element (FE) analysis software. The FE simulations were conducted based on the guidelines presented by Prabha et al. (2007).

These simulations resulted in a moment-rotation relationship for each connection. The rotational spring stiffness was obtained from the linear region of the moment-rotation curve within the elastic range. A previous experimental study conducted by Truter (1997) determined the moment-rotation behaviour of a haunched eave connection. The experimental rotational spring stiffness was used to determine the accuracy of the FE model. Figure 6 shows the experimental results obtained by Truter (1997) with a fourth-order regression line superimposed through the experimental results, as well as the FE simulation results.

The regression line and the FE simulation response yield results that are virtually identical for the initial linear region. Only the linear region is important for this study, as serviceability limit state requirements refer to the elastic response of the structure. Based on the exceptionally good fit, it was accepted that the techniques used in the numerical model yield accurate results for this investigation and they were thus used to obtain all the other rotational spring stiffnesses.

The reader is referred to Albertyn (2011) for a detailed description of how the moment-rotation curves were obtained for each connection. Table 1 presents the rotational spring stiffnesses for each connection.

The differences between the clockwise and anti-clockwise rotations of the eave, as well as the ridge connections, fall within acceptable limits.

**RESULTS AND DISCUSSION**

The vertical and horizontal displacement responses of the 5 m span portal frame with hinged and grouted supports are presented for load cases 2 and 3. The responses of load case 1 are omitted as they yield displacement patterns similar to those of load case 2. Table 2 lists the types of connection investigated for both types of support, with a description of each displacement response used in Figures 7a through 7d and 8.
Results for load cases 2 and 3

Figure 7a presents the vertical displacement of the ridge and Figure 7b presents the horizontal displacement at the column apex when subjected to LC 2. Figure 7c presents the vertical displacement of the ridge and Figure 7d presents the horizontal displacement at the column apex when subjected to LC 3.

For ease of reference, the models that refer to the pin-supported column bases are shown as solid lines. The corresponding models referring to the grouted supports are presented as dashed lines. Table 3 presents a summary of the significant results that were extracted from Figures 7a to 7d.

In Table 3 we notice that the maximum vertical displacement difference between the models is 3.4 mm, with a maximum percentage difference of 6.2% for LC 2. Larger percentage differences occur between the various models of the horizontal displacement for LC 2 and the vertical displacement of LC 3. However, the maximum displacement differences for these load cases are 1.9 mm and 2.3 mm, respectively. Due to the insignificant displacement differences in these models, it can be assumed that the numerical models yield sufficient accuracy. The differences obtained between the numerical and experimental models can be attributed to the accuracy with which the experimental measurements were obtained.

Large displacement and percentage differences occur in the horizontal displacement for LC 3. This is clearly evident between the displacement responses of the experimental hinged and experimental grouted models. There is a displacement difference of 8.7 mm or 30.1% between these experimental models. This implies that the grouted support interface has a significant effect on the horizontal displacement of the portal frame when subjected to a lateral force. A better correlation was expected between the displacement responses of the rotational spring and experimental grouted models since the numerical model incorporates rotational

<table>
<thead>
<tr>
<th>Model name</th>
<th>Hinged base support models</th>
<th>Model name</th>
<th>Grouted base support models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental hinged model</td>
<td>Experimental model: ■ Eaves and ridge connections designed as rigid ■ Base connections designed as hinged</td>
<td>Experimental grouted model</td>
<td>Experimental model: ■ Eaves and ridge connections designed as rigid ■ Base connections designed as grouted</td>
</tr>
<tr>
<td>Conventional numerical model</td>
<td>Numerical model # 1: ■ Eaves and ridge connections modelled as analytically rigid ■ Base connections modelled as hinged supports</td>
<td>Conventional numerical model</td>
<td>Numerical model # 1: ■ Eaves and ridge connections modelled as analytically rigid ■ Base connections modelled as hinged supports</td>
</tr>
<tr>
<td>Updated conventional numerical model</td>
<td>Numerical model # 2: ■ Eaves and ridge connections modelled using rotational springs ■ Base connections modelled as hinged supports</td>
<td>Rotational spring model</td>
<td>Numerical model # 2: ■ Eaves and ridge connections modelled using rotational springs ■ Base connections modelled using rotational springs</td>
</tr>
</tbody>
</table>

Figure 7a Vertical displacement response of the ridge when subjected to LC 2

Figure 7b Horizontal displacement at the column apex when subjected to LC 2
springs at all connections. This led to the conclusion that other influences contribute to the difference in displacement responses which were previously thought to be insignificant, and this led to a sensitivity analysis being conducted on the grouted interface.

**SENSITIVITY ANALYSIS**

The aim of the sensitivity analysis performed on the column base was to identify the possible cause(s) of the inaccurate horizontal displacement behaviour produced by the numerical model with rotational springs. Since only the column base connection was changed, the sensitivity analysis focused on the effect of this connection on the displacement behaviour of the portal frame. This led to an investigation of the effect of the preload on the holding-down bolts.

**Effect of preload on holding-down bolts**

Ordinary bolts in bolted steel connections are tightened according to the “turn of the nut” method (Kulak et al. 2001). Various experimental studies indicate that the method referred to results in a bolt preload of approximately 70% of the bolt proof stress. A further investigation was conducted to determine the effect of this factor on the displacement response of the portal frame, since the bolt preload affects the rotation of the column base connection. This led to determination of the moment-rotation curves at bolt preloads of 0%, 5%, 35% and 70%. Table 4 shows the rotational spring stiffnesses of the grouted column base at these bolt preloads, which were obtained from the moment-rotation curves from the FE analysis.

From Table 4 it can be concluded that bolt preload has a significant influence on the rotational spring stiffnesses. Rotational spring stiffness differences of 14% and 16% are obtained between bolt preloads of 0% and 35%, and 35% and 70%, respectively. This shows that the bolt preload could have an effect on the displacement behaviour of the portal frame.
approximately a quarter of the expected value. This implies that the column base of the experimental grouted model is less stiff than expected. It also suggests that some slip may have occurred between the holding-down bolts and the foundation, or that the holding-down bolts were not torqued to the required 70% bolt preload. After completion of the experimental tests, careful examination of the holding-down bolts revealed that an insignificant slip had occurred at the column base. For this reason the effect that in-plane rotation has on the horizontal displacement of the portal frame was determined. Table 5 presents the magnitude of the horizontal displacement of the portal frame as a function of the column base plate rotation.

From Table 5 it can be observed that an insignificant base plate rotation of 0.26° or 0.00454 radians will result in a horizontal column apex displacement of 6.8 mm. Therefore, if the column base plate rotated insignificantly by 0.26°, this would result in an additional 6.8 mm horizontal displacement of the portal frame compared with when no slip occurs. If no slip of the holding-down bolts occurred, an experimental displacement of 21.8 mm would be observed if we assume that the column base plate rotated by 0.26°. A difference of 1.7 mm or 7.2% is found when this displacement is compared with the 70% bolt preload of the numerical model. Also, a difference of 0.6 mm or 1.9% is observed when the actual experimental displacement is compared with the average displacements of the 0% and 70% bolt preloads. This proves that the numerical model with rotational springs yields accurate results.

The conventional numerical approach overestimates the displacement response, which lies beyond the 0% bolt preload displacement response. The conventional approach overestimates the displacement response by 11.5% and 65.1% compared with the displacement responses of the 0% and 70% bolt preloads.

**CONCLUSION**

This study confirms that the numerical model with rotational springs can be used to model a portal frame, and that it does yield more accurate displacement results than the conventional analysis. The important question that must now be asked is whether the numerical model with rotational springs (updated numerical model) is an economically viable option to use in a consulting engineering practice in South Africa. This question is best answered in terms of the expertise of the designer, available software, time required for the analysis and the potential cost saving.

In terms of time, there is an insignificant computational time difference between the conventional model and the updated numerical model when a two-dimensional second-order non-linear analysis of the portal frame is performed. Accurate rotational spring stiffnesses of each connection are required before the two-dimensional analysis of the updated numerical model can be performed. This was achieved by conducting a finite element analysis of each connection, which required 12 hours of computational time per connection using a four-quad core computer with 32 GB RAM.

Also, considerable expertise in finite element analysis is required to develop a model of all the connections, taking into account the level of complexity to develop accurate numerical models. This could be achieved by an experienced graduate professional with the necessary theoretical and practical knowledge of finite element analysis. These individuals, however, attract a higher cost to company and greater consultancy fees to conduct an analysis.

### Table 5 Portal frame displacement due to base plate rotation

<table>
<thead>
<tr>
<th>Vertical lift of one end of the base plate (mm)</th>
<th>Base plate width (mm)</th>
<th>Base plate rotation (°)</th>
<th>Base plate rotation (radians)</th>
<th>Horizontal displacement of the portal frame (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125</td>
<td>110</td>
<td>0.13</td>
<td>0.00227</td>
<td>3.4</td>
</tr>
<tr>
<td>0.25</td>
<td>110</td>
<td>0.26</td>
<td>0.00454</td>
<td>6.8</td>
</tr>
<tr>
<td>0.50</td>
<td>110</td>
<td>0.52</td>
<td>0.00908</td>
<td>13.6</td>
</tr>
<tr>
<td>0.75</td>
<td>110</td>
<td>0.78</td>
<td>0.01316</td>
<td>20.4</td>
</tr>
<tr>
<td>1.00</td>
<td>110</td>
<td>1.04</td>
<td>0.01815</td>
<td>27.3</td>
</tr>
</tbody>
</table>

### Figure 8 Horizontal column displacement

Figure 8 shows the numerical model's horizontal displacement responses for bolt preloads of 0% and 70%, with the experimental grouted model's displacement response superimposed. Figure 8 shows that the experimental grouted model produced a horizontal displacement of 28.6 mm when a horizontal load of 7.85 kN was applied to the apex of the column. The corresponding displacements at 0% and 70% bolt preload were 34.8 mm and 23.5 mm respectively. To obtain a similar horizontal displacement of 28.6 mm, the column base rotational spring stiffness in the rotational spring model was adjusted to 52 kNm/rad.

From a close examination of Figure 8 it can be seen that the experimental response lies midway between the 0% and 70% bolt preloads. This would suggest that the rotational spring stiffness of the experimental model is approximately the average of the rotational stiffnesses of the two numerical models' rotational stiffnesses, i.e. ± 250 kNm/rad. The actual column base rotational spring stiffness in the numerical model was obtained as 52 kNm/rad, which is
The software required to perform advanced finite element analysis are expensive and usually not readily available in local consulting engineering design offices. Conventional and affordable structural engineering design software used in most design offices cannot perform the required advanced simulations to obtain moment rotation curves of connections. This leaves the design engineer unable to perform sophisticated analysis.

Based on the aforementioned reasons, and since most of the portal frame structures require a limited number of portal frames, the cost saving achieved using the updated numerical model would not make this type of analysis economically viable in a design office.

Thus, based on the findings of this research, it is recommended that portal frames in practice be analysed using the conventional approach, as it is reliable and safe. Since more accurate displacement results were obtained modelling connections as rotational springs, the recommendation is that this approach be followed for structural engineering research applications.

The scope of this study did not include focusing on the buckling behaviour of portal frames. It is therefore recommended that further research on portal frames be conducted at ultimate limit state behaviour accounting for the real behaviour of connections. Further research could also be conducted in the dynamic behaviour of portal frames with connections modelled as rotational springs.

REFERENCES


The application and interpretation of linear finite element analysis results in the design and detailing of hogging moment regions in reinforced concrete flat plates

S A Skorpen, N W Dekker

Finite Element methods have been used by civil and structural engineers since the 1960s, and the theory behind this is well researched. However, there is still a lack of direction on how to use the information obtained from this type of analysis to practically design a structure for strength and serviceability criteria. Design codes are broadly based on simplified calibrated strength models and are consistent with simplified and practical detailing.

In this paper traditional methods of analysis of a simple pad foundation are compared with the linear finite element method, and the results compared to experimental results. The following questions are answered:

- Are the traditional simplified methods adequate with respect to overall strength?
- To what extent may finite element peaks or singularities be averaged or smoothed without compromising durability and serviceability?
- How should the reinforcement obtained from linear finite element methods be detailed?

LIST OF NOTATIONS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>effective depth of reinforcement in a slab or footing (mm)</td>
</tr>
<tr>
<td>h</td>
<td>depth of concrete in a slab or footing (mm)</td>
</tr>
<tr>
<td>B</td>
<td>shorter plan dimension of a footing (mm)</td>
</tr>
<tr>
<td>D</td>
<td>longer plan dimension of a footing (mm)</td>
</tr>
<tr>
<td>Ms</td>
<td>smoothed support bending moment (Nmm)</td>
</tr>
<tr>
<td>Mp</td>
<td>peak moment at the centreline of the support (Nmm)</td>
</tr>
<tr>
<td>Fs</td>
<td>support reaction (N)</td>
</tr>
<tr>
<td>bs</td>
<td>width of support column (mm)</td>
</tr>
</tbody>
</table>

INTRODUCTION

Concrete design codes, SANS 10100 and Eurocode 2, currently in use, contain calibrated strength models enabling the user to calculate a safe resistance of a structural member. In many cases, such models are simplifications of quite complex failure modes. Load effects obtained using appropriate methods of analysis provide values of bending moments, shear forces and axial forces. Local peak effects (singularities) cannot be calculated using the traditional methods of analysis. The simplifications are justified, by and large, by the ductile behaviour of the members. The reliability of the models has been proved by the lengthy process of calibration involved and the many structures that have safely resisted the applied loads. A specific application can be found in a plate with column supports. Significant differentiation in curvature over the supports is regulated by the traditional methods, by simple stepping requirements.

The advent of finite element methods of analysis provides absolute rather than average values of load effects and stresses. Practical detailing of structural elements does not generally take cognisance of the peak values obtained from more sophisticated methods of analysis. Given how long simple methods of analysis have been used, and the reliability attached to the proven methods, the use of more sophisticated methods of analysis should be applied in such a manner as to provide consistent results.

Finite Element (FE) methods have been used by civil and structural engineers since the 1960s (Carlton 1993) and the theory behind these is well researched. However, there is still a lack of direction on how to use the information obtained from this type of analysis to...
practically design a structure for strength and serviceability criteria. Design codes assume that the designer has engineering judgement and a 'feel' for the behaviour of concrete when using FE analysis (Brooker 2006).

FE plate structures are analysed using classic plate theory which has been formulated by considering equilibrium and strain compatibility in plates which are thin enough for shear deformations to not have a significant effect on the behaviour of the slab, and thick enough that in-plane and membrane forces are not important. Park and Gamble (2000) refer to these plates as "medium thick" but they are generally referred to as thin plates.

Thin plate theory is used for flat structures where transverse shear effects are not important, and is based on Kirchhoff's theory. Thick plate theory is used for flat structures where the effects of transverse shear must be included. This is based on Reissener's / Mindlin's theory which takes into account the effect of shear strain.

The basic assumptions of thin plate or shell theory are summarised by Rombach (2005):
- Plane sections remain plane before and after loading
- Linear strain distribution of the slab depth (Navier theory)
- No strain at the middle of the plane
- Stresses in the normal direction can be ignored
- A thin slab (span/depth > 10)
- Constant slab depth
- Small vertical displacements, \( w \ll h \) (Order theory)

The simplified analysis of flat plate type structures, such as slabs and footings, described in most codes (TMH 7, SANS 0100 and BS8110) ignore transverse shear effects and assume that plane sections remain plane.

Where shear effects become important (i.e. deep beams where span/effective depth < 2) the member can be modelled using equivalent truss analogy.

FE analyses can be either linear or non-linear. Linear FE analysis is the most commonly used type, but is limited in its capabilities as it does not take cracking and softening of the concrete into account (Rombach 2004). This type of analysis is suitable for an ultimate limit state design check, but cannot be used to check serviceability deflection and cracking. Non-linear FE analyses model the cracked behaviour of the concrete by means of an iterative process, but are complicated and time consuming to set up, and the software cost is significantly more than a linear FE program. In practice, flat plate type structures are generally designed using a linear FE analysis, and serviceability compliance done with 'rule of thumb' span to effective depth ratio checks.

The main criticisms of linear FE analyses are its use of elastic material properties, which result in overestimated support moments and underestimated deflections (Jones & Morrison 2005), and an impractical required reinforcement contour output. Figure 1 shows the typical transverse bending moment distribution in a pad footing.

A paper by Brooker (2006) gives recommendations for interpretation of a linear FE analysis of flat slabs and advocates averaging the peak moment across a larger area. The recommendation is to use the total bending moment under the FE moment curve, and apportion it as per the detailing rules given in BS 8110 (1997). This requires three quarters of the moment to be resisted by the column strip, of which two thirds are apportioned to the inner column strip. The remaining moments are resisted by the outer column strip and the edge strip.

In this paper traditional methods of analysis of a simple pad foundation are compared with the linear finite element method and the results compared to experimental results. The following questions are answered:
- Are the simplified methods adequate with respect to overall strength?
- To what extent (width) may peak values be averaged or smoothed without compromising durability and serviceability?
- How should the reinforcement obtained from linear FE methods be detailed?

The intention is not to do a theoretical assessment of the finite element method, but rather to provide a practical explanation of how it can be applied to general structural engineering, giving guidelines on the required amount of reinforcement and placement thereof to satisfy ultimate and serviceability limit states.

### Table 1: Apportionment between column and middle strip in footings expressed as a percentage of the total negative (hogging) moment

<table>
<thead>
<tr>
<th>Code</th>
<th>Column strip width</th>
<th>Column strip</th>
<th>Edge strip</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMH 7 1989, Code of Practice for the Design of Highway Bridges and Culverts in South Africa, Part 3 (discussed under footings)</td>
<td>( b_{col} + 3d ) if the width of the footing is greater than 1.5(b_{col} + 3d)</td>
<td>66.67%</td>
<td>33.33%</td>
</tr>
<tr>
<td>SANS 10100 2000, The structural use of concrete, Part 1</td>
<td>( B ) ( \frac{2}{\frac{D}{B} + 1} \times 100 )</td>
<td>66.67%</td>
<td>33.33%</td>
</tr>
<tr>
<td>Eurocode 2, Design of concrete structures EN 1992-1-1:2003 (E)</td>
<td>( b_{col} + 3d ) if the width of the footing is greater than 1.5(b_{col} + 3d)</td>
<td>66.67%</td>
<td>33.33%</td>
</tr>
<tr>
<td>BS 8110:1997 Structural use of concrete, Part 1</td>
<td>( b_{col} + 3d ) if the width of the footing is greater than 1.5(b_{col} + 3d)</td>
<td>66.67%</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

\( b_{col} \) = column dimension in the long direction; \( d \) = depth of the slab; \( D \) = longer plan dimension of footing; \( B \) = shorter plan dimension of footing.

**Figure 1** Typical transverse distribution of bending moments in a pad footing.
DESIGN OF FOOTINGS FOR FLEXURE

The prescribed method for designing footings in most codes is consistent with the design requirements for flat slabs. In the methods described in the South African bridge code, TMH 7 (1989), and the South African concrete design code, SANS 10100 (2000), vertical loads are resisted by an equivalent beam with the same width and depth as the footing. Bending moments are not constant across the width of a footing, and it has been experimentally shown (Regan 1981) that they are highest on a line connecting the columns, and then reduce transversely. For this reason most codes prescribe the design of footings by considering a "column" and a "middle" or "edge" strip, with the column strip resisting approximately two thirds of the load effect and the middle strip one third. This apportionment varies between codes, and Table 1 summarises what various codes require. This approach is aimed at satisfying serviceability requirements by placing more reinforcement in regions of higher bending moment and thereby reducing curvature. In this paper the authors refer to this method as the simplified design (SD) method.

FINITE ELEMENT METHOD

The significant advance in computer software technology in recent years has resulted in a surge in the use of finite element software to analyse the load effects in structures, and in particular flat plate type structures. The finite element method is an approximation in which a continuum is replaced by a number of discreet elements (Zienkiewicz et al 1976). Each component representing the system as a whole is known as a finite element. Parameters and analytical functions describe the behaviour of each element and are then used to generate a set of algebraic equations describing the displacements at each node, which can then be solved. The elements have a finite size and therefore the solution to these equations is approximate; the smaller the element the closer the approximation is to the true solution (Brooker 2006). The output from a linear finite element flat slab analysis is in the form of contour plots of stresses and bending moments. These peak bending moments can vary considerably depending on how the support conditions are modelled, and the element size. It is the opinion of the authors that the basics of using a linear FEM to analyse flat slabs is commonly understood by most designers. However, the modelling of column to flat plate connections is still open to numerous forms of interpretation and designer preference. The most common support models listed by Rombach (2004) are shown in Figure 2.

The models in Figure 2 can be interpreted as follows:

a. Full 3D continuum model – this models accurately, but is very time consuming.

b. Pinned supports over all nodes above the column – this is not suitable where the column is relatively flexible.

c. Encased supports assigned to the edge of the column in the shell model – this is not suitable where the column is relatively flexible.

d. Spring supports assigned to the column area in the shell model.

e. Rigid column head – this allows rotation of the column cross section and is suitable for flexible columns.

f. Point support at one node – this is the least accurate way of modelling a support, but probably the most commonly used. Peak load effects (singularities) in elastic FE models are consistent with high elastic stresses. These peaks are reduced by yielding and cracking, or ‘softening’, of the concrete, and are never actually realised in real structures. In a two-dimensional analysis, the bending moment in a one-way spanning slab supported on pinned supports is generally smoothed using the following equation given by Rombach (2004):

\[
M_p = M_{sb} - \frac{F_s b_s}{8} \tag{1}
\]

where

- \(M_p\) = peak moment at the centreline of the support
- \(F_s\) = support reaction
- \(b_s\) = width of support column

Singualrities in FE analyses commonly occur where pinned supports and concentrated loads are modelled. The stress and bending moment contour output from a finite element modal will indicate peaks as shown in Figure 1.

Flat slab/plate behaviour is three-dimensional and much more difficult to analyse. It is widely accepted that these singularities in a flat slab analysis do not need to be considered in design. However, if this is assumed, then it is not clear to what extent a peak value obtained from a simplified FE model
may be smoothed or averaged in a two-way spanning slab.

The typical output from a linear FE analysis of plate elements gives bending moments in the x and y directions \( M_x \) and \( M_y \) and a local twisting moment \( M_{xy} \) (see Figure 3). This twisting moment takes the three-dimensional behaviour of a flat slab into account. However, it does not act in the direction of the reinforcement. A popular method of including the twisting moment is known as Wood-Armer moments, and most design software will automatically calculate the Wood-Armer moments for the user. The Wood-Armer moments were developed to take complex loading into account, where the twisting moment \( M_{xy} \) needs to be considered (Denton & Burgoyne 1996). There are four components – top (hogging) moments in the x and y directions \( M_{xT} \) and \( M_{yT} \), and bottom (sagging) moments in each direction \( M_{xB} \) and \( M_{yB} \). This method is conservative and these moments form an upper limit envelope of the worst-case design moments. The four components can be used to calculate the required reinforcement for each of the reinforcement layers in a flat slab type structure. (Brooker 2006)

Modern codes allow for nonlinear analysis of reinforced concrete structures, but in practice such a complex analysis is seldom justified due to the large amount of work required and the cost of suitable software. Designs are usually based on linear-elastic material behaviour, assuming that the ductile properties of reinforced concrete allow for a limited redistribution of forces. Rombach (2004) states that the accuracy of such a simplified approach is generally sufficient. A conservative design approach is to have two slab models, one where columns are assumed to be pinned supports to determine the worst case sagging moment, and the second where the column supports are fixed to determine the worst-case hogging moments. Eurocode 2 does not prescribe a specific analysis or dictate how to interpret FE method load effects, which are open to a wide range of interpretations depending on how the column supports are modelled. Most commonly used FE packages give no clear directive on how to detail the reinforcement for flat slabs designed using FE. In general it is accepted that the design engineer will use the required reinforcement contour plots to decide how to place the slab reinforcement. It is, however, obvious that if the FE reinforcement contours are followed exactly this would lead to a very impractical reinforcement layout.

From the above it is clear that when designing and detailing, using FE analysis, a great deal is left up to engineering judgement.

**FLEXURAL CAPACITY OF FOOTING CALCULATED BY THE SIMPLIFIED METHOD AND FINITE ELEMENT METHODS**

In order to assess load effects in a linear elastic FE model, the design of a simple foundation pad footing was undertaken using simplified design methods (SD), and then compared with the results of a linear elastic FE model of the same footing. The moment variation at the critical design section (i.e. at the face of the column) for both methods of analysis was compared.

The analysis of a reinforced concrete pad footing is a multi-parameter problem. The stiffness of the flat slab is significantly influenced by the non-linear properties of concrete (i.e. cracking, which in turn influences member forces and deflections), and furthermore there is the added complexity of soil-structure interaction. The deformation characteristics of the soil can play a significant role in the distribution of the pressure and hence the load effect. For this exercise a conservative uniform bearing pressure under the footing was assumed, with the column acting as a support.

The moments obtained from the FE analysis are also very sensitive to how the supports are modelled in the FEM, and to the slab geometry. The effect of the support model was considered by analysing a square footing and changing the way the supports were modelled, and the effects of geometry was considered by analysing a combined loading rectangular footing and varying the thickness.

The conventional flat slab/footing design method described in the South African bridge code TMH 7 (1989) was used for the simplified method of design. The critical section for the design of the flexural hogging reinforcement in the x and y directions is taken at the face of the column (Figure 4). If the width of the footing is greater than \( b_{col} + 3d \) the model effects FE moments, a square footing was considered by analysing a combined loading rectangular footing and varying the thickness.

To get an indication of how the support model affects FE moments, a square footing was analysed with the different support conditions described by Rombach (2004), and the following noted for each, as summarised in Table 2 and shown Figure 5:

- peak \( M_x \) axis moment (FE \( M_{peak}^{\text{x}} \))
- peak Wood and Armer moment (FE Wood and Armer \( M_{peak}^{\text{WA}} \))
- total \( M_x \) axis moment — the sum under the moment curve at the face of the column (FE \( M_{total}^{\text{x}} \))
- total Wood and Armer moment — the sum under the moment curve at the face of the column (FE Wood and Armer \( M_{total}^{\text{WA}} \))
These peak and total moments were then compared to the SD column strip (SD $M_{col}$) and total moments (SD $M_{total}$), as well as the concentrated SD column moment where two thirds of the column strip moment are concentrated into an inner column strip. The applied load corresponded to a typical pressure under a footing founded on dense sand (SANS 10160 1989).

The pad foundation was modelled using the linear elastic FE program Prokon (2012), which is available to the majority of designers in South Africa (see Figure 5). It consisted of square 0.025 m x 0.025 m plate elements. The elements used to analyse the pad foundations are discreet Kirchoff-Mindlin quadrilaterals which provide good results for both thick and thin plates and are free from shear locking. Shear deformations are not considered here, as the cantilever span to depth ratio was relatively low. The following is a summary of the results shown in Figure 6 and Table 3:

- For each different support model the total FE $M_{x}$ or $M_{y}$ moments were the same as the total SD moment.
- The total FE Wood and Armer moment was greater than the SD moment (up to 20% more), because Wood and Armer moments are design moments, which include the $M_{xy}$ twisting moment.
- The peak FE Wood and Armer moments were higher than the peak FE $M_{x}$ or $M_{y}$ moments because of the $M_{xy}$ twisting moment.
- The $M_{xy}$ twisting moments were significantly affected by how the supports were modelled.
- Different supports (constraint) conditions caused the FE $M_{x}$ or $M_{y}$ peak moment to vary by as much as 36%.
- Different supports (constraint) conditions caused the FE Wood and Armer peak moment to vary by as much as 88%.
- The FE peak $M_{x}$ or $M_{y}$ moment can be more than double the SD column strip moment, depending on the support model.
- The column strip moment approaches the FE peak moment if the concentrated column strip detailing rules specified in SANS 10100 Cl are used.
- The most realistic moment distribution through the footing was obtained from

Table 2 Summary of footing analysis parameters

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat plate type</td>
<td>Pad foundation</td>
</tr>
<tr>
<td>Support in FE analysis</td>
<td>Varies</td>
</tr>
<tr>
<td>Plan dimensions</td>
<td>1.2 m x 1.2 m</td>
</tr>
<tr>
<td>Thickness, $h$</td>
<td>0.15 m</td>
</tr>
<tr>
<td>Concrete strength, $f_{cu}$</td>
<td>36.7 MPa</td>
</tr>
<tr>
<td>Yield stress of reinforcement, $f_{y}$</td>
<td>450 MPa</td>
</tr>
<tr>
<td>Reinforcement (high tensile) Y8</td>
<td></td>
</tr>
<tr>
<td>Concrete modulus of elasticity, $E_c$</td>
<td>33.1 GPa</td>
</tr>
<tr>
<td>Reinforcing modulus of elasticity, $E_s$</td>
<td>200 GPa</td>
</tr>
<tr>
<td>Concrete tensile strength, $f_{t}$</td>
<td>4.9 MPa</td>
</tr>
<tr>
<td>Design uniformly distributed load, $w$</td>
<td>220 kPa</td>
</tr>
<tr>
<td>Design point load applied to column</td>
<td>318 kN</td>
</tr>
</tbody>
</table>

Figure 5: Finite element pad foundation model

![Figure 5: Finite element pad foundation model](image)

Figure 6 Bending moment comparison at the face of the column using different support conditions

![Figure 6: Bending moment comparison at the face of the column using different support conditions](image)

Figure 7 Simplified design moment

![Figure 7: Simplified design moment](image)
Table 3 Effect of support conditions on FE peak and total moment

<table>
<thead>
<tr>
<th>Support type</th>
<th>FE $M_{\text{peak}}$ (kNm/m)</th>
<th>FE W&amp;A $M_{\text{peak}}$ (kNm/m)</th>
<th>SD $M_{\text{col}}$ inner (kNm/m)</th>
<th>FE $M_{\text{peak}}$ / SD $M_{\text{col}}$ inner</th>
<th>FE $M_{\text{total}}$ (kNm)</th>
<th>FE W&amp;A $M_{\text{total}}$ (kNm)</th>
<th>FE $M_{\text{total}}$ (kNm)</th>
<th>FE $M_{\text{total}}$ / SD $M_{\text{total}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Full 3D continuum model</td>
<td>38.86</td>
<td>39.24</td>
<td>33.1</td>
<td>44.12</td>
<td>1.17</td>
<td>0.88</td>
<td>29.68</td>
<td>33.03</td>
</tr>
<tr>
<td>b) Rigid supports (column area)</td>
<td>43.87</td>
<td>47.74</td>
<td>33.1</td>
<td>44.12</td>
<td>1.325</td>
<td>0.99</td>
<td>29.68</td>
<td>32.44</td>
</tr>
<tr>
<td>c) Rigid supports (edges)</td>
<td>50.8</td>
<td>52.6</td>
<td>33.1</td>
<td>44.12</td>
<td>1.53</td>
<td>1.15</td>
<td>29.68</td>
<td>32.03</td>
</tr>
<tr>
<td>d) Rigid supports (corners)</td>
<td>57.6</td>
<td>59.4</td>
<td>33.1</td>
<td>44.12</td>
<td>1.74</td>
<td>1.31</td>
<td>29.68</td>
<td>31.98</td>
</tr>
<tr>
<td>e) Encased supports (fixed at column edges)</td>
<td>52.85</td>
<td>73.8</td>
<td>33.1</td>
<td>44.12</td>
<td>1.6</td>
<td>1.20</td>
<td>29.68</td>
<td>33.12</td>
</tr>
<tr>
<td>f) Spring supports</td>
<td>37.05</td>
<td>42.08</td>
<td>33.1</td>
<td>44.12</td>
<td>1.12</td>
<td>0.84</td>
<td>29.68</td>
<td>33.10</td>
</tr>
<tr>
<td>g) Spring supports at edges</td>
<td>39.74</td>
<td>39.75</td>
<td>33.1</td>
<td>44.12</td>
<td>1.2</td>
<td>0.90</td>
<td>29.68</td>
<td>32.79</td>
</tr>
<tr>
<td>h) Rigid column head (rigid links)</td>
<td>52.85</td>
<td>73.84</td>
<td>33.1</td>
<td>44.12</td>
<td>1.6</td>
<td>1.20</td>
<td>29.68</td>
<td>33.12</td>
</tr>
<tr>
<td>i) Point support</td>
<td>33.48</td>
<td>42.74</td>
<td>33.1</td>
<td>44.12</td>
<td>1.01</td>
<td>0.76</td>
<td>29.68</td>
<td>35.72</td>
</tr>
</tbody>
</table>
modelling the columns support as a 3D continuum and modelling the edge of the column with springs. The peak moment is within 5% of the inner column strip-stepped SD method moment.

■ For both spring models the sensitivity of the model to support models was checked. In general a more realistic moment distribution was obtained as the stiffness of the spring decreased.

■ Ignoring the stiffness of the column, and modelling the support as pinned over the footprint of the column, show reverse curvature in the column area (see Figures 3 and 6). This reduction in moment over the column could be attributed to the fact that fixing the translational degrees of freedom on the column footprint prevents the movement at the nodes, while curvature within the element still occurs and therefore a reduced moment over the footprint of the column is observed. If the rotational degrees of freedom are also fixed in the column footprint, the moment over the column reduces to almost zero, which in reality is impossible.

Adequacy of the simplified method

The pad foundation requires bottom reinforcement in the transverse (x) and longitudinal (y) directions to resist the $M_y$ and $M_x$ hogging moments respectively.

The Simplified Design (SD) method of analysis results in a constant moment, which is then split into a column strip moment and an edge strip moment for the pad foundation. The column strip can be stepped again by concentrating two thirds of the column strip moment into half of the column strip width to form an inner and outer column strip, according to the detailing rules of SANS 10100 Cl 4.6.5.4 (see Figure 7).

Considering the results of the FE analyses with different support constraints, the supports modelled with full 3D continuum are the most realistic and will be used for comparison with the SD analysis. The linear FE moment outputs are the $M_y$ and $M_x$ moments and the commonly used Wood-Armer moments, which include the twisting moment $M_{xy}$.

Figure 8 shows the FE $M_x$ moment contours for the hogging moments in the pad foundation. Figure 9 shows the FE Wood and Armer moment contours for the $M_{xT}$ and $M_{yT}$ hogging moments in the pad foundation.

A section taken through the SD and FE bending moment diagrams at the face of the column in the x and y directions is shown in Figure 10. Both the peak FE $M_y$ and $M_x$ hogging moments occur at the face of the column as there is a cantilever on both sides.
A section through the FE moment contours shows a realistic moment distribution, increasing to a maximum value at the support. To simplify this for the design of flat plates the column strip rules were introduced. The column strip requirement of the simplified design method ensures an increase in the design moment over the column strip as the peak moment in the FE analysis increases. As the FE peak moment increases, the SD column strip reduces, resulting in an increased SD moment.

The integration of the area under the moment diagram gives the total SD and FE load effect. The total SD design moment and total FE $M_x$ moment are the same. This does not change with geometry or constraint model, as the principle of equilibrium has to apply. The total FE Wood and Armer design moments are, however, greater than the SD design moments. These moments were intended for use in design where the twisting moment needs to be considered, and, because of the unique solution and optimisation requirement, the capacity is always greater than the applied moment (Denton & Burgoyne 1996).

**Finite element method:**

**the effect of varying slab geometry**

The influence of slab geometry on an FE analysis was considered by modelling a pad spread footing supporting two columns, and then varying the footing depth. The column size was calculated to meet the criterion of a $0.4f_{cu}$ MPa maximum concrete stress, South African bridge code TMH 7 (1989), in order to maximise local effects. A range of footing depths $h$ was then considered, varying in 100 mm increments from 400 mm to 1 300 mm, with a constant load effect. The chosen variation in footing depth covers a range of reinforcing percentages from maximum to nominal values. It also allows a study of the variation in peak values in FE methods. For the purposes of this study only resistance to hogging (negative) bending moments was considered. Resistance to the sagging (positive) moments, shear and punching forces was not investigated. The authors note that, as a footing’s depth decreases, so punching becomes the governing failure mechanism. Figures 11 and 12 show the overall footing dimensions, and the analysis parameters are summarised in Table 4.

The pad foundation was modelled using the linear elastic FE program Prokon (2012), which is available to the majority of designers in South Africa. The model as shown in Figure 13 consisted of square 0.25mx0.25m plate elements (span/10 is recommended by Brooker (2006)) with column supports modelled as 3D continuum models, as this gives the most realistic moment distribution. The elements used to analyse the pad foundations are discreet Kirchhoff-Mindlin quadrilaterals that provide good results for both thick and thin plates, and are free from shear locking. Shear deformations are not considered here in order to be consistent with the code requirements. The authors note that shear strain and deformation should be considered in thick pad foundations.

The worst-case negative (hogging) bending moment envelope along the face of the support in each direction was then used to design the required finite element model (FE) reinforcement.

**Adequacy of the simplified method**

The pad foundation requires bottom reinforcement in the transverse $(x)$ and longitudinal $(y)$ directions to resist the $M_y$ and $M_x$ hogging moments respectively, and top steel

---

**Table 4 Summary of footing analysis parameters**

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat plate type</td>
<td>Pad foundation</td>
</tr>
<tr>
<td>Support in FE analysis</td>
<td>Full 3D continuum</td>
</tr>
<tr>
<td>Plan dimensions</td>
<td>9 m x 5 m</td>
</tr>
<tr>
<td>Thickness, $h$</td>
<td>Varies from 0.4 m to 1.3 m</td>
</tr>
<tr>
<td>Concrete strength, $f_{cu}$</td>
<td>30 MPa</td>
</tr>
<tr>
<td>Concrete Young’s Modulus, $E_c$</td>
<td>28 GPa</td>
</tr>
<tr>
<td>Concrete tensile strength, $f_t$</td>
<td>2.4 MPa</td>
</tr>
<tr>
<td>Design uniformly distributed load (factored), $w$</td>
<td>347.5 kPa</td>
</tr>
<tr>
<td>Design uniformly distributed load (unfactored), $w$</td>
<td>250.0 kPa</td>
</tr>
</tbody>
</table>

---

**Figure 11** Footing plan dimensions

**Figure 12** Three-dimensional footing model showing moment sign convention

**Figure 13** Finite element footing model

---

**Figure 11** Footing plan dimensions

**Figure 12** Three-dimensional footing model showing moment sign convention

**Figure 13** Finite element footing model
in the longitudinal (y) direction to resist the \( M_y \) sagging moment.

The linear FE moment outputs used were the \( M_x \) and \( M_y \) moments and the Wood and Armer moments. Figures 14 and 15 show the moment contours in the x and y directions for a pad foundation depth of 700 mm. Figures 16 and 17 show the FE Wood and Armer moment contours for the \( M_{xy} \) and \( M_{yt} \) hogging moments in the pad foundation with a depth of 700 mm. The effect of the twisting \( M_{xy} \) moment at the constraints (column) can be seen in the Wood and Armer moment contours.

A section taken through the SD and FE bending moment diagrams at the face of the column in the x and y directions for pad foundation depth of 700 mm is shown in Figure 18. Both the FE \( M_x \) and \( M_y \) moments and the FE Wood and Armer moment are shown in the graph.

The peak FE \( M_y \) hogging moment occurs on the cantilever side of the column, whereas the peak FE \( M_x \) hogging moment is mirrored about the pad foundation centreline, as there is a cantilever on both sides on the column. The Wood and Armer design moments are greater than the SD design moments, as these moments include the twisting \( M_{xy} \) moment. Because of the unique solution and optimisation requirement, the Wood and Armer moment is always greater than the applied moment (Denton & Burgoyne 1996).

The peak FE moment \( M_{peak} \) (maximum FE hogging moment) was affected by the curvature of the pad foundation. As the stiffness of the footing decreased the peak FE moment increased. The SD method of analysis results in a constant moment which is split into a column strip and edge strip moment. The column strip requirement ensures that as the FE peak moment increases with a footing depth decrease, the SD column strip reduces, resulting in an increased SD moment, thus ensuring that the increase in curvature is provided for.

The integration of the area under the moment diagram gives the total SD and FE load effect. The total SD design moment does not vary with the change in pad foundation depth, as the self-weight of the pad foundation does not have an effect on the applied load effect. The total FE \( M_x \) and \( M_y \) moments are also constant with respect to change in footing depth, and are the same as the SD design moments. The FE Wood and Armer moments are affected by the twisting moment, which in turn is affected by how the constraints are modelled, and slab geometry. An increase in slab stiffness leads to an increase in the twisting \( M_{xy} \) moment. A comparison of the total FE \( M_x \) and \( M_y \) moments (equal to total SD
moment) to the total FE Wood and Armer moment for $M_y$ and $M_x$ is shown in Figures 19 and 20.

The total FE Wood and Armer moment $M_y$ (cantilever) was approximately 5.6% greater than the total SD moment at a depth of 400 mm, increasing to approximately 6.9% at a depth of 1 300 mm, i.e. a 23% increase. The principal axis moment was the same as the SD moment.

The total FE Wood and Armer moment $M_x$ (beam) was approximately 10.5% greater than the total SD moment at a depth of 400 mm, increasing to approximately 16.5% at a depth of 1 300 mm. The principal axis moments were the same as the SD moment.

Peak load effects (singularities) in linear FEM

Figure 21 shows the change in the SD moment, column strip and middle strip, compared to the FE peak moment ($M_{\text{peak}}$), as the depth of the pad foundation (h) varies. Both the $M_y$ and $M_x$ moments and Wood and Armer moments were plotted, and the SD requirement of differentiating between the column and middle strip is shown. For a pad foundation width of greater than $1.5 \times (b_{\text{col}} + 3h)$, the transverse distribution of curvature has reduced sufficiently so as to not warrant the differentiation between the column and middle strip.

The $M_y$ FE $M_{\text{peak}}$ moment remained constant as the footing depth increased. The $M_y$ FE $M_{\text{peak}}$ moment decreased with the increase in footing depth, until a footing depth of 1.100 mm, and then levelled out.

Figure 22 shows a comparison of the ratio of the FE $M_{\text{peak}}$ to the SD column strip moment and the SD concentrated column strip moment.

Both the $M_x$ and $M_y$ FE $M_{\text{peak}} / SD M_{\text{column}}$ ratios approach one as the stiffness of the slab decreases, and levels out to a constant value at a depth consistent with the limit for the column strip of the code. Again, showing that the SD column strip requirement ensures that as the FE peak moment...
increases with a footing depth decrease, the SD column strip reduces, resulting in an increased SD moment, thus ensuring that the increase in curvature is provided for.

**Observations from analysis**

From the above numerical analyses the following can be concluded about the simplified method overall strength, finite element peak values and detailing according to linear FE methods:

- The total FE $M_x$ and $M_y$ moments are the same as the total SD moment.
- Both the simplified method and finite element analysis and reinforcement layouts provided adequate and similar flexural capacity.
- The FE peak $M_x$ or $M_y$ moment can exceed the column strip moment by a significant amount, depending on how the support constraints are modelled. It is, however, commonly assumed that this peak is reduced by cracking of the concrete and yielding of the reinforcement.
- The peak and total Wood and Armer moments obtained from a linear FE analysis are significantly influenced by how the supports are modelled. This is because of the change in the twisting moment with the support/constraint model.

**EXPERIMENTAL INVESTIGATION**

Preliminary experiments carried out at the Department of Civil Engineering at the University of Pretoria support the above
Table 5 Footing test parameters

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat slab type</td>
<td>Foundation</td>
</tr>
<tr>
<td>Support</td>
<td>Springs (k = 2 500 kN/m)</td>
</tr>
<tr>
<td>Plan dimensions</td>
<td>1.2 m x 1.2 m</td>
</tr>
<tr>
<td>Thickness, ( h )</td>
<td>0.15 m</td>
</tr>
<tr>
<td>Concrete strength, ( f_{cu} )</td>
<td>36.7 MPa</td>
</tr>
<tr>
<td>Concrete Young's Modulus, ( E_c )</td>
<td>33.1 GPa</td>
</tr>
<tr>
<td>Concrete tensile strength, ( f_{r} )</td>
<td>4.57 MPa</td>
</tr>
<tr>
<td>Reinforcement (high tensile)</td>
<td>Y8</td>
</tr>
<tr>
<td>Yield stress of reinforcement, ( f_{y} )</td>
<td>450 MPa</td>
</tr>
<tr>
<td>Design point load applied to column</td>
<td>318 kN</td>
</tr>
</tbody>
</table>

Figure 23 Reinforcement numbering for SD and FE footings

Figure 24 Spring layout

Photo 1 LVDT placement on the footing supported on springs (Hossell 2012)
numerical analysis. Hossell (2012) undertook tests on reinforced concrete footings supported on springs where two specimens were designed and reinforced, one according to the SD method and the other according to the linear FE method. The influence of the reinforcement layout on the response of the footing to ultimate limit state and serviceability limit state characteristics was observed.

The spring-supported footing test setup is shown in Photo 1, with the springs simulating the founding support conditions. The test parameters are shown in Table 5 and the reinforcement and spring support layouts are shown in Figures 23 and 24. Footing (a) was reinforced according to an SD analysis, and Footing (b) according to an FE analysis. Strain gauges were placed on the flexural reinforcement bars at the critical design section along the face of the column, and LVDTs at the centre of each support spring were used to measure the displacement of the footing. The strain in the reinforcement across the footing was logged at a rate of 1 Hz. The change or variation in strain is shown in Figures 25, 26 and 27; and a summary of the footings response to the load is included in Table 6. It should be noted that, as a result of using the Wood and Armer moments to calculate the FE reinforcement, the FE footing required slightly more reinforcement than the SM footing.

**Figure 25** Strain in reinforcement with applied load for (a) SD and (b) FE footings

**Figure 26** Strain in reinforcement prior to cracking for (a) SD and (b) FE along section AA

Strain in reinforcement with load
Figure 25 shows that the first crack (sudden "jump" in strain) in both footings occurred at very similar loads, as this is primarily dependent on the tensile strength of the concrete. The SM footing test had to be stopped at a load of 412 kN, before failure, as the testing machine piston moved out of alignment. The FE footing failed in punching at 480 kN. At the design load of 318 kN the reinforcement strain in both footings was well below the yield strain. The yield strain was calculated using the 0.2% proof stress method described in TMH 7 (1989).

Transverse variation in reinforcement strain prior to cracking
Figure 26 shows that prior to the concrete cracking the reinforcement at the face of the footings (i.e. design section) is strained the most. The SD footing had a greater variation in strain between the reinforcement under the column and the reinforcement at the edge, than the reinforcement in the FE footing. The reinforcement strain in the FE footing appears to be more uniform across the footing width than when compared with the SM footing.

Transverse variation in reinforcement strain after cracking
Flexural cracking occurred at an applied load of approximately 205 kN in both
footings, as indicated by the sudden increase in strain in the reinforcement shown in Figure 25. The increase in strain in the central reinforcement shown in Figure 27 indicates the formation of cracks at the face of the column, and shows the transfer of force from the concrete to the reinforcement.

Once the concrete cracked, a greater variation in strain was observed in the SD footing, compared to the FE footing, shown in Figure 27. With a larger variation in strain the reinforcing bars beneath the column in the SD footing strained more than the bars towards the edge; indicating that fewer, but larger, cracks developed when compared to the FE footing. The FE footing showed a more uniform variation in strain, indicating that more cracks had formed, but because of the lower strain levels these cracks were smaller.

**Load-deflection curves**

Figure 28 shows the load deflection curves at the centre of the two footings. Cracking and flexural failure can be seen by the change in gradient of the curves. First crack occurred at very similar loads and deflections for both the FE and SD footing.

**Observations from experimental work**

From the above experimental work the following can be concluded regarding SD and FE analysis and design:

- Both the simplified method and finite element designs provided adequate and similar flexural capacity.
- Detailing reinforcement in accordance with the variation in moments produced from a linear finite element analysis results in a more uniform distribution of strains across the width of the footing before and after cracking occurs.
- Cracking would appear to be controlled by reinforcing to follow the finite element peak moment.
- There is no apparent benefit in controlling deflection by reinforcing to follow the FE peak moment.

**CONCLUSIONS**

**Total resistance achieved with FE design compared to that of traditional methods**

Both the numerical analysis and experimental work support the conclusion that at the ultimate limit state there is very little...
difference, if any, between a flat plate analysed and reinforced using the SD method and one analysed with an FE model.

For each different flat plate structure modelled, irrespective of the support model, the total FE $M_x$ and $M_y$ moment was the same as the total SD moment. The total FE Wood and Armer moment was always greater than the SD moment, i.e. design moments which include the $M_{xy}$ twisting moment.

**To what extent the peak moment in an FE analysis can be ignored**

The support (constraint) model has a significant effect on peak moments calculated in an FE analysis. If the stiffness of the column/support was taken into account, peak moments were not observed in the FEM analysis and a very realistic moment distribution was obtained (Figure 29). Pinned supports are not advised, as the stiffness of the support must be taken into account.

The peak moment from the FE model may exceed the SD column strip moment. However, this peak is reduced by cracking of the concrete and yielding of the reinforcement. This may be compensated for by considering a column strip with a reduced width.

**Serviceability performance of an FE design**

Detailing reinforcement to follow the FE moments at the serviceability limit state results in a more uniform distribution of strain across the width of the slab, and therefore more, but smaller cracks.

The reinforcement distribution according to the FE method does not have a significant effect on the overall stiffness of the slab, and therefore does not appear to influence the deflection of the slab. This was shown in both the experimental testing performed on the signal column footing and on the flat slab.

**ACKNOWLEDGEMENTS**

Shane Hossell is gratefully acknowledged for his contribution to this paper.

**REFERENCES**


Hossell, S M 2012. Comparison of different methods of analysis, design and detailing of reinforced concrete footings. Project report for BSc (Civil Engineering), Pretoria: University of Pretoria.


PROKON Software Consultants Ltd. 2012. PROKON suite of structural analysis programs Pretoria, South Africa: PROKON.


Towards a systems thinking approach in allocating infrastructure budgets in local government

G N Kaiser, J J Smallwood

Legislation, systems and processes provide a rigid framework within which local government operates. In a perfect world, perfect information would result in perfect decisions. However, the world is chaotic, people are far from perfect, and priorities vie for resources, as is evidenced through the social inequality and environmental degradation in society. A common way of managing complexity is to break complex processes into manageable portions. In doing so, the advantage of integration is often lost. Furthermore, silo thinking tends to yield duplication, often through localised efficiencies at the cost of overall value. System thinking can be employed to facilitate a constructive, creative space that has potential to evolve in the presence of intuitive trust, combined with a framework with which to measure and verify against. The City of Cape Town serves as example of a local authority grappling with allocation of resources in a manner which is fair and efficient in fulfilling its electoral mandate.

**INTRODUCTION**

Most people are faced with making budgetary decisions daily – only a lucky few can perhaps afford to buy without compromise. Where money as a resource is scarce, trade-offs invariably have to be made. If the household has no money, then there are no choices to be made in allocating money. The complexity of allocating the household budget is thus initially premised on the budget-existing. Once a budget is available, other factors influence complexity of decision-making, such as the structure and size of the household, standard of living, culture and value systems. Whatever the case may be, household budget choices are generally made by the head of the household, who has a fair grasp of the scope and complexity of internal budget options, expectations and priorities. The decision around budget spend is purely financially based for anyone living in poverty, but includes other factors as soon as enough money is available to provide choice – in other words, it becomes an economic choice.

The aim of this research is to provide a model or models on which to base budgetary allocation decisions in the City of Cape Town, in an environment where funding constraints exist.

Local government is charged with the provision of free basic services necessary for a reasonable quality of life, without which either health or the environment could suffer. This includes a vast scope of services requiring significant funding to maintain and augment, including utility services such as water and electricity, mobility infrastructure such as roads and public transport, and social and recreational facilities such as parks, sports fields and community halls. To be of value, such models would need to be simple, inclusive, transparent, flexible and capable of providing repeatable results.

A comprehensive library of legislation cascades from the Constitution down to detailed National Treasury practice notes – guiding, controlling, empowering and regulating the municipal budgetary process. Applying rigorous legislation to complex systems will not necessarily result in good decisions, just as one cannot regulate for honesty or competence. Alternative solutions based on pattern and systems thinking are likely to yield better solutions, and will be explored in the context of the economic impact of budgetary allocations.

**INFRASTRUCTURE AND SCARCE RESOURCES**

South Africa is grappling with the reality of a world where resources are scarce. The term “resources” may be interpreted broadly to include everything from minerals, products, utility services, soft services, skills, and labour, to money. The economy is driven by producers selling their wares at the highest possible price, while consumers aim to purchase at the lowest possible price. When
a particular resource is abundant, the price decreases. The market players thus aim to keep resources scarce as this drives capitalism (Ellis 2009: 15).

Legislation tabled subsequent to the 1994 elections has provided rules, controls and processes in pursuit of equitable allocation of resources across the three spheres of government, with a focus on correcting imbalances created by apartheid. The legislation is sound and can be implemented to the letter, but yet often fails to yield the best result for reasons as simple as having a span of control that is too large to enable fair consideration of all variables and options. Given the need for infrastructure provision, South Africa can ill afford to allocate an oversubscribed budget sub-optimally.

This is emphasised by National Treasury in stating in the introductory paragraph of a recent Municipal Finance Management Act (MFMA) circular that reminded municipalities to do more with existing resources (South Africa 2010).

Allocation of scarce resources has and continues to confound institutions of all kinds. In the public sector, the problem of attributing value to products and services is further complicated by collective social services, such as parks, policing and security (Pauw et al 2009: 50–51).

Despite comprehensive legislation, public participation and political oversight, decision-making around budgetary issues is difficult and the value hard to measure. Infrastructure is under pressure in South Africa, and resources to fund new and maintain existing infrastructure are not abundant.

LITERATURE

The impact of growing worldwide inequality has never been as evident as in 2012 (the year during which this paper was initiated). On the one hand, the world continues along the well-established economic path with little apparent concern of the worldwide impact of growing inequality. On the other hand, inequality is mooted as cause for wide-spread uprising, including:

- The Occupy Wall Street, otherwise known as the 99% movement, and
- The Arab Spring or revolution, which since December 2010 removed leaders from power in Libya, Egypt, Tunisia and Yemen, and sparked widespread protest across much of the Arab world.

In South Africa, Thabo Mbeki warned of the risk attached to the concept of “two nations” occupying the same space. In response to the formal economy’s marginalisation of the poor, a second economy developed, with very little real progress in converging (Manning 2006: 17). Causes of the mass public protest are thought to include a combination of factors, including:

- Inflation causing reduced buying power of currency leading to a drop in standard of living and a lack of access to health care, education and basic services
- Un- and underemployment leading to increased poverty and dependence
- Political and religious oppression
- Autocracy with little or no political opposition or public participation, often controlled by force, such as in a police state
- Social imbalance and inequality. Inequality is commonly justified in that it provides incentive for individuals to work harder and thus prosper. The relative wealth adds to the incentive and drives economic growth (Galbraith 1969: 68). Redistribution of wealth has found less favour than growth has in improving the lives of the impoverished. Eradicating inequality is resisted by the wealthy on the grounds that it undermines arts and education, leads to uniformity and monotony. Galbraith also developed the “theory of social balance” which refers to the balance between private goods and public goods. He argues that public goods are artificially highly valued thanks to demand increased through advertising and emulation of conventional prosperity. So the public is prone to recognise novelty as adding value, while the less attractive social services are grudgingly paid through taxes (Galbraith 1969: 188).

The South African government embarked on a strategy to create a dynamic economy (Joyce 2007: 195) in the development of a macro-economic policy titled the Growth, Employment and Redistribution (GEAR) policy. The policy was approved in June 1996 as a five-year plan to support the implementation of the Reconstruction and Development Programme (RDP). The aim of the policy was to strengthen economic development, reduce unemployment by expanding the employment base, and redistribution of opportunities and income (Padayachee & Valodia 2001: 78).

GEAR relied on a market-friendly environment inspiring private sector investment to drive the social initiatives. By 2001 the economy had stabilised, but GEAR had failed to achieve its goals (Padayachee & Valodia 2001: 78). The next phase was characterised by an expansionary budget aimed at balancing broad objectives of poverty reduction and providing a safe environment, infrastructure investment and maintenance of assets, reducing the tax burden to stimulate job creation and household spending, and reducing interest rates.

This was followed in 2006 by ASGISA (Accelerated and Shared Growth Initiative for South Africa), with principles focused on achieving the millennium goals of halving unemployment and poverty by 2014. The merits of ASGISA lay in the wide consultation during its development, as well as the holistic developmental perspective (South Africa 2006). To date, the benefits have not accrued to any noticeable extent.

The remarkable economic stability encountered under Trevor Manuel (Green 2008: 538) as Minister of Finance did not manage to live up to expectations in terms of job creation. The economy had grown, and could not be faulted for its stability when compared to other economies, but the formal economy had simply failed to provide the number of jobs required to alleviate poverty. Pursuing the number of imperatives requiring prominence that the new government faced, it is not surprising that job creation did not follow the stability of persistent growth (Nattrass 2008: 8).

Fedderke et al (2001: 5) explored the impact of infrastructure spending on economic growth over 125 years, as it relates to rail, roads, electricity and telecommunications. The research results indicated that the economic impact was distinguishable in terms of certain types of infrastructure, but more difficult to quantify in others. The concluding supposition suggested that large-scale infrastructure investment could catapult South Africa beyond the 3% economic growth level. Bogetic and Fedderke (2005: 2–34) further confirmed that large-scale investment in infrastructure was required to remedy the underinvestment which had occurred in the past, and that the value of infrastructure in South Africa was generally comparable to that of other centres in the world. If the available funding were better applied, it could thus increase the value of the end product.

It is important to consider the developmental nature of South Africa, and thus municipalities, to fully comprehend the state of infrastructure and local government capacity, as well as the necessity for the rigorous legislation and controls, all of which are aimed at protecting the people. The Integrated Development Planning (IDP) process was introduced in 2008 in an effort to provide a mechanism whereby newly formed municipalities could perform in a “coordinated, strategic, developmental and fiscally responsible way” (Van Donk et al 2008: 323). Critique of the IDP process is that it is a rigorous technocratic tool – while functionally correct, it is unlikely to provide anywhere close to an optimum outcome.

Since 2010 service delivery protests have become the norm. According to Community
Protests in South Africa (Karamoko 2011: 2), lack of accountability and absence of true public participation are aggravating factors in service delivery protests. While the number of monthly protests varies depending on the criteria used, protest action is widespread throughout the country, with not a month passing without a protest. Since 2007 the Western Cape has been in the top three provinces in terms of protest action.

It is evident that an alternative paradigm of problem solving is required, as the programmes aimed at “fixing society” have been unimpressive. System-thinking approaches any complex system, such as an organisation, city or region, as a whole as an integrated system of which all component parts have an impact or are impacted upon by all others (Senge 2006: 7). Taking a holistic view provides space for intuition. Subconsciously the human mind is capable of integrating and dealing with great complexity. A system approach provides scope for unconventional solutions and a humanist approach to managing, empowering and providing for people. It requires a mind shift away from the conventional analysis of components towards assessing and problem-solving entire systems.

In the context of using system thinking as a way of improving the function of local government holistically, it is necessary to unpack some of the theory behind system thinking as the fifth discipline. Added to the disciplines required of a learning organisation, of personal mastery, shared vision, mental models and team learning, is system thinking, the underlying worldview, in all its complexity. Personal mastery refers to “consistently realising results that matter deeply to them”, anything from a physical skill to a job function. Shared vision, if achieved, inspires people to attain their best effort in all they do, not simply to comply, but with enthusiasm in sincere pursuit of the vision. Mental models are deeply entrenched views of the world, preconceptions, often providing a skewed picture of reality and stifling the opportunity to grow and learn. Team learning requires a creative space where honest conversations can lead to the necessary security and freedom of thought for a group of people to debate constructively, to think together, thereby maximising the possibilities for positive outcomes (Senge 2006: 8–10).

Dealing with complexity is a recurring theme in texts offering alternative models for creating improved solutions. From complexity economics to dealing with complexity in system thinking, the propensity to divide complex organisations into manageable chunks of information has been shown to achieve fragmented results. Organisations tend to organise into hierarchical structures, which completely eliminate any organic environment.

In *The Origin of Wealth*, Beinhocker (2006: 347) argues for an “adaptive social architecture”, which will allow organisations to adapt to change intrinsically and organically rather than through policy and structural provisos. Legislation and policy in government are slow-moving and need to follow a path of public participation and lengthy approval processes, constraining government bodies, and constraining adaptability. In a complex adaptive system, it is necessary to view human behaviour, rather than rational behaviour. The premise of human resources in local government is that, while it is recognised that employees are important in fulfilling the constitutional mandate, they are not valued to the extent that they feel secure in their positions, or are recognised as individuals. The legislation informs that processes should be as mechanistic as possible, not providing for individual thought or input. In a complex system, this is thought to be unlikely to provide much in the line of value.

In *False Economy*, Alan Beattie (Beattie 2009: 247) defines path dependence as resembling “evolutionary biology – the role played by a sequence of events, some of which may come by chance”. An irresistible example quoted by Beattie is that of the configuration of a keyboard – that the layout is designed to slow down typing and specifically to prevent hitting adjacent keys in quick succession. This arose when a particular mechanical typewriter malfunctioned often due to speed of typing. The current layout has long lost the original intent, but inertia and the “network effect” or critical mass of use of a standardised layout means it is unlikely to change. The adoption of QWERTY keyboards in smart phones such as BlackBerry is an indication that the layout is here to stay.

In the *Black Swan*, Nassim Taleb (Taleb 2007: 41) emphasises the security that man has chosen to derive from statistical analysis by adopting models such as the bell curve and linear regression to predict the future. To easily understand the concept, he posits the example of a free-range turkey that roams around the farmyard. The turkey is fed every day for a number of years, say a thousand days. Using linear regression the turkey thus anticipates being fed daily until eternity, until the day arrives that the turkey is not fed, but has its throat cut. This turns the linear regression curve upside down. Taleb’s issue with the bell curve is that the outliers tend to have a far more dramatic and unpredicted effect than ever credited.

The norm is used to assess risk and map the future. His ideology provides scope for alternative modelling.

From an environmental economics perspective, the focus has moved from conventional through green, sustainable and restorative, to regenerative design that considers a living system in its entirety (Reed 2009: 9). The evolving system is based on patterns in nature and is seen to comprise at least the impacts of global warming, water scarcity, destruction of habitat, social justice, pollution, toxins and degradation of raw material.

A model directly related to infrastructure is explained in *Country-led monitoring and evaluation systems* (Segone et al 2008: 56), which is predicated on establishing a strategic intent and pursuing this intent above all else. Common themes throughout these texts lie in dealing with whole systems, which are recognised in all their complexity. The importance of common vision is emphasised, as the work of man is generally performed for the benefit of man.

As a local authority, the City of Cape Town is beholden to legislation, emanating from central government, which is based to some extent on events in the world. Technology enabled the world to become a global village, but it also emphasised that every country is connected, has an impact on, and is impacted on by world events. The system works because countries cooperate. So-called successful countries in particular have no incentive to be the leader in breaking out of the system.

MODELS FOR IMPROVED DECISION-MAKING

A number of models are proposed to assist local authorities, such as the City of Cape Town, in improving decision-making around budgets. It is possible that further benefits will accrue due to the variety of the models proposed. Models provide frameworks, with embedded information, which can make it easier to deal with detail, which in turn provides space for evaluating alternatives and analysing social issues in more detail. Models tend to simplify data, yielding approximations of the truth, resulting in some inaccuracy in detail. Transformation of mind-set is difficult to quantify, but transformed individuals behave in a manner that has intrinsic value. Experiencing an environment of trust, while not tangible, is nonetheless a real experience with real benefits.

Over the years the City of Cape Town has used various prioritisation models, applicable mainly to capital expenditure, although they have been applied to projects funded by the
operating budget. Typically, projects are scored in weighted categories including risk, compliance, IDP alignment and backlog, and a final score achieved per project. The efficacy of such models can be hampered through the crafting of creative motivations for funding – for example the strategic objectives of a local authority tend to encompass an entire spectrum of social and economic services, providing space to motivate virtually any expenditure.

Coversing the business of local government in all its complexity would require a highly specialised model, which would need specialised skills to implement. In reality, any model that attempts to grade initiatives numerically will be of value, particularly if combined with interventions with the potential to enhance trust, integration and cooperation. Thus a combination of models and strategies addressing different angles of the problem has the potential of being more practical, broadly applicable and implementable than a single model.

Teamwork determines that combinations of people derive better solutions in interaction than the sum of the individuals. Diversity multiplies value in any team. This happens through combining different perspectives, interpretations, heuristics and predictive models into a framework or cognitive toolbox (Miller & Page 2007: 8). Although diversity may complicate the process of agreeing rational choices, pursuit of a common goal paves the way for more integrated decisions.

Whichever models are implemented, certain aspects of the budget process are non-negotiable:
- Good corporate governance in the broadest possible sense
- Compliance with legislation
- Fiscal discipline

Regional prioritisation model
Apartheid planning resulted in an undeniable disparity of service delivery within the municipal area. The current administration purports to have a pro-poor focus, but this is not as evident on the ground as it may have been expected to be. A regional prioritisation model proposes to integrate geographic location with service delivery and socio-economic measures.

Sub-councils have been included in the budget process, entrenching a more regional approach providing the opportunity to submit budget priorities. Information may be available in different forms depending on the directorates, but needs to be consolidated to provide a graphical depiction by a sub-council and should include:
- Population, and number of households
- Service statistics
- Income level
- Occupational status
- Education
- Community satisfaction
- Crime statistics
- Unemployment
- Condition and availability of infrastructure
- Bulk votes across boundaries.

The identification of areas in need of basic services does not necessarily require an academic exercise, and could easily be identified visually, with informal settlements displaying the largest lack of service and thus in most urgent need of intervention. Mapping is, however, required to provide structure to the cross-subsidisation that exists, to ensure an incremental improvement without the balance of areas becoming degraded in an uncontrolled manner.

The regional prioritisation model above is indicative of the value that could be harnessed through the combination of geographic information and meaningful statistical data. The actual content can be adapted to the requirements of the leadership using up to date spatial and census information.

Asset management model
The concept behind an asset management model in providing for efficient budgeting is that information is available and accessible, informing the actual requirements to cover the cost of maintaining assets in a serviceable condition, thus avoiding indirect fruitless expenditure in allowing asset deterioration. While there is a broad understanding of the constraints provided by existing municipal service infrastructure, which restricts approvals of new development, detailed information on infrastructure items is required to determine best value investment. In complex organisations such as municipalities, the task of tracking asset life cycles is onerous, but the benefits of having a working model in place can significantly change the required funding cycles.

The City of Cape Town has assets of which the variety and value are vast and way beyond the scope of this study, and the intention is not to develop a management model to suit all assets, but rather to encourage such development. Many resources are available to guide the management of municipal assets, such as the International Infrastructure Management Manual from the Institute of Municipal Engineering of Southern Africa. The asset inventory needs to be part of the SAP ERP system, having the benefit of being managed and recorded in an integrated fashion.

Modelling the IDP to a common vision and value system
Since the introduction of municipal integrated development plans in the MFMA, the IDP has been used to motivate budget applications. As the IDP covers virtually every conceivable reasonable activity, almost any programme can be motivated under at least one of the objectives. Beyond this stated alignment and the annual review of the document as required by legislation, the IDP document is voluminous, creating a challenge in providing a common vision or shared strategic intent.

The eight strategic focus areas of the previous five-year term of office of the IDP were replaced by five pillars in the IDP applicable to the 2012–2017 period. The pillars are the expressed outcome of interaction with the people of Cape Town. The political leadership’s vision for the term of office is to “create a more inclusive society by working towards greater economic freedom for all people of the city” (City of Cape Town 2012: 14). While the political leadership has a clear vision tied to the IDP, the shift from the previous IDP is perhaps not immediately apparent to the average employee. The political leadership would like to make a difference in the lives of the ordinary people of Cape Town. It is suggested that the City of Cape Town determines a clear, inspirational vision statement, crystallises a strategic intent and undertakes a programme to align the staff with this vision. A vision statement should be memorable, moving, energising. Some options are explored in Figure 1.

Benign visions of a local authority include conservative, moderate and mechanistic visions, as reflected in Figure 1, such as the provincial government ambition (Western Cape Provincial Government 2007: 5). At the other end of the spectrum lie radical visions, such as that of Bhutan. Bhutan has developed an alternative measure of wellbeing to GDP, which considers happiness as an important component of wellbeing (Bhutan Government 1999: 6).

The risk in adopting anything other than standard capitalistic goals and measures rests in reduced competitiveness in a global world.

In so far as a value system is concerned, the shortest route to a set of sound values exists in the Batho Pele principles, developed by central government in 1997 and applicable to all government departments. It may be of merit for the municipality to revive these values and ensure assimilation by the wider employee base. The values are displayed across the City of Cape Town in municipal offices (South Africa 2007).
Staff transformation model

The relatively peaceful transition of South Africa into a democratic country has been captured in the annals of history. The lack of social transformation nearly twenty years later, with ever-growing deficits in housing, employment, education and healthcare, and widening gaps in social inequality is a blight on the success story. While official census data for employment and population numbers were released in 2013, the number of unemployed in the City of Cape Town still falls between 300 000 and 500 000, assuming approximately 26% unemployment. The City of Cape Town employs in the region of 27 000 people. By definition of being employed, municipal employees are fortunate, and in a position to have a positive impact on society. To have maximum impact, officials need to focus on working towards the greater good, setting narrow self-interest aside. Modern society has not been very successful in acknowledging the value of socially conscious citizenry. Broader society may thus benefit from the transformation of officials into conscious citizenry and protectors of social justice. If successful, training interventions addressing aspects such as social justice, freedom and kindness could inspire sufficient trust, integrity and commitment to ensure that decisions made at all levels by all officials have a positive impact, which will have a net benefit to the municipality.

Transforming mind sets may be a difficult task, but at a relatively low cost it can have one of the most powerful positive impacts, together with a sense of satisfaction of making a lasting difference extending far wider than the workplace.

If members of staff were to operate in an environment of trust, and act in the best interest of the City of Cape Town, gains could be made. Transformation intervention has the potential to cover a range of imperatives to result in a more motivated, aligned, engaged staff. Transformation is a process rather than an event; while facts illustrate concepts, value is gleaned from the experience rather than a measurable outcome such as test results.

Improved integration model

The City of Cape Town has close to four million inhabitants, with any progress in service delivery resulting in growth caused by influx from elsewhere in South Africa. The administration covers an area of 2 470 km², through approximately 27 000 permanent staff, with a budget of close to R30 billion. The formation of the City of Cape Town was premised on a centralised model, with function-specific directorates. The scale of operations resulted in a silo approach, which provides sufficient texture to deter cross-functional interaction, unless corporately forced.

Over the years, the leadership has instituted various crosscutting committees and reporting structures. Interaction combines political and administrative spheres, and occurs at high level.

The current leadership introduced a cluster system as part of a transversal governance framework. What is, however, lacking is integration at a more operational level. As all staff form part of the hierarchical management structure, line management can play a valuable role in enabling the necessary integration through proper communication, covering functions, progress, policies and strategy, IDP, performance and vision, which will aid in empowering and transforming the workforce as well, and can be reinforced by electronic communication.

Triple bottom line model – financial, social and environmental reporting

The King Committee on Governance, in their report on governance for South Africa, emphasised the importance of leadership, integrated sustainability, corporate citizen- ship and social transformation. The City of Cape Town has accepted King III. Triple bottom line reporting provides the opportunity to further integrate the strategic focus areas of the IDP. By inclusion of comprehensive and holistic information in an area where, historically, finance mattered at the expense of society and the environment, the transformation to a responsible citizenry will be further embedded.

For reasons of size, complexity and history, discrete projects only cover a very small portion of the total operating expenditure. If the City of Cape Town is to transform, it will be necessary to create a more flexible structure, with commitments either reduced

Figure 1 IDP objectives translating into vision statements

<table>
<thead>
<tr>
<th>Possible vision statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conservative: The be the best metropolitan city in South Africa</td>
</tr>
<tr>
<td>Moderate: To build on the City’s status of the best metropolitan city in SA by becoming better at everything we do</td>
</tr>
<tr>
<td>Aspirational: To be a world-class city in every sense</td>
</tr>
<tr>
<td>Idealistic: To build a society based on equity, transparency, social justice and environmental sustainability</td>
</tr>
<tr>
<td>Radical: To pursue peace, redefined prosperity, wellness and happiness of all our people</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Based on the five pillars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mechanistic: To be the leading city in Africa, our goals are entrenched in five pillars:</td>
</tr>
<tr>
<td>1. Opportunity city: Providing an enabling environment for investment and job creation</td>
</tr>
<tr>
<td>2. Safe city: Ensuring the safety and security of all our people</td>
</tr>
<tr>
<td>3. Caring city: Affording people dignity and hope</td>
</tr>
<tr>
<td>4. Inclusive city: Including all people in our plans and activities</td>
</tr>
<tr>
<td>5. Well-run city: An administration that can be relied upon to make sound decisions and follow good corporate governance principles in everything</td>
</tr>
</tbody>
</table>

| Abbreviated: |
| (a) To provide all citizens with opportunity within a safe and caring city, which is well-run, through the leadership of the City of Cape Town, or (b) the City of Cape Town has a vision to create a more inclusive society by working towards greater economic freedom for all people of the city, or (c) this City works for you |

<table>
<thead>
<tr>
<th>Western Cape Provincial Government 2040 ambition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A resilient, inclusive and competitive Western Cape with higher rates of employment producing growing incomes, greater equality and an improved quality of life.</td>
</tr>
</tbody>
</table>
or made more mobile. Commitments extend beyond contractual agreements on projects and include costs that have been accepted as part of the fixed establishment, such as the cost of staff, which contributes roughly a third of expenditure, the cost of accommodating the staff establishment, such as office space, transport, parking and many other support services. The cost of compliance is unavoidable, but it may be interesting to quantify this and then gauge the effectiveness against real or perceived corruption vis-a-vis other initiatives such as empowerment, to curb poor performance and unprofessional behaviour.

Triple bottom line management has to extend beyond compliance and reporting, and should form an intrinsic part of all activities, providing balanced value attributed to social, environmental and financial imperatives.

MODEL VALIDATION

To conclude the development of the model, a validation survey was conducted among senior staff to gauge opinion with respect to the effectiveness of the combined model. Some of the models proposed may include interventions already implemented, and some of the data change continuously. The survey provided a summary of each of the models, and provided for each to be rated in terms of suitability, practical relevance, ease of implementation and potential to change the budget.

The validation was premised on a small sample of senior officials with extensive experience in budgeting at the City of Cape Town. Individuals were targeted to participate based on their involvement and experience in budgeting both from finance and line departments, particularly Utility Services, as it is responsible for approximately two thirds of the city’s budget. The sample consisted of 25 individuals in reporting levels 0–4, ranging from the municipal manager to branch heads.

The results returned on the surveys indicated that all proposed models were deemed to be relevant. The difference between the top and lowest model’s MS calculated as 0.68 of a point, with all scores falling between 3.15 and 3.83. Only five out of the 600 responses indicated that a model “does not” have the potential to impact on the efficacy of budget allocation, equating to less than 1%.

The summary scores are indicated in Table 1 and Figure 2.

The overall average of 3.43 indicates that the survey sample accepts the models proposed as valid. The comments proffered were favourable towards initiatives aimed at increasing value. A caution was, however, raised that, without political support, the best of interventions were likely to have negligible impact.

Table 1 Summary model scores

<table>
<thead>
<tr>
<th>Factor</th>
<th>Suitability</th>
<th>Practical relevance</th>
<th>Ease of implementation</th>
<th>Potential to change budget</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional prioritisation model</td>
<td>3.58</td>
<td>3.65</td>
<td>3.26</td>
<td>3.83</td>
<td>3.58</td>
</tr>
<tr>
<td>Asset management</td>
<td>4.00</td>
<td>4.00</td>
<td>3.48</td>
<td>3.84</td>
<td>3.83</td>
</tr>
<tr>
<td>Common vision &amp; value system</td>
<td>3.28</td>
<td>3.32</td>
<td>3.17</td>
<td>3.32</td>
<td>3.27</td>
</tr>
<tr>
<td>Staff transformation</td>
<td>3.40</td>
<td>3.29</td>
<td>3.00</td>
<td>3.00</td>
<td>3.17</td>
</tr>
<tr>
<td>Improved integration</td>
<td>3.84</td>
<td>3.72</td>
<td>3.21</td>
<td>3.50</td>
<td>3.57</td>
</tr>
<tr>
<td>Triple bottom line</td>
<td>3.39</td>
<td>3.39</td>
<td>2.73</td>
<td>3.09</td>
<td>3.15</td>
</tr>
<tr>
<td>Total average</td>
<td>3.58</td>
<td>3.56</td>
<td>3.14</td>
<td>3.43</td>
<td>3.43</td>
</tr>
</tbody>
</table>

Figure 2 Model mean score by factor
that, for reasons of the complexity of the municipal finance environment, no single model would have sufficient depth, offset with simplicity, to realistically contribute to better decision-making, resulting in a combination of models. The validation survey confirmed that the models could have a positive impact on optimising budget allocation.

REFERENCES
ANALYSING DELAY AND QUEUE LENGTH USING MICROSCOPIC SIMULATION FOR THE UNCONVENTIONAL INTERSECTION DESIGN: SUPERSTREET

H H Naghawi, W I A Idewu

With the increasing demand on today’s roadway systems, intersections are beginning to fail at alarming rates prior to the end of their design periods. Therefore, maintaining safety and operational efficiency at intersections on arterial roadways remains a constant goal. This effort for sustainability has spawned the creation and evaluation of numerous types of unconventional intersection designs. Several unconventional designs exist and have been studied, including the Bowtie, Continuous Flow Intersection, Paired Intersection, Jughandle, Median U-Turn, Single Quadrant Roadway and Superstreet Median. Typically, these designs eliminate reroute conflicting left-turn manoeuvres to and from the minor or collector road. High left-turning volumes are addressed by adding an exclusive left-turning signal. This consequently increases the required number of signal phases and shorter green time for the major through traffic. This paper describes the evaluation of an unconventional intersection designed to lessen the effects of high left-turning traffic. To aid in the evaluation of the unconventional Superstreet design, a comparison of a Conventional intersection’s operation was made. Constructing and analysing a live Superstreet and Conventional intersection design is a massive undertaking. Microscopic traffic models were developed and tested using CORSIM. A variety of scenarios were created by changing the approach volumes and turning percentages on the major/minor roads to reflect different congestion levels that may occur at the intersection on any given day. The total number of created scenarios was 72, i.e. 36 scenarios for each design. Among the general findings of this research was that the Conventional design consistently showed evidence of higher delay time and longer queue length compared to the Superstreet intersection design. The reduction in the network delay ranged from 27.39% to 82.26%, and an approximate 97.5% reduction in average network queue length experienced on the major road's through lanes when the Superstreet design was implemented. This is a significant reduction, especially since the through lane volume of the major road is relatively high. These results are assumed to be due to the additional available green time for the Superstreet intersection design.

INTRODUCTION

On a typical four-leg intersection, one of two intersecting roads services the higher traffic volume. This roadway is referred to as the major or arterial road. The second roadway, which services the lower traffic volume, is referred to as the minor or collector road. When the volume on either road nears capacity, queues begin to form, raising the potential for crashes and unsafe driving manoeuvres. For this reason improving safety and operational efficiency at intersections on arterial roadways remains a constant goal. The Federal Highway Administration (FHWA 2004) studies have shown that conventional methods of adding capacity to an intersection have diminishing results. For instance, the addition of a second through lane adds 15 years to the life of the intersection before it reaches capacity; the addition of a third through lane adds only ten years; and a fourth through lane adds only six years. Simply put, the increase in supply decreases the overall design life. Drivers attracted to the seemingly more efficient road eventually yield larger demand at a faster rate. The demand increase at large intersections can result in longer clearance intervals, more protected left-turn phasing, longer pedestrian clearance times, greater imbalances in lane utilisation, and potential queue blockage caused by the resulting longer cycle length (FHWA 2004). Combined, these factors increase loss time and potential for signal failure, and warrant the need to study and evaluate alternative methods.

In an attempt to improve the operational efficiency and safety characteristics...
of intersections, past research has explored several types of unconventional intersection designs. Several unconventional designs exist and have been studied, including the Bowtie, Continuous Flow Intersection, Continuous Green-T, Parallel Flow Intersection, Jughandle, Median U-Turn, Single Quadrant Roadway, Split Intersection, Roundabouts and Superstreet Median Crossover. These designs are referred to as “unconventional” because they incorporate geometric features or movement restrictions that would normally be allowed at standard intersections. Typically, these designs eliminate/reroute conflicting left-turn manoeuvres to and from the minor or collector cross road. High left-turning volumes are often addressed by adding an exclusive left-turning signal. Unfortunately the addition of a left-turn signal increases the required number of signal phases and shortens green time for the major through traffic, thereby increasing queue formation. Reducing the number of signal phases would improve the overall operation and safety of the intersection by enhancing capacity (with an increase in effective green) and reducing delay when the number of signal phases is reduced (Bared and Kaisar 2002: Reid and Hummer 1999).

This paper describes the evaluation of an unconventional intersection design created to decrease the effects of high left-turning traffic. To aid in the evaluation of the unconventional design named Superstreet, a comparison to a Conventional intersection’s operation was performed. Constructing a live Superstreet and Conventional intersection for evaluation reasons is a massive undertaking and not feasible in many circumstances. For this reason the two intersection designs were modelled and simulated using the microscopic traffic simulation model CORSIM (CORridor SIMulation). CORSIM is a combination of NETSIM and FRESIM. NETSIM, originally called UTCS-1, is a component of CORSIM that is capable of representing complex urban networks. Following distance, lane changing, turning movements, overtaking and driving behaviour are governed by this component of CORSIM. Many Measures of Effectiveness (MOEs) are outputted by NETSIM, including stopped delays, queue lengths, signal phase failures, fuel consumption. FRESIM is another component of CORSIM that is capable of representing complex freeway systems. CORSIM is capable of simulating freeway and surface street operations simultaneously (Papacostas & Prevedouras 2001).

<table>
<thead>
<tr>
<th>Simulation model</th>
<th>Classification</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRANSIMS</td>
<td>Large-scale microscopic</td>
<td>Modelling regions with several millions</td>
</tr>
<tr>
<td>CORSIM</td>
<td>microscopic</td>
<td>Modelling urban traffic conditions and advanced traffic control scenarios</td>
</tr>
<tr>
<td>VISSIM</td>
<td>mesoscopic</td>
<td>Modelling complex dynamic systems such as transit signal</td>
</tr>
<tr>
<td>INTRAS</td>
<td>microscopic</td>
<td>Modelling traffic conditions on freeways, ramps and highway segments</td>
</tr>
<tr>
<td>INTEGRATION</td>
<td>microscopic</td>
<td>Simulating both freeways and arterials, and evaluating ITS scenarios</td>
</tr>
<tr>
<td>MASSVAC</td>
<td>macroscopic</td>
<td>Forecasting freeway and arterial traffic, and evaluating ITS scenarios</td>
</tr>
<tr>
<td>MITSIMLab</td>
<td>microscopic</td>
<td>Model traffic operations</td>
</tr>
<tr>
<td>TransCAD</td>
<td>macroscopic</td>
<td>Conventional static model</td>
</tr>
<tr>
<td>Tranplan</td>
<td>macroscopic</td>
<td>Conventional static model</td>
</tr>
<tr>
<td>EMME/2</td>
<td>macroscopic</td>
<td>Conventional static model</td>
</tr>
<tr>
<td>Dynasmart-P</td>
<td>mesoscopic</td>
<td>Model route choice behaviour</td>
</tr>
<tr>
<td>OREMS</td>
<td>microscopic</td>
<td>Model emergency and disaster evacuation</td>
</tr>
<tr>
<td>DYNEV</td>
<td>macroscopic</td>
<td>Enhanced to model regional hurricane planning process</td>
</tr>
<tr>
<td>NETVAC</td>
<td>macroscopic</td>
<td>Evacuation model</td>
</tr>
<tr>
<td>CTM</td>
<td>macroscopic</td>
<td>Evacuation model</td>
</tr>
<tr>
<td>PARAMICS</td>
<td>microscopic</td>
<td>Provides complete visual display</td>
</tr>
<tr>
<td>CORFLO</td>
<td>macroscopic</td>
<td>Simulates design control devices</td>
</tr>
<tr>
<td>GETRAM</td>
<td>microscopic</td>
<td>Simulates traffic and human behaviour</td>
</tr>
</tbody>
</table>

**Table 1 Simulation models**

**LITERATURE REVIEW**

**Traffic simulation modelling**

Traffic micro-simulation models are widely used to qualify and evaluate the benefits and limitations of traffic operation alternatives. Boxill and Yu (2000) classify traffic simulation models as microscopic, mesoscopic and macroscopic. Models that simulate individual vehicles at small time intervals are termed microscopic, while models that aggregate traffic flow are termed as macroscopic. Mesoscopic refers to models in-between microscopic and macroscopic. The main disadvantage of microscopic simulation models is the extensive data required and the need for advanced computer resources. Microscopic simulation has been used for a long time to simulate project scale cases such as intersection design. What is new about microscopic simulation is that it is now possible to be used at a regional scale, such as simulating hurricane evacuation for a whole region with several million inhabitants (Nagel & Rickert 2000). Table 1 illustrates the most commonly used simulation models found in literature.

**Previous work**

The FHWA (2004) informational guide for signalised intersections classifies intersection...
treatments into three kinds: (1) intersection reconfiguration, (2) indirect left-turn treatments, including Jughandle, Median U-Turn, Superstreet, Continuous Flow Intersection (CFI) and Quadrant intersections, and (3) grade separation treatments. Reid and Hummer (1999) used CORSIM to compare traffic operations along an arterial road that has five signalised intersections for the Conventional Two-Way Left-Turn Lane (TWLTL) design, and two alternative unconventional designs, the Median U-Turn Crossover design and the Superstreet Median Crossover design. Results from the study indicate that the Median U-Turn and Superstreet designs improve system travel time and average speed in comparison with the TWLTL design, and overall there was a peak period travel time reduction of 17% when using the unconventional Median U-Turn and Superstreet designs.

Reid and Hummer (2001) later used CORSIM to compare the traffic performance of seven isolated unconventional intersection designs – the Quadrant, Median U-Turn, Superstreet Median Crossover, Bowtie, Jughandle, Split Intersection, and Continuous Flow Intersection. The simulation results showed that the Superstreet and Bowtie designs were only competitive with the Conventional design when the cross streets configuration was two lanes.

Also, Kim et al (2007) used VISSIM to compare the performance of the Superstreet designs to the Conventional designs. The results showed that the Superstreet design is similar to the Median U-Turn design, but has some additional features that allow for through traffic progression on the major road in both directions by preventing the minor road traffic from crossing the major road.

Description of Superstreet

The Superstreet Median Crossover design, shown in Figure 1, is an extension of the Median U-Turn design. The Federal Highway Administration (FHWA 2004) reports that “the design of a Superstreet Median Crossover is similar to that of a Median U-Turn Crossover. Crossovers should be located approximately 180 m (600 ft) from the main intersection. A semi-trailer combination design vehicle would need a median width of 18 m (60 ft) to accommodate a U-Turn”.

Drivers are not allowed to turn left from the crossroad onto the major road. The through movement for the vehicles on the minor road is accomplished by turning right onto the major road, then making a u-turn, and turning right again to the minor road (FHWA 2004). Figure 2 shows the vehicular movement at a Superstreet Median Crossover. Research has shown that forcing cross street traffic to turn right onto an arterial first, and then turning left back onto the cross street, is generally superior to a left-turn then-right pattern as seen on the Quadrant Roadside Intersection Design. These crossovers create difficult merges from the left arterial, and may only be useful when the cross street volume is small in comparison to the arterial road volume (Mahalel et al 1986).

The Superstreet configuration allows each direction of the major street to operate as two separate three-approach intersections, and allows each direction of the major street to operate on an independent timing pattern. Therefore, two two-phase traffic signals are needed at the main intersection, one for each minor street approach. In addition, two two-phase signals are required at upstream/downstream median crossover. Since the signals of the major road may be controlled and timed independently of the minor street, it is possible to achieve a maximum amount of traffic progression in both directions of the major road. The major road’s through movement benefits the most from the Superstreet Median Crossover design. Left-turning movements are permitted directly from the major street, so they also benefit from decreased delay (FHWA 2004). A typical phasing diagram of the Superstreet Median Crossover design is also shown in Figure 1.

Safety

Conflict points provide a means of comparing relative safety for vehicles between the Conventional four-leg signalised intersection and the unconventional intersection. Superstreet Median Crossover creates a total of 20 conflict points compared to 32 conflict points created by the Conventional four-leg signalised intersection. Table 2 summarises the number of conflict points in a four-leg
signalised intersection and the number of conflict points in a Superstreet intersection design. Hummer and Jagannathan (2008) investigated the safety aspects of the Superstreet by analysing sites in Maryland and North Carolina. Results showed huge reductions in collision frequencies and rates.

Superstreet, Continuous Flow Intersections (CFI), Center Turn Overpass, and Roundabouts are believed to achieve significant reductions in accident frequency, accident severity, stopped delay, and queue length (Kim et al 2007). Noted advantages of the Superstreet design in particular have been:

1. Reduces four-phase signal to two-phase signal
2. Signals for opposite direction of travel can be timed for progression independently

The simplified signal phases are very effective for the progression of through traffic and for reducing delays at the intersection, which will save the overall travel time. Although these advantages have been cited numerous times, the degree to which the Superstreet operation is more advantageous than a Conventional intersection is not well known. Furthermore, the minor street vehicle types, and the crossing and turning volumes vary by location, which makes evaluating and comparing these two intersection designs at a live and active intersection difficult.

The following sections describe the methods used, as well as the results of a comparison between a Superstreet intersection and a Conventional intersection using simulation. Delay and queue lengths were the primary measures of effectiveness used in the evaluation. Delay is arguably the most frequently experienced and troublesome aspect of travel for motorists, while the hazards associated with queue length are a constant concern for city and state traffic officials.

### METHODOLOGY

The proposed methodology for the operational evaluation and comparison between a Superstreet intersection and a traditional four-leg intersection was conducted using CORSIM platform. The two intersections used in the analysis were formed by two roadways, arterial and collector, crossing at a 90 degree angle. For simplicity, each leg of the intersection was considered to be level. The design of each leg was extended approximately 1000 feet from the centre of the intersection. Each intersection was designed in accordance with the Policy on Geometric Design of Highways and Streets (AASHTO) standards for a passenger vehicle, and a design speed of 45 mph. Lane width was considered to be 12 ft and shoulder width was considered to be 4 ft (AASHTO 2004). The design was completed using computer aided design (CAD) software, and then it was imported into traffic simulation software.

The development of the CORSIM microscopic model for the two intersection designs involved primary component steps, including the following:

1. Intersections design
2. Signalisation
3. Developing alternative scenarios
4. Analyses and comparison of all scenarios using appropriate measures.

Model calibration and validation are necessary and critical steps in any model application. However, the primary limitation to the CORSIM model development was the lack of real intersection data to support calibration and validation of the model. The following sections will discuss the details and approach to completing each step listed above.

### Table 2 Number of conflict points at a four-leg signalised intersection compared to a Superstreet Crossover (Source: FHWA 2004)

<table>
<thead>
<tr>
<th>Conflict type</th>
<th>Four-leg signalised intersection</th>
<th>Superstreet Median Crossover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merging/diverging</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Crossing (left turn)</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Crossing (angle)</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>20</td>
</tr>
</tbody>
</table>
Intersection design

The designs of the Conventional and unconventional intersections consisted of a four-lane divided major road and a three-lane undivided minor road. Only one four-phase signal was required for the Conventional CORSIM design network shown in Figure 3. However, a total of four two-phase signals had to be modelled on the Superstreet design. Two were placed at the main intersection and two more were placed at the u-turns. The Superstreet intersection was designed to match the configuration as shown in Figure 1. The final Superstreet CORSIM design model is shown in Figure 4. The circles with the embedded squares (nodes) represent the location of the signals. The additional nodes shown were used to assign turning volumes.

Signalisation

Traffic signal timing is one of the most important tasks in evaluating/comparing the two intersection designs. Since intersections are locations where traffic streams approaching from various directions converge, it is imperative that traffic signals are accurately timed to manage the traffic flow. The green time that each approach has is dependent upon many factors, with the two main factors being: cycle length and the phase plan. Intersections with large approach volumes and a high percentage of left-turning traffic usually require four phases – one phase for the north and south through traffic, one phase for the east and west through traffic, a phase for the north and south left-turning traffic, and a phase for the east and west left-turning traffic. The cycle length, which is the time taken for one approach to witness a red signal twice, is dependent upon the phase plan, volume, expected loss time for each phase, and pedestrian crossing time. Therefore, the optimum cycle length should be calculated for each case to eliminate as much loss time as possible.

Saturation flow is a key input for optimal signal timing. A small variation in saturation flow values could affect changes in cycle length, thereby affecting the efficiency and operations of an urban system. Many studies have identified suitable saturation flows at signalised intersections as being between 1 500 and 2 500 passenger cars per hour per lane (pcphgl). This variation in saturation flow is attributed to site-specific conditions (Williams & Kholslo 2006). The Highway Capacity Manual (HCM 2000) uses a base saturation flow of 1 900 pcphgl and adjusts for factors such as number of lanes, lane width, grade, lane utilisation, etc.

In this study, a saturation flow of 1 800 vehicles per hour (vph) was used, since the percentage of existing trucks was not considered to be large enough to affect the base conditions.

The desired cycle length and timing were determined using the following equation (HCM 2000):

\[ C_{des} = \frac{L}{1 + \frac{V_c}{1 800 \times PHF \times \frac{\nu}{c}}} \]  

where:
- \( C_{des} \) = Desired cycle length
- \( L \) = Total lost time
- \( V_c \) = Critical volume
- \( PHF \) = Peak hour factor
- \( \frac{\nu}{c} \) = Volume to Capacity ratio

Although appropriate for most applications, the formula mentioned above fails when the intersection critical (v/c) ratio is equal to or greater than one, and the cycle length estimate becomes unreasonably large or yields a negative number. When cases such as these arose, a cycle length of 120 seconds was used. Furthermore, 120 seconds is a common cycle length for intersections with very high approach volumes.

### Table 3 Simulation scenarios summary

<table>
<thead>
<tr>
<th>Simulation scenario</th>
<th>Arterial Roadway</th>
<th>Collector Roadway</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Traffic volume (pcph)</td>
<td>Turning movement (%)</td>
</tr>
<tr>
<td>1</td>
<td>1 800</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>2 400</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>1 800</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>2 400</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>1 800</td>
<td>20</td>
</tr>
<tr>
<td>6</td>
<td>2 400</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>1 800</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>2 400</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>1 800</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>2 400</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>1 800</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>2 400</td>
<td>20</td>
</tr>
</tbody>
</table>
Developing alternative scenarios
A total of 72 simulation scenarios (36 each per intersection) were developed with various levels of congestion at the intersection. Congestion was created using two sets of traffic volume inputs, which included 1 800 passenger cars per hour (pcph) and 2 400 pcph on the arterial roadway, and 600 pcph and 800 pcph on the collector roadway. Additionally, three different turning percentages were included in the test. As shown in Table 3, each of the two primary scenarios on each major and minor roadway was accompanied by three sets of traffic turning-movement sub-scenarios. The percentage of right-turn volume on the major roadway was fixed at 15%, with varying percentages for the left-turn volume (10%, 15% and 20%). The percentage of right-turn volume on the minor roadway was fixed at 20%, with various percentages for the left-turn volume (30%, 20% and 15%) to consider the effect of the volume for left-turn traffic. Also, it can be seen that each scenario on the major road is accompanied with three different scenarios on the minor roadway. These scenarios were used to represent varying levels of congestion and turning movements at the intersection.

The through lanes on the major road, for both the Conventional and Superstreet design, carry the highest volume. The through movement on the major road was considered to be the most critical and was therefore one of the main entities used for the analysis.

RESULTS
A total of four individual simulation runs, each using different random seed numbers, were executed for each of the 72 scenarios. This resulted in a total test set of 288 simulation runs. The additional simulation runs were also necessary to establish stochasticity within the output so that statistical testing could be carried out. The results reported in this section reflect the average of the comparative measures of effectiveness computed for each of the four replications. Once the simulations are completed, CORSIM creates reports outlining different measures of effectiveness (MOEs). The two performance measures used for the basis of comparison between the Conventional and the Superstreet intersection designs were the network average delay and the average queue length on the through lanes of the major road. These two performance measures were selected because of their direct effect on traffic operations. They also demonstrated the overall efficiency of the intersection design.

### Average delay
The average delay is a critical operational performance measure on interrupted-flow facilities, which reflects a greater discomfort caused to drivers than travel time (Zhou et al 2002). Tables 4 and 5 provide a comparison of the average delay for all scenarios using the Conventional intersection design versus those using the Superstreet intersection design, with major approach volumes of 1 800 vph and 2 400 vph respectively. The tables show the percentage difference/reduction between the Conventional and Superstreet intersection design. The analyses also show the statistical significance of the difference for the network average delay between the Conventional and the Superstreet intersection design. Statistical analyses of the data were performed using the two sample t-tests at 95% confidence level. The t-testing was used to compare relative effectiveness, by

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Conventional</th>
<th>Superstreet</th>
<th>Percent reduction</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.22</td>
<td>1.44</td>
<td>35.14%</td>
<td>0.0026</td>
</tr>
<tr>
<td>2</td>
<td>2.20</td>
<td>1.47</td>
<td>33.18%</td>
<td>0.001</td>
</tr>
<tr>
<td>3</td>
<td>2.30</td>
<td>1.67</td>
<td>27.39%</td>
<td>0.0048</td>
</tr>
<tr>
<td>4</td>
<td>10.36</td>
<td>2.24</td>
<td>78.38%</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>5</td>
<td>3.42</td>
<td>2.38</td>
<td>30.41%</td>
<td>0.0586</td>
</tr>
<tr>
<td>6</td>
<td>3.35</td>
<td>2.09</td>
<td>37.61%</td>
<td>0.0036</td>
</tr>
<tr>
<td>7</td>
<td>6.54</td>
<td>1.16</td>
<td>82.26%</td>
<td>0.0028</td>
</tr>
<tr>
<td>8</td>
<td>2.37</td>
<td>1.16</td>
<td>51.05%</td>
<td>0.0002</td>
</tr>
<tr>
<td>9</td>
<td>2.49</td>
<td>1.31</td>
<td>47.39%</td>
<td>0.0002</td>
</tr>
<tr>
<td>10</td>
<td>4.62</td>
<td>2.81</td>
<td>38.91%</td>
<td>0.0036</td>
</tr>
<tr>
<td>11</td>
<td>4.42</td>
<td>2.34</td>
<td>47.06%</td>
<td>0.001</td>
</tr>
<tr>
<td>12</td>
<td>4.65</td>
<td>2.07</td>
<td>55.48%</td>
<td>0.0004</td>
</tr>
<tr>
<td>13</td>
<td>2.66</td>
<td>1.48</td>
<td>44.36%</td>
<td>0.0016</td>
</tr>
<tr>
<td>14</td>
<td>2.70</td>
<td>1.26</td>
<td>53.33%</td>
<td>0.0002</td>
</tr>
<tr>
<td>15</td>
<td>3.00</td>
<td>2.48</td>
<td>17.33%</td>
<td>0.0942</td>
</tr>
<tr>
<td>16</td>
<td>5.20</td>
<td>2.69</td>
<td>48.27%</td>
<td>0.0012</td>
</tr>
<tr>
<td>17</td>
<td>4.89</td>
<td>2.56</td>
<td>47.65%</td>
<td>0.0008</td>
</tr>
<tr>
<td>18</td>
<td>4.81</td>
<td>2.21</td>
<td>54.05%</td>
<td>0.0006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Conventional</th>
<th>Superstreet</th>
<th>Percent reduction</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.73</td>
<td>4.28</td>
<td>44.63%</td>
<td>0.0064</td>
</tr>
<tr>
<td>2</td>
<td>7.00</td>
<td>4.02</td>
<td>42.57%</td>
<td>0.0026</td>
</tr>
<tr>
<td>3</td>
<td>6.52</td>
<td>4.23</td>
<td>35.12%</td>
<td>0.0578</td>
</tr>
<tr>
<td>4</td>
<td>11.34</td>
<td>3.87</td>
<td>65.87%</td>
<td>0.0236</td>
</tr>
<tr>
<td>5</td>
<td>8.27</td>
<td>3.66</td>
<td>55.74%</td>
<td>0.0016</td>
</tr>
<tr>
<td>6</td>
<td>8.41</td>
<td>3.98</td>
<td>52.68%</td>
<td>0.003</td>
</tr>
<tr>
<td>7</td>
<td>9.50</td>
<td>4.49</td>
<td>52.74%</td>
<td>0.0269</td>
</tr>
<tr>
<td>8</td>
<td>6.91</td>
<td>4.67</td>
<td>32.42%</td>
<td>0.0528</td>
</tr>
<tr>
<td>9</td>
<td>7.09</td>
<td>3.22</td>
<td>54.58%</td>
<td>0.0006</td>
</tr>
<tr>
<td>10</td>
<td>8.74</td>
<td>4.12</td>
<td>52.86%</td>
<td>0.0052</td>
</tr>
<tr>
<td>11</td>
<td>8.75</td>
<td>3.91</td>
<td>55.31%</td>
<td>0.004</td>
</tr>
<tr>
<td>12</td>
<td>8.76</td>
<td>3.87</td>
<td>55.82%</td>
<td>0.0036</td>
</tr>
<tr>
<td>13</td>
<td>7.24</td>
<td>3.05</td>
<td>57.87%</td>
<td>0.0012</td>
</tr>
<tr>
<td>14</td>
<td>7.28</td>
<td>2.86</td>
<td>60.71%</td>
<td>0.0004</td>
</tr>
<tr>
<td>15</td>
<td>7.37</td>
<td>3.48</td>
<td>52.78%</td>
<td>0.0014</td>
</tr>
<tr>
<td>16</td>
<td>8.86</td>
<td>5.05</td>
<td>43.00%</td>
<td>0.0018</td>
</tr>
<tr>
<td>17</td>
<td>8.86</td>
<td>3.91</td>
<td>55.87%</td>
<td>0.0032</td>
</tr>
<tr>
<td>18</td>
<td>8.92</td>
<td>4.01</td>
<td>55.04%</td>
<td>0.0044</td>
</tr>
</tbody>
</table>
determining if the network average delay on the Superstreet intersection design was shorter than the Conventional intersection design, and the statistical significant difference between the two intersections. The following null and alternative hypotheses were used:
- \( H_0 \): the network average delay on the Conventional and Superstreet intersection design is equal.
- \( H_1 \): the network average delay on the Conventional and Superstreet intersection design differs.

### Table 6 Conventional and Superstreet queue length comparisons (1 800 vph)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average queue length (m)</th>
<th>Percent reduction</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.13</td>
<td>2.63</td>
<td>78.32%</td>
</tr>
<tr>
<td>2</td>
<td>12.00</td>
<td>2.75</td>
<td>77.08%</td>
</tr>
<tr>
<td>3</td>
<td>4.50</td>
<td>2.63</td>
<td>41.67%</td>
</tr>
<tr>
<td>4</td>
<td>75.63</td>
<td>6.25</td>
<td>91.74%</td>
</tr>
<tr>
<td>5</td>
<td>7.00</td>
<td>5.63</td>
<td>19.57%</td>
</tr>
<tr>
<td>6</td>
<td>5.75</td>
<td>5.75</td>
<td>0.00%</td>
</tr>
<tr>
<td>7</td>
<td>38.88</td>
<td>1.00</td>
<td>97.43%</td>
</tr>
<tr>
<td>8</td>
<td>3.75</td>
<td>1.00</td>
<td>73.33%</td>
</tr>
<tr>
<td>9</td>
<td>5.00</td>
<td>0.75</td>
<td>85.00%</td>
</tr>
<tr>
<td>10</td>
<td>18.88</td>
<td>8.88</td>
<td>52.98%</td>
</tr>
<tr>
<td>11</td>
<td>16.00</td>
<td>6.50</td>
<td>59.38%</td>
</tr>
<tr>
<td>12</td>
<td>19.38</td>
<td>5.38</td>
<td>72.26%</td>
</tr>
<tr>
<td>13</td>
<td>5.25</td>
<td>1.63</td>
<td>69.05%</td>
</tr>
<tr>
<td>14</td>
<td>5.63</td>
<td>1.00</td>
<td>82.22%</td>
</tr>
<tr>
<td>15</td>
<td>7.13</td>
<td>5.25</td>
<td>26.36%</td>
</tr>
<tr>
<td>16</td>
<td>24.88</td>
<td>7.63</td>
<td>69.35%</td>
</tr>
<tr>
<td>17</td>
<td>21.25</td>
<td>7.25</td>
<td>65.88%</td>
</tr>
<tr>
<td>18</td>
<td>18.88</td>
<td>5.88</td>
<td>68.87%</td>
</tr>
</tbody>
</table>

### Table 7 Conventional and Superstreet queue length comparisons (2 400 vph)

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Average queue length (m)</th>
<th>Percent reduction</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56.625</td>
<td>13</td>
<td>77.04%</td>
</tr>
<tr>
<td>2</td>
<td>55.75</td>
<td>12.5</td>
<td>77.58%</td>
</tr>
<tr>
<td>3</td>
<td>56.375</td>
<td>12.875</td>
<td>77.16%</td>
</tr>
<tr>
<td>4</td>
<td>87.375</td>
<td>12.25</td>
<td>85.98%</td>
</tr>
<tr>
<td>5</td>
<td>63.125</td>
<td>12.125</td>
<td>80.79%</td>
</tr>
<tr>
<td>6</td>
<td>64.5</td>
<td>13.25</td>
<td>79.46%</td>
</tr>
<tr>
<td>7</td>
<td>77.25</td>
<td>16</td>
<td>79.29%</td>
</tr>
<tr>
<td>8</td>
<td>58.125</td>
<td>12</td>
<td>79.35%</td>
</tr>
<tr>
<td>9</td>
<td>59.75</td>
<td>19.25</td>
<td>67.78%</td>
</tr>
<tr>
<td>10</td>
<td>67.25</td>
<td>13.625</td>
<td>79.74%</td>
</tr>
<tr>
<td>11</td>
<td>67.25</td>
<td>13.5</td>
<td>79.93%</td>
</tr>
<tr>
<td>12</td>
<td>66.875</td>
<td>13</td>
<td>80.56%</td>
</tr>
<tr>
<td>13</td>
<td>59.375</td>
<td>18.5</td>
<td>68.84%</td>
</tr>
<tr>
<td>14</td>
<td>60.5</td>
<td>17.75</td>
<td>70.66%</td>
</tr>
<tr>
<td>15</td>
<td>59.5</td>
<td>17.5</td>
<td>70.59%</td>
</tr>
<tr>
<td>16</td>
<td>67.5</td>
<td>14.875</td>
<td>77.96%</td>
</tr>
<tr>
<td>17</td>
<td>67.125</td>
<td>12.625</td>
<td>81.19%</td>
</tr>
<tr>
<td>18</td>
<td>60.5</td>
<td>13</td>
<td>78.51%</td>
</tr>
</tbody>
</table>

**Queue length**

In the research, queue length was also used as a performance measure of effectiveness for the operational evaluation and comparison of the two intersection designs. Once again, a two-sample t-test was performed at 95% confidence level to determine statistically significant difference in the average queue length between the Conventional and Superstreet intersection designs. The following null and alternative hypotheses were used:
- \( H_0 \): the average queue length on the Conventional and Superstreet intersection design is equal.
- \( H_1 \): the average queue length on the Superstreet intersection design is greater than the Conventional intersection design.

Tables 6 and 7 provide a comparison of the average queue length for all scenarios using the Conventional intersection design versus those using the Superstreet intersection design, with major approach volumes of 1 800 vph and 2 400 vph respectively. The numbers in italics in the shaded rightmost columns of Tables 6 and 7 show that a significant difference existed between the two intersection designs. The tables show that the average queue length for all scenarios was significantly better using the Superstreet intersection design compared to the Conventional intersection design, with a percentage reduction up to 97.43%. Again, these results are thought to be due to the additional available green time for the Superstreet intersection design.

**CONCLUSION**

The aim of this paper was to evaluate and compare the operational efficiency of a Conventional signalised intersection with an unconventional Superstreet Median Crossover intersection using micro-simulation software. For this purpose two CORSIM models depicting the Superstreet Median Crossover and a Conventional intersection were developed and tested. Several scenarios were created by changing the approach volumes and turning percentages on the major/minor roads to reflect different congestion levels at the intersection, resulting in a total of 72 scenarios, i.e. 36 for each model. The optimal signal timing was calculated for each case to eliminate biasness. Each scenario had its own independent output, and the most pertinent variables were extracted from the output and used in the analysis. The variables considered to be of
primary importance were queue length and average network delay, since these measures directly affect traffic operation on major roads’ through lanes.

Among the general findings of this research was the fact that the Superstreet intersection design consistently showed evidence of decreased delay time and queue length when compared to the Conventional design. The percentage reduction in the network delay ranged from 27.39% to 82.26%, and an approximate 97.50% reduction in average network queue length experienced on the through lanes of the major road. Such a large divide is best explained by the signalising methods along the major roads. Generally speaking, increasing green time has positive effects on network delay and queue length. The Superstreet operates on a synchronised two-signal phase, which allows more green time to be allocated to the major roads’ through volume, and consequently decreases the chance for queues to form and signal failure to occur. When the major road’s through lane volume is relatively high and receives green time priority, this effect is more recognisable.

A more detailed investigation of the comparison data showed that the greatest delay and queue length differences occurred when a high percentage of minor road left-turners (approximately 30%) coincided with a moderate amount of major road left-turners (above 15%). This implies that restricting left turns from minor street approaches could result in operational benefits for intersections in most cases.

Although Superstreet design has been suggested to decrease the overall delay at an intersection, an increased delay is experienced for motorist desiring to travel through and turn left from a minor street approach. For this reason, and because of the design of a Superstreet requires extra right of way, careful consideration concerning the minor street traffic composition and nearby stakeholders should be taken. It is suggested that the Superstreet should be considered only where there is adequate right-of-way and where high arterial through volumes conflict with moderate to low cross-street through volumes.

While data presented in this paper provides evidence that the Superstreet is well suited for major street operations, more research is needed to examine the effects of directional volume, lane volume, driver adaptability and expectations in different geometric designs.

ACKNOWLEDGEMENT

The authors of this paper wish to acknowledge the technical assistance provided by Brian Wolshon of Louisiana State University for providing expert information critical to the development of the model and the analysis of the data.

REFERENCES


COMMENT

On reducing the carbon footprint of the concrete industry

It is suggested that a small, but probably significant reduction in — or actually a more accurate assessment of — the carbon footprint of the industry can be obtained by considering the amount of CO₂ sequestered during carbonation of concrete and cement-stabilised pavement materials.

Both the cement minerals and the free lime released during the hydration of Portland-type cements are subject to carbonation in engineering time.

Although the carbonation of concrete in atmospheric air is slow (about 0.1–3 mm/year), far more cement is used in concrete than in stabilisation. However, that of cement-stabilised soil pavement layers is much faster (about 0.5–2 mm/day on all surfaces exposed to atmospheric air) and about 2–50 mm/year from the bottom of the layer upwards due to reaction with soil air (Netterberg 1991). Although these rates tend to decrease with time, this means that a thickness of 150–300 mm of 2–3% cement-stabilised material can thus become completely carbonated within 5–10 years. In both cases the amount of CO₂ taken up will be approximately equal to the amount of CO₂ released by the now carbonated cement during its manufacture.

The same applies to lime-stabilised pavement layers.

This will enable a “cradle-to-grave” estimate rather than just a “cradle-to-gate” estimate to be made.

Reference


Dr Frank Netterberg
fnetterberg@absamail.co.za

RESPONSE FROM AUTHORS

The authors are in agreement that the sequestration of CO₂ during carbonation of concrete serves in lowering the carbon footprint of the industry.

However, for the amount of CO₂ sequestered to be approximated, there first needs to be a more accurate assessment than that done in the published paper of the end applications of the cementitious materials. The proposed study would give details of the stock of concrete structures of all types in South Africa, including stabilised layers in road construction, and their actual service lives. Based on the findings, one can then approximate the cradle-to-grave environmental impacts of concrete structures in SA.

Regarding road-stabilised layers in particular, there is the question of access of CO₂ to the layer, due to (a) overlying wearing courses, and (b) degree of saturation of the layer.

Rachel Muigai
murea001@myuct.ac.za

Prof Mark Alexander
mark.alexander@uct.ac.za

Prof Pilate Moyo
pilate.moyo@uct.ac.za
Publishing particulars of paper under discussion
Vol 55 (3) 2013, Pages 96–102, Paper 841: Structurally efficient housing incorporating natural forms

M Gohnert, A Fitchett, I Bulovic, N Bhikhoo

COMMENT
The above-mentioned paper deals with the catenary shape to structural forms to obviate tensile and bending stresses.

The authors make the following statement on page 98: "Unfortunately, the practice of implementing natural forms seems to have faded in the early 1900s – an incomprehensible digression in structural design."

This statement is incorrect!
In the 1950s Mr R I D M Myburgh, then in charge of the Design Division of the Department of Water Affairs, designed multiple-dome dams on the same principles. He loaded a steel-wire mesh model of the structure with weights representing the forces on the dam – own weight and water load. The mesh model took on a shape, which, when reversed, would be wholly in compression. This shape was measured and expressed as quadratic equations of a series of horizontal ellipses, which were then used in setting out the structure. A vertical section through the centre of the arch is parabolic. Three dams of those shapes were successfully completed – the Wagendrift Dam on the Bushmans River, the Stompdrift Dam on the Olifants River near Oudtshoorn and the Kat River Dam on the river of that name in the Eastern Cape. I was privileged to have been on the construction of the Stompdrift Dam, consisting of three domes, supported by buttresses. These concrete domes are extremely thin, illustrating the economy of materials! All three dams have now successfully been in service for about fifty years!

Dr Theo van Robbroeck Pr Eng
tpvcn@iafrica.com
Thanks to our Referees!

The SAICE Journal Editorial Panel would like to thank the persons listed below, all of whom served as referees during 2013. The quality of our journal is not only a reflection of the level of expertise of participating authors, but certainly also of the high standard set by our referees.

Dr Dotun Adegoke  Mr Anton Holtzhausen  Dr Martin Rust
Prof Mark Alexander  Dr Emile Horak  Prof Pantaleo Rwelamila
Dr Eduardo Alonso  Prof Adesola Ilemobade  Dr Manus Santhanam
Dr Joseph Anochie-Boateng  Prof Schalk Jacobsz  Dr Koos Schoonees
Mr Phumudzo Baloyi  Dr Paul Jagals  Ms Sarah Skorpen
Dr Hans Beushausen  Mrs Anna Jansen van Vuuren  Prof John Smallwood
Mr John Bradley  Prof Kim Jenkins  Dr Gerrit Smit
Mr Chris Brooker  Mr Werner Jerling  Francis Stoffberg
Vanja Bulovic  Mr Tomme Katranas  Dr Breda Strassheim
Dr Artur Cichowicz  Prof Elsabé Kearsley  Prof Derek Stretch
Mr Louis Coetzee  Mr Johan Kotzé  Mr Gianfranco Talocchino
Prof Johan de Beer  Mr Maxim Kvtun  Dr Hechter Theyse
Dr Morris de Beer  Dr Geoff Krige  Mr Tony Uckermann
Mr Gerard de Swart  Mr Adriaan Kurtz  Dr Christo van As
Dr Erik Denneman  Dr Ruinian Li  Emeritus Prof Martin van de Ven
Dr Dirk Du Plessis  Mr Ione Loots  Mr Etienne van der Klashorst
Dr Raëlize du Plooy  Prof Mahen Mahendran  Mr Rein van der Merwe
Prof Raymond Durrheim  Dr Philippe Maincon  Mr Jackie van der Westhuizen
Mr Heinrich Elges  Mr Toit Malan  Dr André van Nierkerk
Prof Alex Elvin  Mr Hannes Marais  Dr Gert van Rooyen
Dr Johann Enslin  Mr Avin Masarira  Prof Fanie van Vuuren
Prof Cornelius Fourie  Dr Seirgei Miller  Prof Gideon van Zijl
Ms Pauline Froshauer  Dr Malcolm Mitchell  Dr Senthilkumar Venkatachalam
Prof Leroy Gardner  Prof David Muir Wood  Prof Christo Venter
Mr Amanauel Gebremeskel  Estimé Mukandila  Mr Harry Viljoen
Dr Adam Goliger  Mr Paul Naude  Emeritus Prof Alex Visser
Prof Hannes Gräbe  Dr Frank Netterberg  Dr Eduard Vorster
Shaahid Hansa  Dr Gerhard Offringa  Dr Fritz Wagener
Dr Johannes Haarhoff  Dr John Onraet  Mr Martin William
Dr Trevor Haas  Mr Alan Parrock  Mr Dave Wright
Dr Theuns Henning  Dr Phil Paige-Green  Mr Henry-John Wright
Dr Francois Heyns  Dr Gianluca Ranzi
Dr Don Hodgkinson  Prof Chris Roth
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 the 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 may be confidential or sensitive. Authors are also responsible for obtaining permission from copyright owners when reproducing material that has been published elsewhere.

SUBMISSION PROCEDURES AND REQUIRED FORMAT

- Hard copies: Hard copies should only be submitted if the author is unable to submit electronically, for whatever reason. In such a case only one unstapled copy, together with an electronic copy on CD, should be posted to the Chief Executive Officer of SAICE, Private Bag X200, Halfway House, 1685, or delivered to Building 19, Thornhill Office Park, Bekker Street, Midrand. Hard copies should be printed on one side of A4 paper.
- Electronic submission: Manuscripts should be e-mailed to the editor at verellen@saice.org.za. File sizes should preferably not exceed 3 MB per e-mail – figures may for example be sent one by one or in groups not larger than 3 MB. Manuscripts should not be sent in PDF format as this precludes review of papers per track changes.
- Format: Manuscripts should be prepared in MS Word and presented in double line spacing, single column layout with 25 mm wide margins. All pages should bear the authors’ names and be numbered at the bottom of the page. With the exception of tables and figures (see below) the document should be typed in Times New Roman 12 pt font. Contributions should be accompanied by an abstract of not more than 200 words and a profile not exceeding 100 words on each of the authors.
- First page: The first page of the manuscript should include the title of the paper, the initials and surnames of the authors, professional status (if applicable), SAICE affiliation (Member, Fellow, Visitor, etc), telephone and fax numbers, and e-mail and postal addresses. The name of the corresponding author should be underlined. Five keywords should be suggested. The copy must not contain instructions to the editor.
- Figures, tables, photos and illustrations: These should preferably be submitted in colour, as the journal is a full-colour publication. Figures, tables, photos and illustrations, complete with captions, should be placed on separate pages at the end of the text. Their positions should be clearly marked in the text as follows: ‘Insert figure 1’.
- Captions for tables should appear above the table. All other captions should appear below the illustration (figures, graphs, photos).
- Only those figures and photographs essential to the understanding of the text should be included. All illustrations should be referred to in the text. Original prints of photographs are required. Alternatively photos should be scanned at 300 dpi minimum and submitted as separate files in JPEG (jpg) format. Care should be taken not to damage photos by writing on the back or using paper clips. Illustrations will not be returned after publication, unless arrangements are made.
- The final paper will be typeset in a three column format. Figures and tables should be 52 mm, 115 mm or 176 mm wide – to fit across one, two or three columns respectively. The maximum height of any figure or table should not exceed 250 mm including the caption.
- Tables should be typed in Times New Roman 9 pt font. They should not duplicate information already given in the text, nor contain material that would be better presented graphically. Tabular matter should be as simple as possible, with brief column headings and a minimum number of columns.
- 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

- Copyright: On acceptance of the paper or technical note, copyright must be transferred by the author 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.
- Proofs: First proofs of papers will be sent to authors in PDF format for verification before publication.

Guidelines for the preparation of papers and technical notes

Authors should comply with the following guidelines when preparing papers for publication in the journal.