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

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

2  Conduit-hydropower potential in the City of Tshwane water distribution system: A discussion of potential applications, financial and other benefits
I Loots, M van Dijk, S J van Vuuren, J N Bhagwan, A Kurtz

14  Evaluating the prioritisation of South African dams for rehabilitation with special focus on risk to human lives
S Reynolds, C Barnardo-Viljoen

25  Effects of safety factors on the deflections in a concrete gravity dam
M Opan

30  Soil collapse from an effective stress perspective
G Brink, G Heymann

34  Parametric study on the behaviour of Y-shaped composite bridges
P Lu, C Shao

44  Effect of activator dosage, water-to-binder-solids ratio, temperature and duration of elevated temperature curing on the compressive strength of alkali-activated fly ash cement pastes
J Shekhovtsova, E P Kearsley, M Kovtun

53  Directional wave spectra on the east coast of South Africa
S Corbella, D D Stretch

65  Review of pump suction reducer selection: Eccentric or concentric reducers
R M Mahaffey, S J van Vuuren

77  Dynamic behaviour of underspanned suspension road bridges under traffic loads
J Oliva, J M Goicoechea, P Antolin, M Astiz
Conduit-hydropower potential in the City of Tshwane water distribution system: A discussion of potential applications, financial and other benefits

I Loots, M van Dijk, S J van Vuuren, J N Bhagwan, A Kurtz

In water distribution networks, water is often fed under gravity from a higher reservoir to another reservoir at a lower level. The residual pressure head at the receiving reservoir is then dissipated through control valves (mechanically or hydraulically actuated), sometimes augmented by orifice plates where there is a propensity for cavitation. There are possibilities to add turbines in parallel and generate hydroelectricity at these locations using the flow and head available.

The benefit of this hydropower generating application is that minimal civil works need to be done, as the control valves are normally inside a control room/valve chamber. No negative environmental or social effects require mitigation, and the anticipated lead times should be short. From a topographical perspective the City of Tshwane has a lower elevation than the bulk service reservoirs of Rand Water, which is the main water supply. Water is distributed through a large water system that includes 160 reservoirs, 42 water towers, 10 677 km of pipes and more than 260 pressure reducing stations (PRS) that operate at pressures of up to 250 m. The top ten hydropower potential sites in the City of Tshwane water distribution network have a total energy generating capacity of approximately 10 000 MWh/a. A number of potential conduit-hydropower sites have shown promise of short payback periods. The identifying and development of these sites in Tshwane to convert water pressure to electricity is ongoing and expedited further.

Various challenges currently exist with reservoir communication in isolated areas due to vandalism and theft of necessary infrastructure, including electricity cables and solar panels. Because conduit-hydropower systems can be housed completely inside chambers, vandalism and theft can be mitigated. Therefore, one of the major benefits of hydropower turbines at these sites is that the hydropotential could be exploited to power telemetry, pressure management, flow control and monitoring/security systems.

Short payback periods, especially when using pumps as turbines, also make conduit-hydropower systems attractive.

INTRODUCTION

Energy is the lifeblood of worldwide economic and social development. The current status of global energy shortages and the emphasis to reduce CO2 emissions stimulate the development of alternative electricity generation methods at all levels of the South African economy. The demand for energy is increasing continuously, primarily due to changing lifestyles and the increase in population. These demands need to be met in order to stimulate worldwide development. They can be satisfied by developing alternative, particularly renewable, energy resources using well-researched technologies. Renewable energy technologies will have to be exploited to effectively support future economic development and satisfy energy demands. Among targeted renewable energy sources available for energy generation in South Africa are solar radiation, biomass, wind and also (rather underdeveloped) hydropower (DoE 2011).

Energy efficiency, optimisation of existing systems and seeking new approaches in conversion of one energy form into another are also spheres of electricity generation where
hydropower generation, with only 6% of the estimated potential exploited. This is not a burden, but an opportunity. Although South Africa has below-average conventional hydropower potential, large quantities of raw and potable water are conveyed daily under either pumped or gravity conditions over large distances and high elevations. The water is supplied typically to residential, industrial and irrigation areas, commonly requiring high security of supply. These water transport systems have to be operated under sustainable water supply regimes, which is a very important aspect in the operation of any hydropower generation system.

There are basically four areas where electricity generation can occur in the water supply and distribution system (WDS), as shown in Figure 1 (adapted from Briggeman 2011).

1. Dam releases – conventional hydropower.
2. At water treatment works (raw water) – the bulk pipeline from the water source can be tapped.
3. Potable water – at inlets to service reservoirs where pressure reducing stations (PRS) are utilised to dissipate the excess energy.
4. Distribution network – in the distribution network itself where residual energy is dissipated (typically with pressure reducing valves (PRV)).

The University of Pretoria (UP), supported by the Water Research Commission (WRC), is engaged in a research project to investigate the potential of extracting the available energy from existing and newly installed water supply and distribution systems. The project aims to enable the owners and administrators of the bulk water supply and distribution systems to install small-scale hydropower systems to generate hydroelectricity for on-site use and, in some cases, to supply energy to isolated electricity demand clusters or even to the national electricity grid, depending on the location, type and size of installation.

An initial scoping investigation highlighted the untapped hydropower-generating potential from pressurised conduits, specifically in the City of Tshwane WDS.

NEED FOR RENEWABLE ENERGY DEVELOPMENT IN SOUTH AFRICA

The awareness of a need for renewable energy development in South Africa was boosted significantly in November 2003 when the South African government introduced the White Paper on Renewable Energy (WP on RE). This document set a 2013 target of 10 000 GWh to be generated annually from renewable sources. Among targeted renewable energy sources available for energy generation in South Africa are solar radiation, biomass, wind and also, rather underrated, hydropower (DoE 2011).

South Africa, as one of the signatories of the Kyoto Protocol (February 2005), committed itself to reducing emissions by 34% below projected emissions level in 2020. The emissions level from all sources in South Africa is currently estimated at about 500 000 000 tons of carbon dioxide equivalent (CO2e) per annum. Accordingly, South Africa has committed itself to an emissions trajectory that peaks at 34% below a “Business as Usual” trajectory in 2020 and 40% in 2025, remaining stable for around a decade, and declining thereafter in absolute terms. To provide a suitable enabling environment for emissions reduction and reliable energy supply for the South African economy, the Department of Energy (DoE), with endorsement from the National Energy Regulator of SA (NERSA), introduced the Integrated Resource Plan (IRP) for electricity in South Africa 2010–2030. The IRP 2010 had been subjected to public scrutiny and comments, and eventually the whole process manifested in the Final Policy Adjusted IRP 2010: New-build Technology Mix. The DoE subsequently allocated different capacities across various renewable energy technologies from the total development capacity of 3 725 MW. The hydropower sector has been allocated overall capacity of 75 MW to be commercially operational by June 2014. One of the critical qualification requirements is that only small-scale hydropower installations above 1 MW are to be included in the
forthcoming selection process. Effectively all pico (up to 20 kW as shown in Table 1), micro (20 kW to 100 kW) and mini (100 kW to 1 MW) renewable energy installations are below the level of interest of the authorities at this stage.

Internationally, small hydro is considered to be the best proven of all renewable energy technologies, ideal for the electrification of remote communities, assisting in peak supply, and can be used to balance out variations present in wind and solar power production. Both wind and solar technologies require energy storage facilities, typically provided by hydraulic infrastructure (e.g. dams, reservoirs, pipelines, canals, etc).

Hydroelectricity generation from small-scale installations is gaining unprecedented world-wide interest, mainly due to its social, environmental and financial benefits, particularly if hydropower technology is added to existing infrastructure.

Various challenges currently exist with reservoir communication in isolated areas due to vandalism and theft of necessary infrastructure, including electricity cables and solar panels. Because conduit-hydropower systems can be housed completely inside chambers, vandalism and theft can be mitigated. Therefore, one of the major benefits of hydropower turbines at these sites is that the hydroelectric potential could be exploited to power telemetry, pressure management, flow control and monitoring/security systems.

Alternatively, or additionally, other local demand and/or (depending upon the quantum of energy available) off-site energy demand clusters, or even a municipal or national grid could also be serviced by these power stations. The capacity of hydroelectric installations can vary to suit the application for the amount of power to be generated or needed.

### CONDUIT-HYDROELECTRIC INSTALLATIONS

The turbine/generator set is typically installed just upstream of the inlet pipe to the service reservoir or could be placed inline. Water is discharged into the service reservoir under atmospheric pressure. There are a few technical issues to be borne in mind when developing conduit-hydropower:

1. The service reservoir operates as a tailrace.
2. The water inflows into the reservoir should equal the outflows.
3. The head fluctuation within the service reservoir and the system head losses dictate the operating head of the turbine installation.
4. The flow available for hydroelectricity generation is dependent on the water demand, which in turn is subject to community water use patterns, and seasonal variations.
5. The base demand determines likely flow available to the turbine installation.
6. There are transient pressures which could be developed, typically caused by load rejection (i.e. danger of damage if water hammer exceeds design conditions).
7. A turbine installation will not be feasible if the water pipeline is not structurally sound (e.g. age, type of material, etc).
8. A turbine by-pass piping system might need to be installed to allow for excess water flows to be diverted directly into the reservoir.
9. Operational optimisation of series-connected systems may prove difficult.
10. Reliability of supply should not be compromised.
11. Further upgrading of a pipeline system could be offset against potential income from the generated hydropower.

In South Africa there are several inter-basin water transfer schemes (WTS) that can be considered for hydroelectric development. The systems identified to date are mainly under corporate administration of Eskom and the Department of Water Affairs (DWA) and include: Assegaaai to Vaal (KwaZulu-Natal to Mpumalanga), Vaal River Eastern Sub-system Augmentation Project (VRESAP), and the Orange-Fish Tunnel WTS. Eskom is currently conducting feasibility studies to install hydropower capacity in the latter.

Several water supply utilities (former water boards) and metropolitan municipal water supply systems, with configurations comprising a gravity pipeline connecting two reservoirs, are suitable for in-line hydroelectric installation. The turbine/generator set can be installed on the delivery or by-pass and the excess pressure can be exploited for hydroelectricity generation.

Current conduit-hydropower projects under way include:
- **Rand Water** – the utility determined that among its 58 service reservoirs there is a firm hydroelectric potential of 15 MW. It has subsequently been estimated that a further 50 MW of capacity is hidden within the utility’s water supply and distribution systems. A tender recently closed where this type of energy generation is planned at four locations, totalling almost 13 MW.
- **eThekwini’s Water and Sanitation Department** – the department is considering installation of six mini hydropower sets within the eThekwini potable water system. These proposed reservoir sites are situated at Sea Cow Lake, Kwa Mashu 2, Aloes, Phoenix 1, Phoenix 2 and Umhlanga 2 Reservoirs.
- **Bloem Water** – the water utility is considering micro hydropower installations at the Uitkijk and Brandkop Reservoirs on the Caledon–Bloemfontein pipeline. These two sites could generate approximately 400 kW each. Currently a 96 kW installation at Brandkop is being developed.
- **Lepelle Northern Water** – various hydropower options have been investigated and it was found that the supply conduit to the purification works at Ebenezer Dam is a viable economic development option, with potential in excess of 100 kW.
- **City of Tshwane** – various sites are under investigation and will be described in more detail below.

### CITY OF TSHWANE WATER DISTRIBUTION SYSTEM DESCRIPTION

The City of Tshwane (now including Metsweding) receives bulk water from Rand Water, Magalies Water and from own sources including boresholes, water purification plants and springs. Water is then distributed as shown in Figure 2, through a large water system that includes 160 reservoirs, 42 water towers, 10 677 km of pipes and more than 10 000 service connections. Approximately 1 000 000 people are served daily.

<table>
<thead>
<tr>
<th>Table 1 Hydroelectric capacity applications from small-scale categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydropower category</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Pico</td>
</tr>
<tr>
<td>Micro</td>
</tr>
<tr>
<td>Mini</td>
</tr>
<tr>
<td>Small</td>
</tr>
</tbody>
</table>

NB: All installations above 10 MW are classified as macro (or large) hydropower plants.
260 pressure reducing stations (PRSs) that operate at pressures up to 250 m.

The investigation into the development of potential hydroelectric sites in the Tshwane WDS started in 2008 when the first low-cost pilot plant was constructed and tested at the Queenswood Reservoir (Van Vuuren 2010). This was followed up with a 15 kW installation at the Pierre van Ryneveld Reservoir that was completed in October 2011.

IDENTIFYING HYDROPOWER POTENTIAL IN A WDS

A decision support system (DSS) that can be used to identify conduit-hydropower potential in South Africa, developed as part of the WRC research project, provides guidance for the development of identified potential sites (Loots 2013). A system of flow diagrams and tools has been compiled to identify and develop conduit-hydropower sites.

A systematic approach, consisting of the following three phases, is followed when assessing hydropower potential in a distribution network to ensure that all relevant factors are considered:

- **First Phase (Pre-feasibility Investigation):** The only input required in this section is the average daily flow, the average pressure head, if available, the static energy head (if the average head is not known) and, if applicable, the distance to the grid connection and power demand. The output in this section includes the theoretically available power and the ratio of the energy demand vs available energy, in the case of on-site or islanded systems. The Economic Analysis Section does not require any input, except the design life of the project, unless better information than the default values is available. The output from this section includes initial estimates of the net present value (NPV), internal rate of return (IRR) and payback period of the proposed project.

- **Second Phase (Feasibility Study):** The input at this stage becomes more detailed, with measured flow and pressure records required. Some of the output in Phase 2 includes an optimum design flow and head, initial turbine selection, flow rating curve and economic analysis based on the turbine selected. Environmental, social and regulatory assessments are also conducted during this phase.

- **Third Phase (Detailed Design):** The input and output of Phase 3 are to some extent similar to that of Phase 2, but with additional detail input required. Specifically, a complete flow and head data set, all costs and income expected in the life cycle of the project, and criteria for when the system should be functional, are needed. This phase also requires detail design of all civil and electro-mechanical components and infrastructure. Each phase has its own process flow diagram linked to the Conduit-Hydropower Potential Tool (CHD Tool). Some of the aspects of the study will occur in two or more of the phases, but are dealt with in increasing detail as the project progresses. A fourth phase, dealing with operation and maintenance aspects, falls outside the scope of this system, but is also an important phase to consider when designing a conduit-hydropower facility.

The first phase of the DSS was utilised in the identification and analyses of the viability of developing the sites.

HYDROPOWER POTENTIAL IN THE CITY OF TSHWANE WDS

As a first step a desktop study was conducted where the ten larger reservoir sites in the City of Tshwane were identified (Van Vuuren 2010). The use of the potential energy stored in the pressurised closed conduit water systems in Tshwane is, however, not limited to the 10 larger sites as listed in Table 2. The scope of using all available pressurised water systems in Tshwane to convert potential water energy to electricity is still to be investigated and exploited further.

In the Tshwane water supply area (TWSA), there are a number of reservoirs receiving water from Rand Water at a pressure of up to 250 m. The initial conservative assumptions which were used to calculate the potential annual hydropower generation from these pressurised supply pipelines were:

- The fraction of the available static head that can be used to generate power is 0.5.
- The hours per day when power can be generated are only six hours

Based on the above assumptions, the potential annual hydropower generation at reservoirs in the TWSA was calculated. This analysis is a conservative low estimate of the hydropower generating capacity. In the case of the power generation from reservoirs in the TWSA, the fraction which has been used to calculate the hydropower generation is only 12.5% (0.5*6/24*100) of the potential maximum power generation.

Figure 3 indicates the potential hydropower generation capacity at different reservoirs in the TWSA. This analysis was based on utilising the available data in the IMQS information system.

The capacity of hydroelectric installations can vary to suit the application for the amount of electricity to be generated or need ed. An example may be the necessity to have communication with reservoirs in isolated areas due to various operation, maintenance and infrastructure management reasons. It is not practical to have personnel driving hundreds of kilometres at high costs to inspect installations and monitor water levels in these isolated areas on a daily basis, while the potential for hydro energy is available. To supply electricity only a relatively small power source for these reservoirs and PRV installations in isolated areas is required to power telemetry, pressure management, flow control and 24-hour monitoring/security systems.

The use of hydropower generators in water installations also has the security benefit that the installation is inside chambers and buildings. The City of Tshwane

### Table 2 Potential annual hydropower generation capacity at the ten most favourable reservoirs in the City of Tshwane Water Distribution System

<table>
<thead>
<tr>
<th>Reservoirs</th>
<th>TWL (m.a.s.l)</th>
<th>Capacity (kl)</th>
<th>Pressure (m)</th>
<th>Flow (l/s)</th>
<th>Annual potential power generation (kWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garsfontein</td>
<td>1 508.4</td>
<td>60 000</td>
<td>165</td>
<td>1 850</td>
<td>3 278 980</td>
</tr>
<tr>
<td>Wonderboom</td>
<td>1 351.8</td>
<td>22 750</td>
<td>256</td>
<td>470</td>
<td>1 292 471</td>
</tr>
<tr>
<td>Heights LL</td>
<td>1 469.6</td>
<td>55 050</td>
<td>154</td>
<td>510</td>
<td>843 673</td>
</tr>
<tr>
<td>Heights HL</td>
<td>1 506.9</td>
<td>92 000</td>
<td>204</td>
<td>340</td>
<td>745 062</td>
</tr>
<tr>
<td>Soshangueve DD</td>
<td>1 249.5</td>
<td>40 000</td>
<td>168</td>
<td>400</td>
<td>721 859</td>
</tr>
<tr>
<td>Waverley HL</td>
<td>1 383.2</td>
<td>4 550</td>
<td>133</td>
<td>505</td>
<td>721 483</td>
</tr>
<tr>
<td>Akasia</td>
<td>1 413.8</td>
<td>15 000</td>
<td>190</td>
<td>340</td>
<td>693 930</td>
</tr>
<tr>
<td>Clifton</td>
<td>1 506.4</td>
<td>27 866</td>
<td>196</td>
<td>315</td>
<td>663 208</td>
</tr>
<tr>
<td>Magalies</td>
<td>1 438.0</td>
<td>51 700</td>
<td>166</td>
<td>350</td>
<td>624 107</td>
</tr>
<tr>
<td>Montana</td>
<td>1 387.6</td>
<td>28 000</td>
<td>82</td>
<td>463</td>
<td>407 829</td>
</tr>
</tbody>
</table>

Total calculated annual power generation in Tshwane ± 10 000 000
Figure 3 Hydropower generation capacity at different reservoirs in the City of Tshwane Water Distribution System
experiences frequent vandalism and break-ins in all of its isolated infrastructure (removal of solar panels, batteries, electronic equipment and precious metal components).

Table 3 indicates the sensitivity of the assumption used in the calculation of hydropower generation at the ten reservoirs listed in Table 2 for a number of alternative scenarios.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of available head used for power generation</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>0.75</td>
</tr>
<tr>
<td>Generating hours per day</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Fraction of total potential energy generated</td>
<td>0.125</td>
<td>0.175</td>
<td>0.233</td>
<td>0.281</td>
</tr>
<tr>
<td>Yearly generation capacity, MWh</td>
<td>10 000</td>
<td>14 000</td>
<td>18 640</td>
<td>22 480</td>
</tr>
</tbody>
</table>

Table 4 Potential yearly income from power generation at the Queenswood Reservoir

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>80</td>
<td>m</td>
</tr>
<tr>
<td>AADD</td>
<td>181.03</td>
<td>m³/h</td>
</tr>
<tr>
<td>Efficiency</td>
<td>40</td>
<td>%</td>
</tr>
<tr>
<td>Maximum available power</td>
<td>39.46</td>
<td>kW</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Available head (m)</th>
<th>Potential annual energy production (kWh)</th>
<th>% of AADD that could be used to generate electricity</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>27 657</td>
<td>31 114 34 571</td>
</tr>
<tr>
<td>50</td>
<td>34 571</td>
<td>38 892 43 214</td>
</tr>
<tr>
<td>60</td>
<td>41 485</td>
<td>46 671 51 856</td>
</tr>
<tr>
<td>70</td>
<td>48 399</td>
<td>54 449 60 499</td>
</tr>
<tr>
<td>80</td>
<td>55 314</td>
<td>62 228 69 142</td>
</tr>
</tbody>
</table>

PILOT PLANT AT QUEENSWOOD RESERVOIR

The only civil works that were required at the Queenswood Reservoir (location shown in Figure 3) were the installation of a bypass onto the existing pipeline and the fitting of a turbine, generator and other essential electrical equipment.

As turbines used in small-scale hydropower are fairly difficult to procure, are expensive and have long delivery periods, it was decided for the preliminary investigation to use a pump as a turbine (PAT), i.e. to utilise the pump in reverse.

The pump which was used in the setup as shown in Figure 4 is a Sulzer AZ–100/400 pump. Its best efficiency point (BEP) is at a flow rate of 180 m³/hour and a 50 m head, with 34 kW power required. It should be noted that, because the pump was operated as a turbine, the inlet of the pump became the outlet and vice versa.

Estimates of characteristic curves for the selected pump operating as a turbine have been provided by Sulzer SA for the purpose of this pilot project. An approximation of
In order to determine the power output of the generator, a ballast load was connected directly to the generator, effectively ‘throwing away’ the electricity generated. A load has to be connected in order to be able to measure the current and the voltage produced so that the power output can be calculated. The ballast load used was six 4 kWh geyser elements connected to the generator in pairs (in series) as shown in Figure 5. The geyser elements were placed in a tank with water, therefore consuming the electricity generated.

In the case of a permanent installation, the output of the generator would have to be regulated to ensure that it is at the correct frequency (50 Hz). The generator output is an alternating current (AC), but because of the fluctuations in turbine operating conditions, the output of the generator also fluctuates and a variable frequency and voltage output is produced. For on-site generation the output of the generator would have to be connected to a rectifier which converts the current into direct current (DC), and then connected to an inverter which converts the current back into AC but regulates the frequency to a constant and stable 50 Hz.

Based on the AADD (Average Annual Daily Demand) from Queenswood Reservoir of ± 180 m³/hr, the potential energy generation can be determined as shown in Table 4.

PILOT PLANT AT PIERRE VAN RYNEVELD RESERVOIR

The Pierre van Rynelvd (PvR) Reservoir is located south of Pretoria, as shown in Figure 3. Although the site is not one of the top ten favourites listed in Table 2, the site was selected due to the construction of a new 15 ML reservoir near the existing reservoir. This provided the opportunity to construct the second conduit-hydropower pilot plant on the existing reservoir in the Country Lane Estate.

The generated power is utilised on-site for lighting, alarm, communication, etc. The home owners association of the Country Lane Estate has also indicated that they would like to utilise the power for street lighting.

In order to identify the generation potential at this site, some basic data needed to be recorded. The variation in flow rate and available at the site needed to be captured. The basic set-up was to measure and record pressure heads at relevant points along the supply line from the off-take of a Rand Water bulk supply line up to the reservoir.

The outcome of the three extensive field experiments provided confirmation that there is sufficient flow and pressure at the inflow to the Pierre van Rynelvd Reservoir to generate electric power on a pico scale. The results of testing also indicated that small pressure surges occur in the system; this will be used as a benchmark to ensure that the hydropower plant does not become an increased risk for the pipe system.

The pilot plant was constructed on the roof of the old 7.6 ML reservoir (see Figure 6), utilising a cross-flow turbine and a synchronous generator (Figure 7). The maximum capacity is ± 15 kW of renewable, zero-emissions
electricity, but depends on the flow and head pressure conditions at any given time.

The off-take from the main supply line to the hydropower plant on the roof of the reservoir is shown in Figure 8, and the completed installation in Figure 9.

On 29 November 2011 the Pierre van Rynveld Conduit Hydropower Plant was launched jointly by the City of Tshwane, the Water Research Commission and the University of Pretoria, where the City of Tshwane Metropolitan Municipality switched all the site lighting from the conventional municipal grid over onto the hydropower generated on-site.

Some of the problems and challenges faced at this installation, and that are currently being attended to, include:

- Frequency control of the generator
- Sudden load rejection of the system
- Hunting of the PRV due to slow response time
- Significant variability in supply (of flow and pressure head)
- The fact that the system had to be operated and controlled manually.

DEVELOPMENT OF HYDROPOWER SITES IN THE CITY OF TSHWANE

From the potential sites listed in Table 2, five sites were selected for possible development as listed in Table 5.
### Table 5 Selected sites for hydropower development in the City of Tshwane

<table>
<thead>
<tr>
<th>Reservoir</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wonderboom Reservoir</td>
<td>There is one reservoir on this site with a new reservoir being built a few hundred metres northwest of the existing reservoir.</td>
</tr>
<tr>
<td>Akasia Reservoir</td>
<td>There is one reservoir on this site. This site is situated just south of the Hartebeeshoek site of Rand Water, which is one of the reservoir sites of Rand Water being considered for hydropower development.</td>
</tr>
<tr>
<td>Garsfontein Reservoir</td>
<td>There are three reservoirs on this site.</td>
</tr>
<tr>
<td>Heights Reservoir</td>
<td>There are two reservoirs on this site – Low Level (LL) and High Level (HL).</td>
</tr>
<tr>
<td>Waverley Reservoir</td>
<td>There is one reservoir on this site.</td>
</tr>
</tbody>
</table>

It is considered that conduit-hydropower is the “low-hanging fruit” in terms of viable renewable energy which could be developed. The City of Tshwane is in the advantageous situation that excess energy is currently being dissipated, and this could be utilised to generate clean sustainable energy instead.
As an example, a weeklong pressure and flow data record for the Garsfontein Reservoir (number one on the generation potential list in Table 2, i.e. the City of Tshwane’s bulk reservoirs) is depicted in Figure 10. The average pressure upstream of the PRVs is 117.4 m and the average flow is equal to 0.671 m³/s for this specific week. The required pressure downstream of the PRVs is 8 m. Data at this site has been recorded since April 2011, and the average pressure upstream of the PRVs is 120 m and the average flow is equal to 0.77 m³/s.

The hydropower potential for this selected week based on a conservative low efficiency of 70% for the turbine/generator set is shown in Table 6. The two computation methods used are:
- 15 minute intervals – flow and pressure readings for every 15 minutes are recorded. The generation potential is calculated for each of these intervals and then totalled. This would require the selection of a turbine which could operate with a high efficiency over a wide flow and pressure range.
- Weekly averages – the average pressure and flow for this week is determined and used to calculate the generation potential. This would require operational changes to the supply and reservoir system to provide a more stable flow and pressure range. This would, however, result in an increase in potential electrical output. The annual potential is simply an extrapolation of the recorded potential during this week.

A water supply distribution system consists of a complex network of interconnected pipes, service reservoirs and pumps that deliver water from the treatment plant to a consumer. The distribution of water through the supply system is governed by complex, non-linear, non-convex and discontinuous hydraulic equations (Keedwell & Khu 2005). Adding to this complex network, the hydropower plant from which the maximum benefit needs to be extracted requires a systematic procedure to evaluate the inter-relationship between: storage volumes, supply/demand patterns, turbine selection, operational flexibility and reliability of supply.

The required turbine based on the average flow (0.77 m³/s) and corresponding pressure (110 m) is 550 kW, but if a turbine is sought that follows the fluctuating supply patterns, a 950 kW system will be required. The final selection of the correctly sized turbine will depend on multiple factors. This still requires further investigation. Similar calculations were done for the other four sites.

<table>
<thead>
<tr>
<th>Table 6 Garsfontein Reservoir hydropower potential</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generation potential</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Weekly potential</td>
</tr>
<tr>
<td>Annual potential¥</td>
</tr>
</tbody>
</table>

Note: ¥ Allowance for two days per year for reservoir maintenance

---

<table>
<thead>
<tr>
<th>Table 7 Cost components of a conduit-hydropower scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost and formula</strong></td>
</tr>
<tr>
<td>Initial Planning Cost (IPC) = C\text{investigation} + C\text{environment &amp; social} + C\text{legal &amp; regulatory}</td>
</tr>
<tr>
<td>Capital Cost (CEC) = C\text{design} + C\text{purchase} + C\text{installation} + C\text{start-up}</td>
</tr>
<tr>
<td>Operation and Maintenance Cost (OMC) = C\text{operation} + C\text{maintenance} + C\text{management} + C\text{refurbishment}</td>
</tr>
<tr>
<td>Disposal Cost (DC) = C\text{disposal} + C\text{environmental}</td>
</tr>
</tbody>
</table>

FEASIBILITY OF DEVELOPMENT
South Africa does not have many turbine manufacturers, and thus turbines and generators must usually be imported. There is great disparity (factor of four) between purchase costs of turbines from different manufacturers across the world, with some developing countries supplying turbines at a much lower cost (although the durability of some of these machines could be questionable) (Van Vuuren et al 2011). A broad guideline is that a 1 MW hydropower installation would cost between R16.5 and R22 million. This is composed of the cost components as listed in Table 7.

As an alternative to the generally costly micro turbines, pumps-as-turbines (PATs) can be considered. Pumps have the advantages of being more readily available, easy to operate and maintain, and are generally less expensive than micro turbines. Various companies produce purpose-made PATs that can run at efficiencies of as high as 90% (Sanjay & Patel 2014), but in principle the impeller of a centrifugal pump can be turned around to produce a PAT with efficiency of around 30%. The application of PATs has been extensively documented (Sanjay & Patel 2014; Williams 2003).

Typically, all engineering projects will incur on-going revenue costs, maintenance costs and ultimately replacement costs. Therefore the long-term cost must be considered examining the relationship between the value of money and time. Life cycle costing (LCC) includes all costs associated with a system (or component) as applied over the defined life cycle. However, as a first order analysis, the expected payback period of...
capital cost could be calculated to provide an indication of feasibility.

The preliminary selected sites were sized, utilising a conservative load factor of 0.8 and a turbine system efficiency of 70%. The preliminary feasibility based on a Megaflex tariff (2013 base year) of 58 c/kwh indicates that all five these sites are feasible, as shown in Table 8. All these sites have a payback period of between six and seven years.

The preliminary feasibility results in Table 8 are based on the sizing of a turbine to operate at the average flow and pressure at the site. If a turbine is sought that follows the fluctuating supply patterns, a larger system will be required, which could increase the costs. The correct sizing and selecting of the turbine can only be done after flow and pressure data have been obtained, and the operating range of the system has been determined.

CONCLUSIONS

Hydropower represents a nexus of water and energy, and in municipalities and water utilities there are several locations where a feasible conduit-hydropower scheme could be implemented. A technically feasible scheme assists in reducing operating costs, mainly due to energy increases, and provides a sustainable solution whilst having a positive environmental impact. A number of water utilities have started taking the initiative in developing this type of hydropower and it is believed that there is significant potential in South Africa.

There are numerous benefits for developing conduit-hydropower in the City of Tshwane’s water distribution network:

- Hydroelectric energy is a continuously renewable energy source.
- Hydroelectric energy technology is proven technology offering reliable and flexible operations.
- Hydroelectric stations have a long life – many existing stations have been in operation for more than half a century and are still operating efficiently (an example of this is in Cape Town).
- Micro hydropower stations achieve high efficiencies. Purpose-made PATs also operate on efficiencies of up to 90% at best efficiency point, but in general PATs achieve efficiencies of around 30%.
- Conduit-hydropower uses the available water distribution infrastructure, and thus, as long as there is a demand for water, hydroelectric energy can be generated.
- The operational life of the existing pressure reducing valves can be extended.
- Conduit-hydropower “piggybacks” onto existing water infrastructure resulting in minimal environmental impact.

The preliminary feasibility studies indicate short payback periods.

Conduit-hydropower has the potential of mitigating vandalism of local power sources (e.g. solar panels) at remote reservoirs required for reservoir status monitoring. Depending on the generating potential of the installation, local domestic energy clusters could also benefit. The feasibility and construction of two pilot plants were also discussed, and it was shown that it is technically possible and feasible to install turbines in pressurised water supply pipes to utilise excess pressure head.

It is considered that conduit-hydropower is the “low-hanging fruit” in terms of viable renewable energy which could be developed. The City of Tshwane is in the advantageous situation that excess energy is currently being dissipated, and this could be utilised to generate clean sustainable energy instead.

ACKNOWLEDGEMENT

The research presented in this paper emanated from a study funded by the Water Research Commission (WRC) whose support is acknowledged with gratitude.

REFERENCES


Table 8 Feasibility of sites

<table>
<thead>
<tr>
<th>Reservoir site</th>
<th>Estimated average capacity (kW)</th>
<th>Estimated annual generation potential (kwh/a)</th>
<th>Estimated development cost (R)*</th>
<th>Estimated revenue (year 1) (R)**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wonderboom #2</td>
<td>330</td>
<td>2 312 640</td>
<td>7 260 000</td>
<td>902 000</td>
</tr>
<tr>
<td>Heights #3</td>
<td>380</td>
<td>2 663 040</td>
<td>8 360 000</td>
<td>1 040 000</td>
</tr>
<tr>
<td>Akasia #4</td>
<td>260</td>
<td>1 822 080</td>
<td>5 720 000</td>
<td>711 000</td>
</tr>
<tr>
<td>Waverley #5</td>
<td>80</td>
<td>560 640</td>
<td>1 760 000</td>
<td>220 000</td>
</tr>
<tr>
<td>Garsfontein #6</td>
<td>550</td>
<td>3 854 400</td>
<td>12 100 000</td>
<td>1 504 000</td>
</tr>
<tr>
<td>Total</td>
<td>1 600</td>
<td>11 212 800</td>
<td>35 200 000</td>
<td>2 873 000</td>
</tr>
</tbody>
</table>

Notes:
- * Initial planning, design and capital costs
- ** Based on IMQS data (no historical data available)
- Utilising a conservative load factor of 0.8, turbine system efficiency of 70%, averaged Megaflex tariff of 50 c/kwh and subtracting anticipated O&M costs
Evaluating the prioritisation of South African dams for rehabilitation with special focus on risk to human lives

S Reynolds, C Barnardo-Viljoen

Existing decision criteria for the prioritisation of dam rehabilitation in South Africa are evaluated. In this context risk to human lives and economic considerations are the most important decision drivers, although other considerations are also taken into account by the Department of Water Affairs. The Department’s assessment of risk to human lives is compared to international best practice and prediction models to show that the assessment is currently somewhat inconsistent, resulting in over-conservative decisions for low-consequence situations and possibly too low safety levels for severe-consequence situations.

Reduction of the life-safety risks comes at significant costs, which may be better spent elsewhere. Society’s Willingness to Pay is used to determine the lower boundary for acceptable investments in life safety. Investments for improved safety could also be made for economic reasons. Economic optimisation would often imply higher safety levels than required by Society’s Willingness to Pay. These concepts are applied to case studies of actual South African dam rehabilitation projects, allowing evaluation of the quality of decisions taken.

Based on the above, life-safety criteria that incorporate a measure of the economic efficiency of the proposed rehabilitation are developed, and suggestions are made to improve the current decision criteria used by the Department of Water Affairs.

INTRODUCTION

In South Africa, the Department of Water Affairs (DWA) is a national governmental department that is the custodian of a large number of dams. A risk-based model has been developed and is currently applied by the DWA to aid decisions regarding the adequacy of dam safety levels. For a specific dam, the estimated probability and consequences of dam failure are combined to define risks. These risks are evaluated against multiple acceptability criteria on five impact diagrams to assess the risk to human life, and the economic, social, socio-economic and environmental impacts of dam failure. A sixth diagram is used by the DWA, the risk level diagram, where the annual risk of fatalities per exposed hour is evaluated against the annual risk of financial losses (Hattingh & Oosthuizen 2009). If any of these risks are considered unacceptable, the rehabilitation of the dam to improve its safety may be recommended.

In the years 2004/2005, the DWA identified 166 of the 314 South African government-owned dams as being in need of rehabilitation works. This encouraged the initiation of the dam safety rehabilitation programme, and it is estimated that the total expenditure for rehabilitation works since the start of the programme in 2005/2006, up until the 2011/2012 financial year, is more than R1.5 billion (Segers 2012).

The DWA recently identified the need to review its acceptability criteria for risk to human life.

Dam rehabilitation should reduce the probability of dam failure, thereby reducing the risk to human life. Since rehabilitation comes at large costs and it is society that essentially finances dam rehabilitations via public taxes or charges, it should be ensured that these investments in life safety are actually worthwhile for society. In this sense it must be noted that the societal resources that can be allocated to improving life safety through dam rehabilitation are limited. If the cost of reducing the risk to human life is disproportionate to the actual risk reduction, these resources may be better redirected into other sectors, for example into health care, transportation services or education, to improve the quality of life of society. Society’s Willingness to Pay (SWTP) is a utility function which may be used to determine the acceptable level of expenditure into life safety required by society (Pandey et al 2006).

On the other hand, investments in dam rehabilitation works are not always driven by societal preferences for life safety, but...
are often made for economic reasons. In these cases, larger investments than what is required by society can be justified, and their magnitudes can be determined by economic optimisation. These additional economic considerations could include economic motivations for the existence of the facility, damage costs and compensation costs for lives lost due to dam failure or the environmental, socio-economic and social implications of dam failure.

It is unlikely for the SWTP boundary to govern the investment into rehabilitation works, as economic optimisation would typically dictate the decision (Rackwitz & Streicher 2002). Thus, criteria which effectively incorporate this consideration into the decision process are needed.

The aim of this paper is:
- To briefly review international best practice methods to quantitatively evaluate risk to human lives.
- To evaluate the current DWA life-safety criteria by comparing to international best practice criteria.
- To define the lower boundary for investments in life safety required by society, using SWTP.
- To investigate economic motivations for further investments, above the lower boundary for life safety defined by SWTP.
- To propose criteria, primarily to evaluate life safety, but which implicitly incorporates some measure of the economic efficiency of the rehabilitation works.

INTERNATIONAL BEST PRACTICE METHODS FOR QUANTITATIVELY EVALUATING RISK TO HUMAN LIVES

Internationally, risk to life is most commonly quantitatively assessed as the expected fatalities per year against FN-criteria on an FN-diagram (Faber 2009). FN-diagrams have a double logarithmic scale with the x-axis representing the number of fatalities (N) and the y-axis the annual probability (F) of N or more fatalities occurring (Kroon & Maes 2008).

FN-criteria could typically be defined by two properties, namely the intersection with the y-axis and the slope of the criterion line. If the criteria intercept the y-axis at a lower probability (F) of N or more fatalities occurring, the criteria are more stringent. The slope of the criterion line describes the risk aversion of a society, which is the additional public opposition to an event which kills a large number of people over a series of smaller events that collectively result in the same number of fatalities (Ball & Floyd 1998). A slope of –1 represents a “risk neutral” society, while an increased slope, for example –2, is more stringent and describes “risk aversion”.

Furthermore, the criteria can define different regions for risk – risks that are so high that they are to be judged as unacceptable/intolerable, risks that are so low that they are to be judged as acceptable/negligible, and risks that are regarded as tolerable only if they are reduced to be As Low As Reasonably Practicable (ALARP) (Ball & Floyd 1998).

The implementation of the ALARP principle requires consideration of the trade-off between the risk and the time, the cost and the physical difficulty of implementing the risk reduction measure. If the cost of a safety measure is disproportionate to the actual risk reduction, it is not reasonably practicable to implement the safety measure, and a higher risk is accepted (HSE 2001).

Internationally, FN-criteria have been developed for life-safety risks associated with large-scale facilities, including nuclear and offshore facilities and the transport of dangerous goods. According to Ball and Floyd (1998) there are similarities between the criteria as it developed for these industries in the United Kingdom, the Netherlands and Hong Kong. The upper limit of tolerability is often set at 10^{-4} for 10 or more fatalities (or 10^{-5} in the Netherlands). The acceptable/negligible line tends to be located two or three factors of 10 (100 or 1 000) lower on the frequency (F) scale. Most of the criteria used in the UK and in Hong Kong have a gradient of –1, whereas the Netherlands criteria are generally set at –2. This is due to different regimes of control. According to Ball and Floyd (1998) there is no compelling rationale for incorporating risk aversion into the FN-criteria defined for these industries, and generally a slope of –1, corresponding to risk neutrality, is regarded as good practice.

The criteria that have been developed for these industries may not necessarily be directly applied to dam safety, since it may be reasonably impracticable to accept the same safety levels. Expert judgement should be applied to establish industry-specific criteria.

The International Commission on Large Dams (ICOLD) outlines the current application of risk-based methods in international dam safety in Bulletin 130 on Risk Assessment in Dam Safety Management (2005). According to ICOLD (2005), many countries acknowledge that risk-based tools are useful within dam safety, but there are contradicting views and opinions, and some countries are hesitant to explicitly define FN-criteria for life safety.

In Australia, the Australian National Committee on Large Dams (ANCOLD) proposed FN-criteria as shown in Figure 1. The following properties can be observed from the criteria:
- Different criteria are defined for new and existing dams. According to ANCOLD (2003), the marginal cost of reducing risk at existing dams is generally more than at new dams. Thus, it is not reasonably practicable to accept same safety levels.
- The acceptability limit is set at 10^{-4} for ten or more fatalities for existing dams, and at 10^{-5} for new dams, corresponding to the trend observed by Ball and Floyd (1998) for other industries.
- A risk-neutral slope of –1 is used, corresponding to the Ball and Floyd (1998) recommendation.

![Figure 1 ANCOLD FN-criteria for new and existing dams (ANCOLD 2003)](image-url)
The criteria have a lower probability of failure cut-off. According to ANCOLD (2003), technology does not allow for the construction of dams with lower probabilities of failure, and it is not reasonably practicable to reduce dam safety levels to more stringent criteria.

The ANCOLD criteria are thus based on engineering judgement, implicitly incorporating cost considerations for reasonable practice.

Several other dam safety organisations, such as the New South Wales Dam Safety Committee (NSW-DSC) also in Australia, and the US Army Corps of Engineers (USACE) in the USA, have based their criteria on the ANCOLD criteria. It was therefore decided to compare South African dam safety criteria for risk to human life to the ANCOLD criteria.

**EVALUATION OF SOUTH AFRICAN DAM SAFETY CRITERIA FOR RISK TO HUMAN LIVES**

The Department of Water Affairs (DWA) quantitatively estimates life-safety risks as the combination of the annual probability of dam failure \( P_f \) and the population at risk (PAR), i.e. the number of people exposed to the dam-break flood. These risks are evaluated against criteria presented on a PAR diagram, with the x-axis representing the PAR and the y-axis the probability of occurrence, as shown in Figure 2 (Hattingh & Oosthuizen 2009).

In Figure 2 an example is shown how the DWA depicts a dam’s risk on this type of graph. To demonstrate the uncertainty associated with estimating risk, the DWA estimates ranges for the PAR and the probability of occurrence.

International methods assess risk to life most commonly as expected fatalities. Thus, since two different consequence measures are used, the South African dam safety criteria could not be directly compared to ANCOLD criteria.

The DWA uses its own in-house developed model to predict what portion of the PAR would become fatalities, based on assumptions related to warning times (WTs) available to the PAR in the event of a dam break. The number of fatalities, expressed as the loss of life (LOL) by the DWA, can be estimated using the diagram shown in Figure 3.

The DWA criteria are compared to ANCOLD criteria by finding, for a certain \( P_f \), the implied warning time (WT) needed such that the DWA-predicted loss of life would correspond to that of ANCOLD.

The implied WTs needed for the DWA criteria to adhere to ANCOLD criteria for new and existing dams are summarised in Table 1. For existing dams, at high probabilities of failure, long WTs are needed for DWA
criteria to adhere to ANCOLD criteria. In practice these high WTs are often not realistically achievable, and thus such DWA dams are accepted at less stringent safety levels than ANCOLD dams. At low probabilities of failure, the WTs are small and probably easily achievable. Consequently DWA overdesigns for low probability events, implying risk-averse behaviour.

A similar pattern is observed when the implied WTs needed for new dams are computed. As the probability of failure decreases, the implied WT decreases. The WTs needed for the criteria to correspond are higher than for existing dams, implying that new DWA dams are accepted at less stringent safety levels than ANCOLD new dams. This is expected because the DWA does not differentiate between new and existing dams in its acceptance criteria.

The DWA model for predicting loss of life was developed based on historical data for dam failures (Hattingh & Oosthuizen 2009). The statistical basis is not documented. It was therefore decided to validate the DWA prediction model by comparing it to another internationally accepted prediction model.

The DeKay and McClelland (1993) model uses a regression approach to predict the loss of life due to dam failure from the population at risk and assumptions related to the warning time, similar to the DWA model. It is based on the statistical analysis of actual historical data compiled from the 1950s onwards, which applies to a wide range of populations at risk. The DeKay and McClelland (1993) predicted life loss compares well with the actual historical data.

The DeKay and McClelland (1993) model incorporates an additional factor, the force factor, accounting for the severity of the dam-break flood on the predicted life loss. The High Force (HF) condition refers to the scenario where the PAR is located in a canyon and the flood waters due to dam failure are very deep and swift. The Low Force (LF) condition is where the PAR is located on a plain and the flood waters are shallow and slow. The DeKay and McClelland (1993) equations for determining the loss of life (LOL) from the population at risk (PAR) and warning time (WT) for both HF and LF conditions are:

\[
LOL_{HF} \approx 0.075(PAR_{HF}^{0.560})e^{-2.982(WT_{HF})+3.790} \quad (1)
\]

\[
LOL_{LF} \approx 0.075(PAR_{LF}^{0.560})e^{-0.759(WT_{LF})} \quad (2)
\]

The DeKay and McClelland (1993) predicted loss of life was compared to DWA-predicted values for a range of population at risk and for three different warning times, namely a “small” WT (0 minutes), a “medium” WT (30–45 minutes) and a “large” WT (90 minutes). The comparison between the two life-loss prediction models is shown in Figure 4 and Figure 5 for HF and LF conditions respectively.

For both HF and LF flood conditions the DeKay and McClelland (1993) model predicts the loss of life within a narrower range than the DWA model. For HF conditions DWA generally over-predicts the loss of life for small and medium WTs. This may lead to conservative decision-making regarding life safety where severe consequences are expected. In this way the DWA unwittingly incorporates risk aversion in decision-making. For LF conditions the DWA severely over-predicts the loss of life for small and medium WTs. Thus, the DWA predictions are too conservative for conditions where low consequences are expected. For large WTs, the DWA generally
under-predicts the LOL in comparison to DeKay and McClelland (1993).

The conservative life-safety decisions implied by the DWA life-loss prediction model may, to some extent, off-set the moderate life-safety decisions of the DWA that are implied when the warning times needed for DWA criteria to adhere to ANCOLD criteria are computed. Until better South African data can be found to recalibrate the current DWA life-loss prediction model, we would, however propose that it be replaced by the DeKay and McClelland (1993) model, which has a well-documented and rational scientific basis.

A further comparison of the DWA criteria to ANCOLD can be made by using the DeKay and McClelland (1993) prediction model to convert "population at risk" to "loss of life". For different probabilities of failure ($P_f$), the corresponding PAR is obtained from the PAR-criteria used by the DWA (Figure 2). For the PAR and an assumed WT the loss of life (LOL) is determined through the DeKay and McClelland (1993) prediction model. The LOL is plotted against the $P_f$ to obtain a DWA equivalent FN-criterion line, which can be directly compared to the ANCOLD FN-criteria.

DWA equivalent FN-criteria are obtained for three assumed warning times (WTs) and for HF and LF conditions, as shown in Figure 6 and Figure 7.

Different DWA criteria lines were obtained, depending on assumptions regarding the available warning time and flood severity condition. This result implies a fundamental flaw in the currently used DWA criteria: while warning time and flood severity will influence the risk level that is associated with a given dam, the criteria that dictate what level of risk is deemed to be acceptable should be independent of the underlying characteristics of individual dams. The ALARP principle may be used as an argument to define different acceptability criteria for broad categories where reasonable practice may dictate less stringent safety requirements. For example, this is the argument behind ANCOLD's different criteria lines for new and existing dams.

For small WT and HF conditions the DWA equivalent FN-criteria were less stringent than the ANCOLD criteria, implying less stringent safety levels in cases where severe consequences are expected. For large WT and LF conditions, the DWA equivalent criteria were more stringent, implying too conservative decision-making in cases where low consequences are expected.

Further, the gradient of the DWA FN-criteria lines are steeper than the risk-neutral slope of -1 used by ANCOLD.

Thus, several problems with the current formulation of DWA life-safety criteria came to light by comparing them to the equivalent ANCOLD criteria. We propose that the DWA eliminates these by switching to the ANCOLD life-safety criteria, for the following reasons:

- The ANCOLD criteria evaluate risk to life using fatalities as a consequence measure, which is most commonly used internationally. Using PAR as a consequence measure is fundamentally flawed, since the numbers of fatalities that may come from the PAR are greatly dependent on factors such as warning time and flood severity, which in turn are dam-specific.
- An internationally recommended risk neutral slope of -1 is used by ANCOLD.

Switching to ANCOLD would not imply a substantial change in the current DWA safety levels. It would, however, imply a more consistent treatment of risk across the board of different warning times and flood severity levels. Also, using ANCOLD criteria would not imply more risk analysis effort than what is currently required, since the DWA already...
estimates the loss of life as part of its standard risk analysis procedures.

The risk to life for eleven case studies of DWA-owned dams that have been identified to be in need of rehabilitation are evaluated against ANCOLD FN-criteria for existing dams. The estimated initial probability of failure and LOL are obtained from DWA dam safety evaluation reports as shown in Table 2.

For Bospoort Dam two different scenarios were considered in the DWA risk analysis – Case 1a, where the sluice gates were assumed to function normally during failure, and Case 1b, with the gates not functioning during failure.

In Figure 8 it is seen that the risk to life for the case studies are mostly within the unacceptable region of the ANCOLD criteria for existing dams, justifying the original DWA rehabilitation decision.

The ANCOLD criteria do not only consider the societal preferences for investments in life safety, but implicitly take other considerations into account, for example economic considerations for reasonable practice. Societal Willingness To Pay is proposed to determine the acceptable level of these life-safety investments required by society.

**SOCIETAL WILLINGNESS TO PAY (SWTP) AS A LOWER BOUNDARY CONSTRAINT ON DAM SAFETY LEVELS**

The reduction of the life-safety risks of a dam through rehabilitation works comes at a cost. Society essentially finances dam rehabilitation via public taxes or charges; therefore it should be ensured that these investments in life safety are actually worthwhile to society. SWTP is a utility function which effectively determines the lower boundary for investments in life safety required by society (Pandey et al 2006). It is based on the Life Quality Index (LQI), which jointly considers the social indicators of a nation to give a measure of the quality of life of a society (Pandey & Nathwani 2004). In a simple form, the LQI can be written as:

\[
L = GqE
\]

where \(G\) represents the Gross Domestic Product (GDP) per person, \(E\) the life expectancy at birth, and \(q\), a parameter which reflects the trade-off placed on consumption and the value attached to length of life. The parameter \(q\) depends on the fraction of time spent producing \(G\), and the remaining time, the leisure time, available for the enjoyment of \(E\). It is the ratio of average work time \((w)\) to leisure time \((1-w)\).

An investment in life safety should lead to an improved life quality. A small change in the LQI due to the implementation of a safety measure is shown (Nathwani et al 2008) as:

\[
\frac{dL}{L} = \frac{dG}{G} + q \frac{dE}{E}
\]

where \(dG\) corresponds to the monetary cost of implementing the project (negative), \(dE\) the change in the life expectancy due to a
change in the risk associated with the project and $K = 1/q$.

The LQI net benefit criterion requires that an investment into life safety, which influences both $G$ and $E$, should lead to a positive change in the LQI, i.e. $dL/L ≥ 0$ (Pandey & Nathwani 2004). SWTP defines the lower boundary for acceptable decisions and may be obtained as the exact value ($dL/L = 0$) of Equation 4:

$$-dG ≥ SWTP = GKdE/E = GKCdμ[R/person/year]$$

(5)

Society requires that an investment, $−dG$, into a life-saving activity should at least be equal to the SWTP for a marginal increase in life expectancy (Fischer et al 2011).

The parameter $dE/E$ may not always be easily quantified; instead it may be calculated as the product of the mortality change ($dμ$) and a demographic constant ($C_γ$). The demographic constant takes age-averaging and discounting into account. For age-averaging two mortality reduction schemes may be considered, namely the $π$-regime, where the change in mortality is proportional over the age distribution, i.e. it implies that persons who are more susceptible to mortality (typically due to weakened physical state) are more subject to the phenomenon, and the $Δ$-regime, where the change in mortality is uniformly distributed over all ages, i.e. it implies that a phenomenon will affect every member of a society, regardless of each individual's age (Lentz 2007). The discount rate, also referred to as the time preference for consumption, compensates for the fact that individuals tend to undervalue the prospect of future consumption compared to current consumption.

If investments are made into risk reducing activities, a “technology curve” may be obtained, as shown in Figure 9. As the investment cost into life safety increases ($ΔC$), the risk to life is reduced ($ΔN$).

The shape of the curve depends on the effectiveness and cost of life-saving measures. For different activities, projects and technologies, the curves differ, since different risk reduction options are typically available, some more effective than others. Society should implement all the safety measures that are more effective than the threshold set by SWTP, i.e. if the investment cost per marginal life saved ($ΔC/ΔN$) is less than the SWTP for a marginal increase in life expectancy, the investment in life safety is efficient and should be made.

The absolute lower boundary for investments in life safety required by SWTP can be defined as:

$$SWTP = \frac{ΔC}{ΔN} \frac{R/life}{R/year} = \frac{R/year}{lives/year}$$

(6)

Thus for a dam rehabilitation project, for the investment into life safety ($ΔC$) to be considered efficient by society through SWTP, the minimum required reduction in risk to life ($ΔN$) could be determined.

Considering the basic principle that expected risk is the product of probability and consequence, the minimum reduction in risk to human life ($ΔN$) could be expressed as a function of the reduction in the probability of dam failure $ΔP_f$ (which depends on the effectiveness of the rehabilitation strategy) and the estimated number of lives lost (LOL) in case of dam failure:

$$ΔN = ΔP_f \cdot LOL$$

(7)

This relationship implies that a rehabilitation investment may be considered inefficient if:

- the rehabilitation strategy available is not effective, leading to a small reduction in $ΔP_f$
- the dam was already fairly safe, thus rehabilitation leads to only a small improvement in $ΔP_f$
- the number of expected lost lives (LOL) due to failure is already so low that the risk is considered acceptable.

The lowest number of expected lost lives (LOL) for which an investment into life safety is still considered efficient by society can be determined by rearranging Equation 7:
For an initial probability of failure before the dam is rehabilitated, if the LOL estimated by the DWA is more than LOL2, the investment into rehabilitation works is required by society. In this way, FN-criteria lines can be developed to reflect the SWTP threshold for a specific dam, assuming a range of initial probabilities of failure, known rehabilitation costs and the final (rehabilitated) probability of failure.

SWTP-criteria are developed for the same eleven case studies of DWA-owned dams that have been evaluated in terms of ANCOLD criteria. The reduction in the probability of failure ($\Delta P_f$) is determined as the difference between the initial probability of failure ($P_{f\text{, initial}}$), shown in Table 2, and the final (rehabilitated) probability of dam failure ($P_{f\text{, final}}$). The value for $P_{f\text{, final}}$ is assumed to be between $1E-5$ and $1E-6$ per year, which is equivalent to the DWA assumption for a well-engineered dam with no known deficiencies (Oosthuizen 2002). The estimated investment cost for rehabilitation works is obtained from DWA design reports, as shown in Table 3.

To develop SWTP criteria which may be applied to South African dam safety, a reasonable SWTP value should be used. Rackwitz (2008) demonstrates the relationship between the life expectancy at birth and GDP per person for different countries. In Figure 10 it is seen that the two factors are highly correlated across countries.

In South Africa the relationship between life expectancy and GDP per person is an outlier compared to other countries at similar levels of development. Our life expectancy is comparatively low due to factors such as HIV, and our GDP is comparatively high due to our richness in mineral resources (i.e. the GDP is not purely dependent on the income produced through the work time of South African citizens). Furthermore, the low employment rate in South Africa may lead to the misinterpretation of parameter $q$.

A low value for work time ($w$) leads to a higher value for leisure time ($1-w$), implying that South African citizens prefer enjoyment of life over spending time earning a higher income.

Thus, the SWTP value for South Africa may not be a true reflection of our society’s preference regarding investments in life safety. Instead, an Earth value for SWTP (ESWTP) developed by Faber and Virgues-Rodriguez (2011) is used. The ESWTP is based on observations for more than 70% of the Earth’s population and conforms well to the underlying assumptions of the LQI derivation, i.e. the joint development of health and life safety (life expectancy at birth), economy (GDP per person) and the necessary time to work (described by $q$).

If a discount rate (time preference for consumption) of 3% (Arrow 1995), and a uniform mortality reduction scheme ($\Lambda$-regime) are assumed, the ESWTP obtained from Faber and Virgues-Rodriguez (2011) is US$517 000/life. US dollars are converted to the South African currency, rand, using the yearly average exchange rates obtained from the International Revenue Service (2012) from the years 2006 to 2011, within the time frame where the investment cost for rehabilitation works for the case studies were estimated. The average of the exchange rate values results in an ESWTP of R4.048 million/life.

The estimated investment cost for rehabilitation works ($\Delta C$), and the reduction in the probability of dam failure $\Delta P_f$, SWTP criteria lines were developed for the eleven case studies, as shown in Figure 11.

Since the available best practice technologies for rehabilitation works are case specific, the investment cost for reducing risk to life depends on the dam under consideration. Consequently different SWTP criteria lines

\[
\text{LOL}_1 = \frac{\Delta N}{\Delta P_f} 
\]
are obtained for each dam. The positions of these lines are, however, within one log cycle (factor of 10) of one another, implying a fairly low level of sensitivity to factors such as the rehabilitation cost and the SWTP value. Therefore, as long as the values are estimated within the correct order of magnitude, useful criteria may be derived.

The estimated risk to life for the case studies (as the combination of the \( P^{(\text{initial})} \) and LOL shown in Table 2) is evaluated against these criteria lines. In Figure 11 it is seen that only two of the eleven case studies (case studies 4 and 11) required rehabilitation in terms of their SWTP criteria. ANCOLD criteria required rehabilitation for all the cases. ANCOLD, however, implicitly incorporates economic considerations for reasonable practice, while SWTP only accounts for societal preferences for life safety.

Further investments should be made if required by the decision-maker or owner of the facility on the basis of economic optimisation.

### ECONOMIC OPTIMISATION AS A DECISION TOOL FOR EVALUATING SOUTH AFRICAN DAMS FOR REHABILITATION

Economic optimisation requires evaluating the profitability of a project, ensuring a maximum benefit at the lowest cost (Rackwitz 2002). It typically implies higher safety levels than those required by SWTP (Rackwitz & Streicher 2002). But, if the economic optimum is at a lower level than dictated by SWTP, the SWTP minimum safety level should be enforced.

Considering investments in dam rehabilitations, the objective function for the monetary net benefit is:

\[
Z = B - C
\]

where \( B \) represents the benefit and \( C \) the cost of the rehabilitation works. \( B \) does not consider the incomes generated from the existence of the facility, but considers only the additional benefit derived from rehabilitation works, i.e. a reduced probability of dam failure which, in combination with the cost of failure, results in reduced expected cost of failure.

The cost of failure considers the economic losses and the compensation costs for lives lost due to dam failure. The DWA estimates direct and indirect economic losses through the risk analysis methodology, where the direct economic losses could include the damage to the structure, loss of agriculture and the costs of emergency relief, while the indirect economic losses could include the loss of future benefits (Hattingh & Oosthuizen 2009).

The compensation costs for lives lost are determined as the product of the estimated lives lost (LOL) and the Societal Value of a Statistical Life (SVSL). Similar to SWTP, SVSL is derived from the LQI concept (Faber & Virgules-Rodriguez 2011):

\[
SVSL = GKE \ [\text{R}]
\]

where \( G, K \) and \( E \) are as defined for Equations [3] and [4].

SWTP and SVSL should not be confused with each other – SVSL is the amount which should be compensated for each fatality, while SWTP defines the acceptable level for investments in life safety.

In each case study, the net benefit \( (Z) \) was determined for two decision alternatives, namely “do-nothing” and “rehabilitate”. The two alternatives are compared to each other in terms of the expected cost of failure and the associated implementation costs, as shown in Table 4. Economic optimisation requires that the alternative with the highest net benefit should be preferred.

For the eleven dam cases considered in this study, the reduced probability of failure is determined as the difference between the \( P^{(\text{initial})} \), shown in Table 2, and the \( P^{(\text{final})} \) after rehabilitation, assumed to be between 1E-5 and 1E-6. To determine the costs of failure, the direct and indirect economic losses and LOL are obtained from DWA dam safety evaluation reports. The estimated cost of rehabilitation works is obtained from the DWA design reports, as shown in Table 3.

Five of the eleven cases required rehabilitation on the basis of economic optimisation, while SWTP only required rehabilitation in two of these five cases. Therefore economic optimisation in most cases recommended higher safety levels.

The original DWA decision was to rehabilitate all eleven dams. The decision is not only based on life safety and economic considerations, it also considers the environmental, social and socio-economic impacts of dam failure and the risk level of dams, as described in the **Introduction** of this paper. The DWA dam safety evaluation reports for the other six cases reveals that two cases did not really require rehabilitation based on DWA criteria (although a number of risks were judged to be fairly high in these cases). The four remaining cases were rehabilitated based on environmental, social, socio-economic and risk level considerations.

Since economic optimisation in most cases dictated the rehabilitation decision, criteria which effectively incorporate these observations into the decision process are

<table>
<thead>
<tr>
<th>Decision Alternative</th>
<th>Probability of Failure (( P_f ))</th>
<th>Cost of Failure (( C_f ))</th>
<th>Expected Cost of Failure (Combined ( P_f ) and ( C_f ))</th>
<th>Implementation Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Do nothing</strong></td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>No costs</td>
</tr>
<tr>
<td><strong>Rehabilitate</strong></td>
<td>Lowered</td>
<td>High</td>
<td>Lowered</td>
<td>Costs of rehabilitation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case Study No</th>
<th>Dam</th>
<th>Average ( \Delta P_f ) (%/yr)</th>
<th>Cost Estimate (R/yr)</th>
<th>( C/\Delta P_f ) (R/%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Bospoort Dam (gates functioning)</td>
<td>5.49E-01</td>
<td>R6.11 mil/yr</td>
<td>R11.12 mil/%</td>
</tr>
<tr>
<td>1b</td>
<td>Bospoort Dam (gates not functioning)</td>
<td>5.50E+00</td>
<td>R6.11 mil/yr</td>
<td>R11.11 mil/%</td>
</tr>
<tr>
<td>2</td>
<td>Klein Maricopooit Dam</td>
<td>5.45E-02</td>
<td>R2.85 mil/yr</td>
<td>R52.34 mil/%</td>
</tr>
<tr>
<td>3</td>
<td>Toleni Dam</td>
<td>2.74E-01</td>
<td>R1.71 mil/yr</td>
<td>R6.25 mil%</td>
</tr>
<tr>
<td>4</td>
<td>Lakeside Dam</td>
<td>1.09E-01</td>
<td>R1.83 mil/yr</td>
<td>R16.68 mil%</td>
</tr>
<tr>
<td>5</td>
<td>Vaalkop Dam</td>
<td>1.05E-02</td>
<td>R1.76 mil/yr</td>
<td>R167.98 mil%</td>
</tr>
<tr>
<td>6</td>
<td>Rust De Winter Dam</td>
<td>2.70E-02</td>
<td>R1.54 mil/yr</td>
<td>R57.32 mil%</td>
</tr>
<tr>
<td>7</td>
<td>Makotswane Dam</td>
<td>1.64E-01</td>
<td>R1.23 mil/yr</td>
<td>R7.47 mil%</td>
</tr>
<tr>
<td>8</td>
<td>Kromellenboog Dam</td>
<td>1.09E-01</td>
<td>R1.39 mil/yr</td>
<td>R12.68 mil%</td>
</tr>
<tr>
<td>9</td>
<td>Albert Falls Dam</td>
<td>5.45E-02</td>
<td>R1.20 mil/yr</td>
<td>R22.00 mil%</td>
</tr>
<tr>
<td>10</td>
<td>Glen Brock Dam</td>
<td>5.49E-01</td>
<td>R1.28 mil/yr</td>
<td>R2.32 mil%</td>
</tr>
<tr>
<td>11</td>
<td>Wentael Dam</td>
<td>6.10E-01</td>
<td>R1.03 mil/yr</td>
<td>R1.69 mil%</td>
</tr>
</tbody>
</table>
needed. Also, the current DWA evaluation does not take the cost of rehabilitation works into account in any way, and so could be improved.

**SINGLE-EVALUATION CRITERIA FOR EVALUATING DAMS FOR REHABILITATION**

It could be argued that a three-phase approach would be the best, where the acceptability of risk to life is first evaluated using SWTP, followed by economic optimisation as possible motivation to rehabilitate, and finally incorporating environmental, socio-economic, social and risk level considerations into the evaluation. The first two tools do, however, require fairly involved estimations of the expected failure cost and the investment cost for rehabilitation works.

It is proposed to replace the first two steps with a single-evaluation criterion, which accounts for both considerations and would be more convenient and easy to use. For this purpose FN-criteria are developed primarily to evaluate risk to life, but which implicitly incorporate the economic efficiency of rehabilitation works. The FN-criteria are similar to ANCOLD criteria, but instead of using descriptive differentiation as in the case of ANCOLD ("new" vs "existing" dams), the ratio of the investment costs for rehabilitation to the reduction in the probability of failure \( C/\Delta P_f \) is used as an efficiency measure on which stringency levels for safety are based. If a large reduction in the probability of failure \( \Delta P_f \) can be achieved at a small cost, it is very efficient and reasonably practicable to implement more stringent safety criteria for these dams. On the other hand, if only a small \( \Delta P_f \) is achieved at a large cost, it might not be reasonably practicable to rehabilitate, and less stringent criteria should apply to such cases. FN-criteria with different levels of stringency were developed for "small", "medium" and "large" efficiency ratios \( C/\Delta P_f \).

The eleven case studies considered in this study were used to define practical ranges for the efficiency ratios. The \( C/\Delta P_f \) ratio is computed, as shown in Table 5.

To differentiate between the efficiency of rehabilitation works, the \( C/\Delta P_f \) ratios are divided into intervals to obtain practical levels for what can be considered as a "small", "medium" and "large" ratio:

- Small: R1 mil%/ \( <C/\Delta P_f < R10 \text{ mil}/% \)
- Medium: R10 mil%/ \( <C/\Delta P_f < R100 \text{ mil}/% \)
- Large: \( C/\Delta P_f > R100 \text{ mil}/% \)

The "small" efficiency ratio was defined to coincide with the ANCOLD criterion line for new dams, the "medium" efficiency ratio with the ANCOLD criterion line for existing dams, and an additional "large" efficiency ratio criterion line, one multiple of 10 less stringent than the previous two were defined.

The risk to life for cases studied was evaluated in terms of these newly developed criteria as shown in Figure 12. For simplicity the average value of the \( P_f \) and the LOL was used to define the risk to life as a singular point instead of a block (with "S" for "small", "M" for "medium" and "L" for "large" efficiency ratios).

Rehabilitation works were required for all the cases, corresponding to the original DWA decision. The case studies within the "medium" or "large" intervals were, however, located closer to, or on the border of, acceptability of their criteria lines.

Based on the SWTP and economic optimisation outcomes, some dams should not have been rehabilitated. This could suggest an argument for moving the criteria lines to an even less stringent safety level, but for this to be properly motivated more case studies should be considered.

The proposed \( C/\Delta P_f \) criteria could be used as a first step to evaluate South African dams for rehabilitation. It is easy and convenient to use. The rehabilitation decision could then be tested by applying the SWTP and economic optimisation criteria, but these methods require more involved estimations of rehabilitation and failure costs.

It should be noted that the \( C/\Delta P_f \) criteria are by no means perfect. The benefit of an investment in rehabilitation works only considers a reduction in the probability of failure, while economic optimisation additionally considers the costs of dam failure.

The criteria do not consider other factors, such as the socio-economic, social and environmental impacts and the risk level, any of which could require dam rehabilitation works. These factors should be considered separately and require expertise in different areas.

**SUMMARY AND CONCLUSIONS**

In summary, when evaluating South African dams for rehabilitation works, the diagram shown in Figure 13 could be considered. The DWA-estimated risks, as the combined probability and consequences of dam failure, could be evaluated against the FN-criteria proposed in this study, as shown in Figure 12, which primarily evaluates risk to life, but implicitly incorporates a measure of economic efficiency. It is well-aligned with ANCOLD criteria that are based on good engineering practice and judgement. In addition, the risks should be evaluated against the existing DWA multiple acceptability criteria for economic, environmental, social and socio-economic impacts of dam failure and the risk level of dams. If a more refined analysis is required, the risks, together with a detailed estimate of the cost of proposed rehabilitation works, could be used to re-evaluate the rehabilitation decision in terms of the SWTP criteria, economic optimisation and the DWA multiple acceptability criteria.

The criteria developed in this study do not serve as absolute criteria, but are
guidelines which should be considered in conjunction with engineering expert judgement. They could serve as useful tools for validating and prioritising dam rehabilitation. Through this the DWA could make informed decisions and efficiently allocate financial resources to the improvement of dam safety in South Africa.

REFERENCES

ANCOLD (Australian Committee on Large Dams) 2003. Guidelines on risk assessment. Hobart, Tasmania, Australia: ANCOLD.


DWA (Department of Water Affairs, Pretoria, South Africa) 2006–2008. DWA dam design reports for dam safety rehabilitation:
- Rix, A et al 2006 (Vaalkop Dam)
- Van Wyk, W et al 2006 (Makotswane Dam)
- Cameron-Ellis, D G 2007 (Bospoort Dam)
- Pienaar, R A & Badenhorst, D B 2007 (Toleni Dam)
- Van Wyk, W & Badenhorst, D B 2007 (Wentzel Dam)
- Badenhorst, D B & Rix, A P 2008 (Lakeside Dam)
- Badenhorst, D B & Trumpelmann, M 2008 (Kromellenboog Dam)
- Badenhorst, D B & Van Wyk, W 2008 (Albert Falls Dam)
- Van Wyk, W et al 2008 (Klein Maricoopoa Dam)
- Van Wyk, W et al 2008 (Rust De Winter Dam)
- Chaloner, A 2009 (Glen Brock Dam).

DWA (Department of Water Affairs, Pretoria, South Africa) 1994–2006. DWA dam safety inspection reports:
- Hattingh, L C 1994 (Wentzel Dam)
- Nightingale, P A 1994 (Rust De Winter Dam)
- Hattingh, L C 1996 (Albert Falls Dam)
- Oosthuizen, C 1999 (Lakeside Dam)
- Muller, H 2000 (Toleni Dam)
- Slabbert, P 1 J 2000 (Vaalkop Dam)
- De Lange, F J 2002 (Wentzel Dam)
- Coetzee, C J 2003 (Rust De Winter Dam)
- Nightingale, P A 2004 (Albert Falls Dam)
- Hattingh, L C 2005 (Bospoort Dam)
- Kelefeetse, S E 2005 (Klein Maricoopoa Dam)
- Naidoo, R 2005 (Makotswane “Buffeldoorn” Dam)
- Nightingale, P A 2005 (Vaalkop Dam)
- Segers, I 2005 (Kromellenboog Dam)
- Van Vuuren, A 2005 (Lakeside Dam)
- Brink, J 2006 (Glen Brock Dam).


Effects of safety factors on the deflections in a concrete gravity dam

M Opan

The aim of this study was to determine the effects of safety factors on the deflections of a concrete gravity dam. Overturning and sliding safety factors for a selected concrete gravity dam with base width b and height H were specified, using pseudo analysis for the b/H ratios and earthquake acceleration values between 0.1 g and 0.4 g. Deflection values for specified parameters were obtained from the structural analysis program SAP2000. Deflection safety factor curves, determined from the b/H ratios, were obtained, as well as the earthquake acceleration values. The results of this analysis showed that safety factors reduced while strain values increased.

INTRODUCTION

Rapid developments in computer technology over the last few decades have made it possible to apply advanced structural analysis techniques to predict the behaviour, and therefore the safety, of concrete dams during earthquakes. However, many uncertainties remain in predicting the behaviour of dams during severe earthquakes. Amongst these are crack formation and propagation in the concrete body of the dam, nonlinear behaviour of the foundations, and the interaction of the water body in the reservoir with the dam structure (Ghrib et al 1997; Chopra 1998; Kreuzer 2000). Static and dynamic approaches, and linear or nonlinear finite element models have been combined with statistical methods in studies of stability analysis to evaluate the safety of dams (Hall 1998; Tinawi et al 2000; Leclerc et al 2003). A further issue is determining the most appropriate definition for seismic input, as this has a very significant effect on the seismic design and seismic safety evaluation of dams. Many uncertainties still exist in this regard, but progress has been made in defining the maximum credible earthquake (MCE), maximum design earthquake (MDE) and safety evaluation earthquake (SEE). With more seismic records available now, the effects of the very short duration peak ground acceleration (PGA) (which may be from 0.5 g to 1.0 g, or even higher in severe earthquakes), and the sustained effective seismic acceleration (which could be about 0.5 – 0.67 of PGA), are now better understood. An indirect result of seismic activity is the effect of increased seepage on uplift forces, due to increasing pressure acting through cracks in the concrete body or under the foundations. This must be taken into account in the stability analysis for seismic design, as it may have a very significant effect on the safety of the dam. Opan and Temiz (2007) reported that safety factors decrease with earthquake acceleration and pressure reduction.

The aim of this study was to determine the relationship between the deflections and safety factors at topping, and at toe, for a concrete gravity dam (CGD) during an earthquake. A CGD with base width b and height H was specified, using pseudo analysis for the b/H ratio and earthquake acceleration values between 0.1 g and 0.4 g. Deflection values for specified parameters were obtained from the structural analysis program SAP2000. Graphs were plotted for the selected dam from the results of the b/H ratio and earthquake acceleration, and the deflection and safety factors were evaluated.

FORCES ON A CONCRETE GRAVITY DAM DURING AN EARTHQUAKE

The forces on a CGD under earthquake conditions are the horizontal hydrostatic load, the horizontal hydrodynamic load, the vertical pore water pressure force, the weight of the CGD and the earthquake force due to the CGD weight. These forces are shown in Figure 1, where $F_E$ is the horizontal hydrodynamic load, $F_h$ is the horizontal hydrostatic load, $W_U$ is the vertical pore water pressure load, $W$ is the weight of the concrete dam, $W_E$ is the concrete dam weight and earthquake load on the horizontal direction, and $b$, $c$, $h$ and $H$ are the parameters of the concrete dam.

The horizontal hydrostatic load is:

$$F_h = \frac{1}{2} \gamma_W b h^2 \quad (1)$$

The horizontal hydrodynamic load is:

$$F_E = a \alpha \gamma_W b h^2 \quad (2)$$

Keywords: concrete gravity dam, pseudo analysis, deflections, safety factors
where \( a \) is the reduction factor (\( a \) is between 0.543 and 0.555) for the earthquake force (Westergaard 1933).

Vertical pore water pressure force is:

\[
F_{U} = \frac{1}{2} \gamma_{w} bh(1 + a) \tag{3}
\]

The weight of the concrete dam wall is simplified by assuming that the cross-section of the wall is a triangle with base \( b \) and height \( h \). In other words \( H = h \). The weight of the concrete gravity dam is:

\[
W = \frac{1}{2} \gamma_{c} bh \tag{4}
\]

Earthquake force due to concrete gravity dam weight is:

\[
WE = \frac{1}{2} \alpha \gamma_{c} bh \tag{5}
\]

where \( \gamma_{c} \) and \( \gamma_{w} \) are the specific weights of concrete and water respectively and \( m \) is the pressure reduction factor for the water pressure force (0 ≤ \( m \) ≤ 1). The ratio \( \alpha \) is the earthquake acceleration divided by gravitational acceleration.

**PSEUDO ANALYSIS**

Pseudo analysis for dams is based on dam stability. The stability of a CGD is provided by its moments. If the moment that attempts to tip the dam at the toe point is lower than the resisting moment, the dam remains stable. This is demonstrated in Equation 6:

\[
W \cdot \frac{b^{2}}{3} \geq FX \cdot \frac{h^{3}}{3} + FU \cdot \frac{2b^{3}}{3} + WE \cdot \frac{h}{3} \tag{6}
\]

With the substitution of Equations 1 to 5, this is equivalent to:

\[
\left[ 2 + \frac{\alpha \gamma_{c}}{\gamma_{w}} \right] - 2 \gamma_{w}(1 + \alpha) \frac{b^{2}}{h^{2}} \geq (1 + 1.41 \alpha) \tag{7}
\]

After simplification, it becomes:

\[
\frac{b}{h} \geq \frac{(1 + 1.41 \alpha)}{(2 + \alpha) \frac{\gamma_{c}}{\gamma_{w}} - 2(1 + \alpha)m} \frac{1}{2} \tag{8}
\]

Equation 8 provides a relationship between the \( b/h \) ratio, and \( m \) and \( \alpha \) are obtained. This relationship is shown in Figure 2 where \( \frac{\gamma_{c}}{\gamma_{w}} \) is 2.5.

The overturning safety factor is obtained from the ratio of resist moment to subvert moment. According to this definition, the overturning safety factor is obtained below:

\[
\epsilon_{d} = \frac{W \cdot \frac{2b}{3}}{FX \cdot \frac{h}{3} + FU \cdot \frac{2b^{3}}{3} + WE \cdot \frac{h}{3}} \tag{9}
\]

After substitution and rearrangement, it becomes:

\[
\epsilon_{d} = \frac{2 \cdot \gamma_{c} \left( \frac{b}{h} \right)^{2}}{\left( 1 + 1.41 \alpha \right) + \left( \frac{\gamma_{c}}{\gamma_{w}} \right) \alpha} \tag{10}
\]

where \( \frac{\gamma_{c}}{\gamma_{w}} \) is accepted as 2.5. The ratio \( b/h \) is found versus \( m \) and \( \alpha \), changes of \( \epsilon_{d} \) connected with \( m \) and \( \alpha \) are obtained, and \( \epsilon_{d} \) decreases with increasing values of \( m \) and \( \alpha \).
The sliding safety factor is attained from a ratio of resist force to sliding force. According to this definition, the sliding safety factor is obtained as follows:

$$\varepsilon_k = \frac{(W - F_U)k}{F_X + F_Y + WE}$$  \hspace{1cm} (11)

With the rearrangement of this equation, Equation 12 is formed:

$$\varepsilon_k = \frac{\gamma_C \gamma_W \gammaW - m(1 + \alpha)}{(1 + 1.11\alpha) + \gammaC \gammaW \alpha \cdot \frac{b}{h}}$$  \hspace{1cm} (12)

where $\gammaC = 2.5$ and $k = 1.0$ are accepted, and $b/h$ is obtained versus $m$ and $\alpha$ in Equation 8.

Changes of $\varepsilon_d$ and $\varepsilon_k$ related to $m$ and $\alpha$ are calculated as shown in Figures 3 and 4, where it can be seen that $\varepsilon_d$ and $\varepsilon_k$ decrease with increasing values of $m$ and $\alpha$.

**STRUCTURAL ANALYSIS USING SAP2000**

A CGD is modelled by using the shells of the finite element in the SAP2000 program. The principal axis of the dam in this program is shown in Figure 5. Shells are separated into $0.50 \times 0.50$ meshes, as shown in Figure 6. Small dimensions are chosen for these meshes.
The effects of safety factors on the deflections in a concrete gravity dam during an earthquake were determined in this study. A concrete gravity dam with a base width $b$ and height $H$ was specified, using pseudo analysis for the $b/H$ ratios and earthquake acceleration values between 0.1 g and 0.4 g. Values for specified parameters were obtained from the structural analysis program SAP2000. Graphs were plotted for the selected dam from the results of the $b/H$ ratio and earthquake acceleration, and the deflection and safety factors were evaluated.

\[ \gamma_C = 2.5 \text{ and } k = 1.0 \text{ were accepted in the pseudo analysis.} \]

In Figure 2 it can be seen that the $b/h$ ratio increased with an increase in $m$ and $\alpha$. In Figures 3 and 4 the safety factors decreased with increasing $m$ and $\alpha$. Deflection of the topping point related to $m$ and $\alpha$ at $b/H = 0.6$, 0.8 and 1.0 are shown in Figures 7, 8 and 9 respectively. It can therefore be said that deflection increased with increasing $m$ and $\alpha$. Deflection of the toe point related to $m$ and $\alpha$ at $b/H = 0.6$, 0.8 and 1.0 are shown in Figures 10, 11 and 12 respectively. Deflections increased with increasing $m$ and, and deflections decreased with an increase in the $b/H$ ratio.

**CONCLUSIONS**

In this study, tipping and sliding safety factors of a dam during an earthquake, and deflections at toe and topping points, were determined. Pseudo analysis was performed to define safety factors. Changes to these safety factors versus $m$ and $\alpha$ were investigated, and it was seen that safety factors decreased with increasing $m$ and $\alpha$. SAP2000 was used to determine deflections at the toe and topping points. According to the changes in these deflections versus $m$ and $\alpha$ it can be said that deflection increased with increasing $m$ and $\alpha$. In addition, safety factors increased with increasing $b/H$, while deflection decreased with increasing $b/H$. Further research should cover analysis and studies regarding crack formation and dispersion in a concrete gravity dam during an earthquake.

**REFERENCES**


INTRODUCTION

Collapsible soil was first recognised and studied by Jennings and Knight (1957), and Knight (1961). Soil collapse is typically evaluated in terms of total stress, as shown in Figure 1, where an unsaturated soil sample is placed in an oedometer at natural moisture content (A), loaded in stages to the desired vertical total stress (B), before the sample is inundated with water and the vertical strain is recorded. In South Africa, a collapse test inundated at 200 kPa is referred to as a Collapse Potential Test, and the vertical strain (B to C) during inundation is taken as the Collapse Potential. The sample is then loaded to point D and unloaded to point E under saturated or near-saturated conditions. This method is superficially attractive due to its simplicity, and is widely used as an indicator test to allow judgement on the severity of possible foundation settlement due to soil collapse. However, it is well known that soil behaviour depends on its effective stress, and a fundamental interpretation of the collapse phenomenon should be made in terms of effective stresses.

A number of theoretical frameworks have been developed to express the effective stress for unsaturated soil. Bishop (1959) formulated the following equation to quantify effective normal stress \( \sigma' \) for unsaturated soil in terms of total stress \( \sigma \), pore air pressure \( u_a \) and pore water pressure \( u_w \):

\[
\sigma' = (\sigma - u_a) + \chi(u_a - u_w)
\]

(1)

\( \sigma - u_d \) is termed the net normal stress and \( u_w - u_a \) is the matric suction. \( \chi \) is a parameter that quantifies the contribution of the matric suction to the effective normal stress.

Keywords: creep, Collapse Potential Test, double oedometer test, effective stress, matric suction, soil collapse, unsaturated soil, vertical strain

Figure 1 Collapse Potential Test
stress. $\chi$ varies between 0 for a completely dry soil and 1 for a completely saturated soil. For these two conditions Equation 1 reduces to the classic Terzaghi equation ($\sigma' = \sigma - u_s$) for effective normal stress of a soil where the pores are completely filled either with air or water.

The difficulty with applying Equation 1 is quantifying the parameter $\chi$. It is determined experimentally by conducting either shear tests or volumetric strain tests and quantifying the contribution of the matric suction to the behaviour of the soil. Bishop (1959) suggested that $\chi$ is a function of the degree of saturation ($S_r$), soil structure and whether the soil is in a wetting or drying phase. Alonso et al (2010) suggested that $\chi$ also depends on the pore size distribution.

The relationship $\chi = S_r^x$ has been found to give good results with $x$ depending on the type of soil. Vanapalli and Fredlund (2000) developed the relationship for $\chi$ shown in Figure 2 by conducting unsaturated triaxial tests. Data from Alonso and Romero (2011) from volumetric strain tests are also shown in Figure 2.

Rust et al (2005) proposed a conceptual yield model to explain soil collapse, using effective stress paths as shown in Figure 3 and briefly explained below.

Consider an in situ collapsible soil with stress condition A in the figure. The stress condition is inside the “dry” yield surface. If wetted, the path will move to point A’, and the in situ stress state will still be inside the now smaller “wet” yield surface for the in situ soil, which is now weaker than before it had been wetted. But since the stress state stayed inside the shrinking yield surface during wetting, no collapse of the in situ soil profile occurred. If the in situ soil at point A is loaded, the stress state may move to point B. Alternatively, if the soil is first sampled, it will move to point C where it does not experience any shear stresses ($q = 0$), before it is loaded in the oedometer to point B. At point B the soil is still at its in situ moisture content and hence still inside the “dry” yield surface. If the soil is now wetted, the yield surface will shrink as the soil weakens, and the stress state will move toward B’ where the yield surface and stress state coincide with the one-dimensional strain ($K_o$-line), and large strains will occur as the soil collapses.

Even though this proposed framework has not been verified experimentally, primarily due to experimental difficulties, it is useful since it attempts to describe the collapse phenomenon in terms of effective stresses. It also shows that, from a fundamental effective stress perspective, the behaviour is much more complex than the behaviour shown in terms of total stresses in Figure 1.

**SAMPLING AND TESTING**

According to Brink et al (1982) and Schwartz and Yates (1980), the typical dry densities that could be expected for collapsible soils in South Africa range from 900 to 1 600 kg/m$^3$, although Rust et al 2005 argued that collapse in materials with dry densities outside this range cannot be excluded. According to Schwartz (1985), it was originally assumed that the collapse phenomenon was largely restricted to loose aeolian deposits, with the result that most of the research and work dealt almost exclusively with such deposits. However, in South Africa numerous cases of collapse have been reported for the residual granitic soils of the Basement Complex (Schwartz 1985). Extensive foundation problems have been encountered with these soils, both in the central Highveld and eastern Lowveld regions of South Africa. The collapsible nature of the residual soils derived from these granites is associated with the deeply weathered soil profiles above the African Erosion Surface, encountered in the humid eastern parts of South Africa, with annual water surplus. During the weathering process, quartz remains unaltered in the form of sand grains, and mica particles remain partially unaltered, but the feldspars are thoroughly kaolinised through the chemical interaction with water charged with carbon dioxide, forming what
is commonly known as ferralitic soils (Brink 1979). As the name suggests, ferralitic soils are rich in iron and aluminium hydroxides. Due to the duration of the weathering cycle and the moist nature of the climate, these ferralitic soils have a very low base status and are characterised by a friable and porous structure and the presence of 1:1 lattice kaolinitic clays (Brink 1979).

An undisturbed sample of decomposed residual granite was collected from the Bushbuck Ridge area, which is part of the Nelspruit Granite Suite. The material in its natural state was described as being slightly moist, dark pink with orange stains, gravelly silty sand with a medium dense to dense consistency and pinholed structure. It had a dry density of 1.392 kg/m³, natural moisture content of 7.2% and specific gravity of 2.64.

The particle distribution consisted of 11% clay, 20% silt, 57% sand and 12% gravel-sized particles. The liquid limit was 36% and the plasticity index 13%. Quantitative XRD analysis indicated that more than 50% of the mineral composition was made up of quartz and kaolinite, with secondary percentages of microcline, muscovite and albite.

A modified oedometer was used to observe the collapse behaviour of the residual granite in the laboratory. It is similar to a standard oedometer apparatus, but allows the incremental addition of water to the sample through a small hole at the bottom of the loading cell during the test. The undisturbed sample was cut to fit into the oedometer ring with an inside diameter of 75 mm and a height of 20 mm. Once the sample had been cut into the ring, the sample surfaces at the top and bottom were carefully trimmed in order to minimise the effect of bedding errors during loading. Once the loading cap was in position the entire cell was sealed using petroleum jelly in order to minimise moisture loss. The cell was then connected to a water-filled burette using a small-diameter rubber tube to allow for the introduction of known quantities of water to the soil specimen. Since the gauge pore air pressure (uₚ) was zero throughout the test (equal to atmospheric pressure), the net vertical stress was equal to the total stress and the applied vertical stress.

At natural moisture content, the sample was loaded in increments, up to a total stress of 200 kPa. Dial gauge readings were taken and the load was only increased once the rate of creep of the sample was equal to or less than 0.25%/hour (Heymann 2000). Up to this point the testing procedure is similar to that of the widely used Collapse Potential Test. Once creep of the sample at a total vertical stress of 200 kPa had subsided sufficiently, additional moisture was introduced for the first time. Moisture was added in increments from the burette and the amount of water entering the sample was recorded. The process was continued until the sample no longer accepted free water. The sample was then inundated and left overnight. The next morning the loads were increased up to the final maximum load.

The moisture retention curve was determined using the filter paper method (Chandler & Gutierrez 1986) and the calibration relationship proposed by Hamblin (1981) for Whatman no 42 filter paper.

**DISCUSSION**

Figure 4 shows the behaviour of the soil in terms of total vertical stress (numerically equal to the applied vertical stress) and the effective vertical stress (Equation 1). For the purpose of these calculations χ was taken as equal to Sₑ with κ = 1.94 at a plasticity index of 13% (Figure 2). The relationship between total stress and void ratio (e) is equivalent to that of a conventional Collapse Potential Test, except that the volumetric strain occurred incrementally as water was introduced to the specimen in stages. The Collapse Potential of 5.6% is taken as the vertical strain during inundation at a total stress of 200 kPa (point A to B in Figure 4).

At the start of the test the soil was at its in situ moisture content of 7.2% and it experienced a matric suction of 2.238 kPa. This translates to a 110 kPa difference between effective and total stress vertical at a degree of saturation of 21%. As the loading was increased prior to introducing water, the increase in effective vertical stress was different to the increase in total vertical stress. This is due to the fact that the moisture content remained constant, but the sample volume decreased. Hence the degree of saturation (Sₑ) increased, changing χ. This may be an oversimplification of a more complex phenomenon. When water was introduced to the specimen (points A and C in Figure 4), the effective vertical stress reduced rapidly and the void ratio decreased. This behaviour is in contrast with classical soil mechanics where a reduction in void ratio requires an increase in effective stress. As more water was introduced, the effective vertical stress eventually tended towards the total vertical stress as the matric suction approached zero. The collapse process appears to consist of two mechanisms. The first part is from point C to D in Figure 4, with a rapid reduction in effective stress which coincides with a reduction in void ratio. For the second part (D to B) sufficient water has been introduced to the specimen for the matric suction to approach zero, but void ratio reduction continues. This suggests that the introduction of water triggered the collapse process by reducing the effective stress, but that the reduction in effective stress cannot fully account for the reduction in the void ratio, as much of the void ratio reduction occurred at constant effective stress. Clearly creep is also an important mechanism contributing to volume reduction during soil collapse. Subsequent load increments, after inundation, occurred at near full saturation, and the behaviour was similar to a conventional one-dimensional oedometer test.
CONCLUSIONS
The data shown in this article indicates that interpretation of soil collapse in terms of total stress leads to an oversimplification of the behaviour of the soil. This masks the contribution of the change in effective stress. The results indicate that the reduction in effective stress plays an important role during soil collapse. It appears to trigger the collapse process, but much of the collapse settlement may also be due to creep where volume reduction occurs at constant effective stress.

Effective stress due to matric suction may be higher than the stress applied to the soil as a result of operations related to construction of infrastructure. It follows that changes in moisture content will have a greater effect on the effective stress, and hence the strength and stiffness of the material, than stress changes due to engineering activity. Engineers therefore need to consider not only the effect of loading the soil due to construction, but also the consequence of moisture change on the effective stress experienced by the soil.

REFERENCES
Parametric study on the behaviour of Y-shaped composite bridges

P Lu, C Shao

A composite Y-shaped bridge was analysed as a three-dimensional structure, using commercially available software. A finite element model to simulate the structural behaviour of the bridge was used to study its static and dynamic behaviour in order to gain a better understanding of the interaction between the main bridge and the ramp. An extensive parametric study was conducted, in which a typical composite Y-shaped bridge model was analysed and tested. The key parameters considered in this study were: loading conditions, ramp radius of curvature, ramp longitudinal slopes, bifurcate diaphragm plate stiffness, and loading cases. The structural behaviour of composite bridges is discussed in detail. The results from this study could enable bridge engineers to design complex composite Y-shaped bridges more reliably and economically.

INTRODUCTION

Composite Y-shaped bridges have become an important component of highway systems, especially in urban grade separation bridges where multilevel interchange structures are necessary. As the curvature of highways increase due to space limitations, such bridges become more suitable. A typical bridge consists of a main bridge composite with a ramp or multi-ramps. The cross-section form of a composite bridge structure mainly adopts the box girder structure, which is favoured for its considerable torsional stiffness to resist applied loads. With the increase in the construction of such bridges, the analysis of the bridge structure is becoming more complicated, especially with regard to restrained torsion, distortion angle, distortion warp and the shear lag effect. Various scholars have studied the design and structural analysis methods of the composite bridge structure (Sennah et al 2004; Samaan 2004; Cheung & Megnouni 1991; Hamblky 1982; Morreua et al 1996; Mitra et al 2004; Vo & Lee 2007).

The typical composite structure is one of the most important parts of urban grade separation bridge structure. As shown in Figure 1, the typical composite bridge structure includes the main bridge, the curved ramp girder and crotches of irregular composite (Y-shaped bridge). Due to the interaction between the main bridge and the ramp, the mechanical behaviour of the composite structure is very complex and difficult to analyse in practical design. Kim and Kim (2003) analysed the thin-walled curved box beam under in-plane flexure. Analytical studies on cross-sectional deformations of curved box girders have been reported by several researchers. These studies are based on the thin-walled beam theory, the folded-plate theory, the finite-strip method and the block-finite-element method. Huang (2008) presented a procedure for obtaining the dynamic response and basic impact characteristics of thin-walled curved box girder bridges under truck loading. Dong & Sause (2010) presented a finite element method for considering material in-elasticity, second-order effects, initial geometric imperfections, and residual stresses of curved tubular flange girders. More recently, Okeil and El-Tawil (2004) reported on the effects of warping stress levels in 18 curved box girder bridges (adapted from the Florida Department of Transportation inventory) and pointed out that additional work was needed to define relevant parameters which could be used to identify bridges where warping calculations are warranted. In addition there are several theoretical and experimental studies on single curved girders (Thevendran et al 2000; Shannugam et al 1995; Zureick et al 2000; Sennah & Kennedy 2001).

The main bridge member connects the curved girder of the ramp with crotches of irregular composite. The irregular crotches...
are used to resist interaction between the main bridge and the ramp members. This typical composite bridge structure was largely used in urban grade separation bridges in China. In recent years, with the increasing use of composite Y-shaped bridges, a number of Chinese urban grade separation bridges have collapsed, including grade separation bridges in Baotou (2007), Zhejiang (2008) and Tianjin (2009) (Figure 2).

In order to enhance the understanding of this area, and to calibrate the analytical model developed by the authors, full-scale testing and computer simulations were carried out on a curved concrete box girder bridge as suggested by Senthilvasan et al. (2002). Heins and Sahin (1979) conducted static and dynamic tests on a steel two-span curved continuous box girder bridge in Seoul, South Korea, in order to calibrate several analytical procedures. Very few investigations have been conducted on composite Y-shaped bridges. A parametric study on the impact factors for 180 curved continuous multiple-box girder bridges was conducted by Samaan et al. (2007). Lu et al. (2007) presented the thin-walled beam theory, combined with finite element techniques, to provide a new thin-walled box beam element in the Y-shaped bridge. In an attempt to reduce collapse of this type of bridge structure, this paper investigates the behaviour of a typical Y-shaped composite bridge through experimental and analytical studies. The interaction between the main bridge and the ramp, and the dynamic performance of this combination, should always be considered in the design of a composite bridge. The design of Y-shaped bridges usually adopts thin-walled box girder sections, which affect both the constrained torsion and distortion of the structure. However, the mutual connection and restriction characteristics of the two bifurcate parts of the bridge structure under static and dynamic load cause constrained torsion and distortion, which could result in bridge failure. It is therefore necessary to study the mechanical behaviour of the Y-shaped bridge so that the study results can contribute towards the modelling of this interaction, and can be incorporated into bridge design.

The mechanical behaviour of the Y-shaped bridge structure is complicated, particularly with regard to strained torsion and distortion, angle distortion, warp and shearing lag effect. The three different methods of analysing bridge structures in the past were mainly as follows: Firstly, the Y-shaped bridge structure was generally divided into two structures, which were then analysed independently. However, this method does not take the interaction between main bridges and ramps into account. Secondly, the analysis of the Y-shaped bridge structure was carried out according to the elementary beam theory, but this method does not take the restrained torsion, the distortion angle, the distortion warp, and the shearing lag effect into consideration. Thirdly, the analysis of Y-shaped bridge structure adopted the shell theory, but the direct internal forces cannot be obtained with this method, and the post-process workload is heavy. The above-mentioned methods therefore cannot reflect the structural mechanical behaviour of the bridge, and may in fact contribute to the lack of safety of the structural design. To counter the problems mentioned above, and in order to enhance user safety of the
Y-shaped bridge, two design methods are normally used in actual practical engineering, as illustrated in Figures 3(a) and (b).

The aim of this study is to conduct a parametric study in three dimensions to examine the effect of key parameters and loading conditions on composite Y-shaped bridges that might influence the impact factors of such bridges. The data generated from the study is used to deduce empirical expressions for impact factors relating to basic mechanical bridge behaviour.

**EXPERIMENTAL SPECIMEN AND FINITE ELEMENT MODEL**

**Description of bridge prototypes and test modelling**

In this study, a typical composite Y-shaped bridge, from existing common urban grade separation bridges, was analysed (see Figure 4(a)), and then reduced in scale to an organic glass model as the test model (scale 1/30), as illustrated in Figure 4(b). All the members in the specimen were made of organic glass. The material properties of these glass members were determined by tensile coupon tests, as prescribed by the relevant standards, and the elasticity modulus $E_s$ and Poisson’s ratio $\nu$ of the composite bridge are $2.77 \times 10^3$ MPa and 0.367 respectively. As the similarity ratio of stress between the prototype and the model is 5:1, the stresses that result from the model test will be the stresses for the prototype. The specification of a geometrical scale of 1/30 (see Figure 5) between the test model and the prototype in the specimen is given in Table 1. The ranges of the parameters considered in this study were based on an extensive survey of existing composite Y-shaped bridges.

**Finite element modelling**

The three-dimensional finite element model used in this study was verified by experimental results. The bridge models were instrumented to measure deflections, strains and support reactions. The experimental test results for longitudinal strains, deflections and support reactions were compared to the theoretical values derived from the same finite element model used herein.

To extend the interpretation of the results and observations obtained in the test, and to gain a better understanding of the behaviour of a composite bridge, a parametric study on the composite Y-shaped bridge specimen was carried out using a finite element analysis (FEA) program. The three-dimensional FEA model is shown in Figure 6. Figure 6(a) was modelled with shell elements, Figure 6(b) was modelled with beam elements, and for

---

**Table 1** The specification of geometrical similarity ratios between test model and prototype/mm

<table>
<thead>
<tr>
<th></th>
<th>Top</th>
<th>Side web</th>
<th>Middle web</th>
<th>Bottom</th>
<th>Diaphragm</th>
<th>Cross beam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prototype</td>
<td>220</td>
<td>350</td>
<td>700</td>
<td>200</td>
<td>600</td>
<td>1 200</td>
</tr>
<tr>
<td>Model</td>
<td>8</td>
<td>12</td>
<td>24</td>
<td>8</td>
<td>20</td>
<td>40</td>
</tr>
</tbody>
</table>

---

**Figure 4(a) Typical composite bridge (dimensions in cm)**

**Figure 4(b) Model of typical composite bridge (dimensions in mm)**

**Figure 5 Scaled model tests of a typical composite bridge**

Figure 6(c) the grillage analysis method was chosen to simulate the bridge.

The analysed Y-shaped bridge, with three-span main and two-span ramp, is a typical Y-bridge structure loaded at middle span on the upper surface of the concrete. The loaded area is 10 x 10 mm. The study uses the FE program to simulate the mechanical behaviour of
the composite bridge, and establish a three-dimensional FE model (space grillage model and shell model) to analyse the behaviour of the Y-shaped bridge. In the shell model (commercial software), the four-node shell elements (shell63) were employed to model the concrete slab. The shell63 has both bending and membrane capabilities. The element has six degrees of freedom at each node, and the element is defined by four nodes. According to the bridge structure size and unit grid control precision, there are 397 shell elements in the shell model. In addition, in the space grillage model (numerical program considering the effect of torsion and distortion for new beam element) space grillages were modelled by two-node three-dimensional beam elements (considering the effect of torsion and distortion for a new beam element). In the space grillage model, the two-node primary new beam elements were employed to model the box beam. The new beam element has eight degrees of freedom at each node, and the element is defined by two nodes. According to the bridge structure size and unit grid control precision, there are 422 new beam elements in the grillage model. To study the effect of boundary conditions on the performance of the structure, three kinds of different boundary conditions are used in the model test and the analysis of the Y-shaped bridge respectively, as follows: (1) the fixed double support was set only in the bifurcation of the Y-shaped bridge, while the other places were set up as vertical single loading; (2) the fixed double support was set both in the bifurcation of the Y-shape bridge and at the end of the beam, while other places were set up as vertical single loading; (3) on the basis of two kinds of constraint conditions, the vertical single loading between the two curved ramps was set to vertical eccentric loading.

In this study the corresponding method of grillage analysis for the bridge at different boundary conditions was used by equivalent stiffness. The equivalent stiffness of the grillage analysis model was presented from the grillage analytical method obtained by equivalent stiffness principles. The change in different grillage stiffness for the bridge during the analysis was taken into account. The finite element model did not include the influence of cracks in the organic glass model. In this study the numerical simulations (shell model and grillage model) were carried out based on the above FE model, and numerical analysis results were calibrated against the corresponding experimental data. The shell element model cannot determine internal forces directly, and the post-process work load is heavy. The analytical method of grillage by essentially increasing the degree of freedom is an effective method to predict properties of this class of bridge. To simulate the experiments, the same loading conditions and constraints as used in the experiments were used in the finite element analysis.

**PARAMETRIC STUDY OF STATIC BEHAVIOUR**

The testing equipment of the Bridge Engineering Department of Guangzhou University’s research centre was used to test the model bridge under six different load conditions.
cases. The axle loading corresponding to the number of steel blocks loaded are presented in Table 2. In order to detect bending behaviour, six static loading positions were determined to be the most unfavourable (stress-inducing) loading positions corresponding to each related span and girder. The detailed longitudinal and transverse loading positions are shown in Figure 7. Measurements of strain in the control section were conducted by means of electric resistance gauges at sections I–I, II–II and III–III. Deflections and transverse deformations were measured by means of deformation gauges at the control section. In the static and dynamic model test, the instruments and equipment used included a static data acquisition instrument (TDS-303 with a range of ± 16 000 με), a vibration testing system (DASP with a range of 0–20 kHz), and Analysis software.

Each control section carried six loading cases for the static bending test. Initial readings of all gauges were recorded with no loads on the model bridge. The loads were then applied and positioned at the predefined critical load positions on the model bridge. Strain and deflection readings were recorded for each loading case. After each loading case, the steel blocks were removed from the model bridge and another zero load reading was taken. The measured data was immediately displayed and compared with the analytical results to reveal any anomalies.

**Influence of loading conditions on girder stresses**

Bridge-loading conditions are normally considered to be major criteria affecting the girder stress of bridges, especially in Y-shaped composite bridges. Results obtained from the model composite bridge are shown in Tables 3 – 6. It was observed that girder stresses appear to change considerably with different loading conditions. For loading I conditions, the values of the girder stresses were generally greater than for loading II and loading III, especially in the ramp curve control section, and the analysis shows that loading II is close to loading III. Results furthermore showed the use of the powerful cross beam in the bifurcation of the Y-bridge. While the fixed double bearings support the bifurcation transverse beam, the powerful cross beam is used to reduce the interaction effect between the main bridge and the ramp. The fixed double bearing enhances the torsional performance of the cross beam, which reduces the huge positive moment and distortional warping bi-moment. When under outside load, the curved ramp of the composite bridge will lead to inside unloading, that is, increased outside deflection with gradually decreasing inside deflection. By comparing the different loading results calculated from the finite element analytical program, it is shown that eccentric loading produces reasonable internal forces, while the torsional double-loading of the end-diaphragm sections in a Y-shaped composite bridge is suitable to resist the torsional and warping effects induced by grade separation curvatures.

**Influence of ramp radius of curvature on girder torsion and distortion**

In the space grillage model (and while considering the effects of torsion and distortion), a new beam element was introduced, increasing the degrees of freedom from six per node to eight. Based on the three-dimensional space grillage FEA model of the new beam element, and Saint-Venant’s principle, the relationships between the model composite bridge torsion, distortion and ramp radius of curvature are presented in Figure 8. The graph shows that the ramp radius of curvature has no significant effect on the torsion and distortion effect. However, it is observed that stress intensification factors of the middle-span section appear to increase considerably with decreasing ramp radius.

---

### Table 2 Detailed longitudinal and transverse loading positions

<table>
<thead>
<tr>
<th>Load case</th>
<th>Concentrated load</th>
<th>Loading position</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>P = 150 N</td>
<td>middle span right web of main bridge I</td>
</tr>
<tr>
<td>C2</td>
<td>P = 150 N</td>
<td>middle span right web of main bridge II</td>
</tr>
<tr>
<td>C3</td>
<td>P = 150 N</td>
<td>middle span right web of main bridge III</td>
</tr>
<tr>
<td>C4</td>
<td>P = 150 N</td>
<td>middle span right web of ramp I</td>
</tr>
<tr>
<td>C5</td>
<td>P = 100 N</td>
<td>middle span right web of ramp II</td>
</tr>
<tr>
<td>C6</td>
<td>P = 100 N</td>
<td>middle span left web of ramp II</td>
</tr>
</tbody>
</table>

### Table 3 Comparison between experimental and analysis results

<table>
<thead>
<tr>
<th>Loading cases</th>
<th>Shell element</th>
<th>Beam element</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>strain (με)</td>
<td>deflection (mm)</td>
<td>strain (με)</td>
</tr>
<tr>
<td>C1</td>
<td>237</td>
<td>0.32</td>
<td>175</td>
</tr>
<tr>
<td>C2</td>
<td>240</td>
<td>0.44</td>
<td>185</td>
</tr>
<tr>
<td>C3</td>
<td>270</td>
<td>0.47</td>
<td>234</td>
</tr>
<tr>
<td>C4</td>
<td>247</td>
<td>0.27</td>
<td>141</td>
</tr>
<tr>
<td>C5</td>
<td>139</td>
<td>0.36</td>
<td>156</td>
</tr>
<tr>
<td>C6</td>
<td>158</td>
<td>0.25</td>
<td>154</td>
</tr>
</tbody>
</table>

### Table 4 Comparison between experimental and analysis results under loadings case 1/με

<table>
<thead>
<tr>
<th>Section</th>
<th>Loading position</th>
<th>Loading I</th>
<th>Loading II</th>
<th>Loading III</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>80/73</td>
<td>74/70</td>
<td>84/79</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>11/11</td>
<td>11/10</td>
<td>10/3</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>–10/10</td>
<td>–10/7</td>
<td>–7/5</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>–4/3</td>
<td>–4/3</td>
<td>–3/2</td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>13/22</td>
<td>11/15</td>
<td>11/18</td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>–9/8</td>
<td>–9/7</td>
<td>–8/8</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>–3/2</td>
<td>–2/1</td>
<td>–2/1</td>
<td></td>
</tr>
</tbody>
</table>

### Table 5 Comparison between experimental and analysis results under loadings case 2/με

<table>
<thead>
<tr>
<th>Section</th>
<th>Loading position</th>
<th>Loading I</th>
<th>Loading II</th>
<th>Loading III</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>19/19</td>
<td>19/17</td>
<td>17/26</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>58/58</td>
<td>54/55</td>
<td>63/62</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>72/59</td>
<td>72/69</td>
<td>55/53</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>27/26</td>
<td>27/29</td>
<td>25/22</td>
<td></td>
</tr>
<tr>
<td>S11</td>
<td>–45/49</td>
<td>–41/40</td>
<td>–44/38</td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>23/25</td>
<td>23/22</td>
<td>24/24</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>4/4</td>
<td>4/4</td>
<td>3/4</td>
<td></td>
</tr>
</tbody>
</table>

### Table 6 Comparison between experimental and analysis results under loadings case 3/με

<table>
<thead>
<tr>
<th>Section</th>
<th>Loading position</th>
<th>Loading I</th>
<th>Loading II</th>
<th>Loading III</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>17/15</td>
<td>17/16</td>
<td>18/23</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>28/29</td>
<td>23/24</td>
<td>30/37</td>
<td></td>
</tr>
<tr>
<td>S5</td>
<td>–31/32</td>
<td>–30/43</td>
<td>–27/36</td>
<td></td>
</tr>
<tr>
<td>S7</td>
<td>19/20</td>
<td>19/23</td>
<td>18/23</td>
<td></td>
</tr>
<tr>
<td>S8</td>
<td>8/6</td>
<td>7/7</td>
<td>6/11</td>
<td></td>
</tr>
<tr>
<td>S13</td>
<td>43/47</td>
<td>40/43</td>
<td>40/44</td>
<td></td>
</tr>
<tr>
<td>S15</td>
<td>18/15</td>
<td>17/13</td>
<td>18/13</td>
<td></td>
</tr>
</tbody>
</table>
Influence of ramp span length on girder torsion and distortion

Five different lengths of 15, 20, 25, 30 and 35 m for each span subjected to loading cases 5 and 6 were considered in the parametric study. Based on the three-dimensional FEA model, the distribution of the warping and distortion bi-moments for the main bridge and ramp are presented in Figures 9 and 10. The graphs show that the maximum bi-moment of both torsion warping and distortion warping for the middle-span section remained almost unchanged; that is to say, the ramp span length has no significant effect on the torsion warping and distortion warping effect.

Influence of ramp longitudinal slope on girder stresses

A longitudinal transitional slope is essential between the Y-shaped composite bridge deck.
and the ground traffic. Longitudinal ramp bridge slopes of 3%, 5% and 7% were studied. Tables 7 and 8 show the effect of longitudinal slopes on girder stresses. Figure 11 shows the deformed shapes of different ramp longitudinal slopes for a Y-shaped bridge. It is obvious that the stresses of the control middle section increase with increase in the longitudinal slope, while the stresses of the loading section decrease with increase in the longitudinal slope. For example, the stress in the S7 section decreased by 5.46% when the longitudinal slope increased from 0% to 7%. Similarly, the stress in the middle section increased by almost 14.28%. These results therefore show that changing the longitudinal slope considerably affects the internal force of Y-shaped composite bridges.

Table 7 The stress of the control sections under loadings case 2 (Pa)

<table>
<thead>
<tr>
<th>Section position</th>
<th>i = 0%</th>
<th>i = 3%</th>
<th>i = 5%</th>
<th>i = 7%</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>1.228E+07</td>
<td>1.240E+07</td>
<td>1.292E+07</td>
<td>1.363E+07</td>
</tr>
<tr>
<td>S3</td>
<td>3.716E+07</td>
<td>3.679E+07</td>
<td>3.419E+07</td>
<td>3.159E+07</td>
</tr>
<tr>
<td>S7</td>
<td>3.124E+07</td>
<td>3.074E+07</td>
<td>3.024E+07</td>
<td>2.953E+07</td>
</tr>
<tr>
<td>S8</td>
<td>1.627E+07</td>
<td>1.647E+07</td>
<td>1.627E+07</td>
<td>1.724E+07</td>
</tr>
<tr>
<td>S13</td>
<td>4.639E+06</td>
<td>4.320E+06</td>
<td>4.639E+06</td>
<td>3.386E+06</td>
</tr>
<tr>
<td>S15</td>
<td>1.415E+06</td>
<td>1.454E+06</td>
<td>1.494E+06</td>
<td>1.557E+06</td>
</tr>
</tbody>
</table>

Table 8 The stress of the control sections under loadings case 2 (Pa)

<table>
<thead>
<tr>
<th>Section position</th>
<th>i = 0%</th>
<th>i = 3%</th>
<th>i = 5%</th>
<th>i = 7%</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>1.228E+07</td>
<td>1.240E+07</td>
<td>1.292E+07</td>
<td>1.363E+07</td>
</tr>
<tr>
<td>S3</td>
<td>3.716E+07</td>
<td>3.679E+07</td>
<td>3.419E+07</td>
<td>3.159E+07</td>
</tr>
<tr>
<td>S7</td>
<td>3.124E+07</td>
<td>3.074E+07</td>
<td>3.024E+07</td>
<td>2.953E+07</td>
</tr>
<tr>
<td>S8</td>
<td>1.627E+07</td>
<td>1.647E+07</td>
<td>1.627E+07</td>
<td>1.724E+07</td>
</tr>
<tr>
<td>S13</td>
<td>4.639E+06</td>
<td>4.320E+06</td>
<td>4.639E+06</td>
<td>3.386E+06</td>
</tr>
<tr>
<td>S15</td>
<td>1.415E+06</td>
<td>1.454E+06</td>
<td>1.494E+06</td>
<td>1.557E+06</td>
</tr>
</tbody>
</table>

Table 9 The ratios of maximum to minimum tensile stress at the control sections

<table>
<thead>
<tr>
<th>Loading cases</th>
<th>II–II (main bridge)</th>
<th>II–II (ramp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>1</td>
<td>1.05</td>
</tr>
<tr>
<td>C2</td>
<td>1</td>
<td>1.05</td>
</tr>
<tr>
<td>C3</td>
<td>3.3</td>
<td>3.3</td>
</tr>
<tr>
<td>C4</td>
<td>1</td>
<td>3.33</td>
</tr>
<tr>
<td>C5</td>
<td>1.07</td>
<td>1.06</td>
</tr>
<tr>
<td>C6</td>
<td>1</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Influence of loading types on stress
A comprehensive experimental study of the stress behaviour of control sections under six types of loading was performed, involving all the double fixed bearings, and the results were used to estimate the in-homogeneous distribution of the control section stress. Table 9 shows the effect of the tensile stress from maximum to minimum, which indicates that warping and distortion of the control sections subjected to eccentric concentrated force are greatly heightened, and stress distribution in the control sections is obviously non-homogeneous. Distribution of the part-normal stress is shown in Figures 12 and 13. The non-homogeneous distribution of stress at the control sections subjected to the eccentric uniform force is obvious, as is the decreasing stress. The warping and distortion of both the main bridge and the ramp are almost negligible. It should be pointed out that eccentric load should be taken into account in the warping and distortion control of Y-shaped bridges.

Influence of section height on girder stresses
Girder heights ranging from 1.0 m to 1.7 m were subjected to loading case 4 to determine the distribution of maximum tensile stress. Distortion stiffness, bending stiffness, warping stiffness and torsion stiffness for the structure subjected to the eccentric concentrated force increased correspondingly with an increase in the girder height. The stresses in the control sections decreased markedly with increase in the girder height, especially bending stress. The analysis showed that, for girder heights of 1.7 m to 1.0 m, the corresponding stresses in control sections may reduce by 54%, including mainly bending stress and warping stress, while distortion stress did not change significantly. Figure 14 shows the maximum stress versus beam height curve.

Influence of diaphragm width and stiffness on girder stresses
It is widely known that laboratory tests require a great amount of time, are very expensive and, in some cases, can even be
impractical, while, on the other hand, the finite element method has in recent years become a powerful and useful tool for the analysis of a wide range of engineering problems. A comprehensive finite element model permits a considerable reduction in the number of experiments. Nevertheless, in a complete investigation of any structural system, the experimental phase is still essential. Taking into account that numerical models should be based on reliable test results, experimental and numerical/theoretical analyses complement one another in the investigation of a particular structural phenomenon.

The diaphragms are located at the Y-shaped connection of the main bridge and ramp. Using a finite element model to reflect the effect of three different diaphragm stiffnesses on the mechanical behaviour of the bridge structure, three diaphragm widths of 0.5 m, 1.0 m and 1.5 m were modelled respectively. The analysis showed that, when the diaphragm width was increased to 0.5 mm, high moment capacity and rotation capacity were achieved. The stresses of the structure under loading case 4 decreased correspondingly with increase in the diaphragm width, and the stresses in sections I–I and II–II for diaphragm widths from 1.5 m to 0.5 m corresponding stresses at control sections may be reduced by 19.5% and 3.2% respectively. The modelling results show that the diaphragm width has an important influence on the stress in the control section of the main bridge and ramp. In addition, four diaphragm stiffnesses were modelled respectively two to five times. The results are shown in Figure 15.

From Figure 15 it can be seen that, when the diaphragm stiffness was increased two to five times, the strain of the control sections decreased by about 7%, 13%, 20% and 29% respectively. A stiffened bifurcation diaphragm would provide better stiffness for the connection between the main bridge and the ramp. Mechanical bifurcation diaphragm stiffening may be considered to prevent premature distortion, torsion or warping failure, and retain high stiffness and overall stability.

**PARAMETRIC STUDY OF DYNAMIC BEHAVIOUR**

A DASP test system, with a vibration signal collecting system, charge filter amplifier and accelerometer, was used for the dynamic testing (Figure 16). Thirteen points were selected to measure the dynamic behaviour of the model bridge. Accelerometers 1 to 7 were instrumented at main bridge sections, while accelerometers 8 to 13 were instrumented at ramp sections. Figure 17 shows the locations of the accelerometers on the composite bridge model. At the same time, to comprehensively reflect the dynamic characteristics of the Y-shaped bridge structure in numerical analysis, the space shell finite element model was adopted to determine boundary conditions, radius of curvature and diaphragm, etc.

**Influence of loading on mode shapes and frequencies**

Bridge loading is normally considered a major criterion affecting the fundamental frequencies of bridges, especially in the Y-shaped composite bridge. Results obtained from the model composite bridges are shown in Figures 18 and 19. It can be seen that fundamental frequencies change considerably under different loading conditions. It is interesting to note that, for each loading condition, the values of the fundamental frequencies vary within torsional frequencies. Mode shapes were also derived for all the bridges studied. For the model composite bridges, the first mode shape was always purely flexural. Comparisons revealed that, with different loading conditions, the first mode shape for the model composite bridges varied in torsion mode, while the second mode shape became more predominant.

**Influence of ramp radius of curvature on mode shapes and frequencies**

The relationships between the fundamental frequency and ramp radius of curvature are presented in Figure 20. The graph shows that increasing the ramp radius of curvature significantly increases the fundamental frequency. For instance, the fundamental frequency of mode bridges with a ramp radius of curvature of 1 667 mm increased by about 15.4% when the ramp radius of curvature increased from 1 800 mm to 2 200 mm.
Influence of bifurcate diaphragm plate stiffness

The effect of bifurcate diaphragm stiffness on the natural frequencies was examined. The results, shown in Figure 15, showed that bifurcate diaphragm stiffness has an effect on the modes. Providing bifurcate diaphragms at the bridge supports enhances the torsional stiffness significantly, and accordingly increases the natural frequencies. When the bifurcate diaphragm stiffness was increased to 10 mm, significant improvement in the torsional modes was shown. For instance, the fundamental frequency of mode bridges with bifurcate diaphragm plates of original stiffness increased by about 6.5% when the bifurcate diaphragm stiffness increased from ten to fifty times. The relationships between the fundamental frequency and bifurcate diaphragm stiffness are presented in Figure 21. Bifurcate diaphragm stiffness can therefore be used to improve the torsional frequency modes of Y-shaped composite bridges.

CONCLUSIONS

Parametric studies were conducted to evaluate the mechanical behaviour due to static and dynamic loading in which boundary condition, diaphragm stiffness, longitudinal slope, loading types, span length and ramp curvature, etc., were varied. The major conclusions from the investigation are as follows:

1. Based on a finite element program and existing commercial software (SAP2000), by comparing the results of the analyses of different boundary conditions, it is shown that the design method for eccentric loading in a Y-shaped bridge ramp is reasonable regarding the stress of the bridge structure. The torsional double-loading of the Y-shaped composite bridge is suitable to resist torsional and warping effects.

2. The ramp radius of curvature has no significant effect on torsion and distortion. However, it was observed that stress intensification on the middle-span section considerably increases with decreasing ramp radius.

3. In the middle-span section, the maximum bi-moments of torsion warping and distortion warping remained virtually unchanged; that is, the ramp-span length does not have a significant effect on torsion warping and distortion warping.

4. The results also show that changes to the longitudinal slope considerably affect the internal forces of Y-shaped composite bridges.

5. Warping and distortion are almost negligible when subjected to concentrated load. It must be pointed out that eccentric load should be taken into account in the warping and distortion of Y-shaped bridges.

6. The reductive stresses of the control sections of the Y-shaped bridge included mainly bending stress and warping stress when the girder height reached a certain value, but the distortion stress did not change significantly.

7. Sufficient stiffness of the bifurcation diaphragm is necessary to prevent premature distortion, torsion or warping failure and overall instability. The main bridge and ramp are both connected into a whole by the bifurcation diaphragm. Sufficient stiffness of the bifurcation diaphragm, together with torsional double-loading and pre-eccentric loading of the ramp, effectively controlled the effects of the torsion and distortion of the main bridge and ramp.

8. Comparisons revealed that, with different loading conditions, the first mode shape for the model composite bridges varied in torsion mode, while the second mode shape became more predominant.
9. Increasing the ramp radius of curvature significantly increases the fundamental frequency. For instance, the fundamental frequency of mode bridges with a ramp radius of curvature of 1 667 mm increased by about 15.4% when the ramp radius of curvature increased from 1 800 mm to 2 200 mm.

10. Bifurcate diaphragm stiffness is adequate to improve the torsional frequency modes for Y-shaped composite bridges.

ACKNOWLEDGEMENTS
The writers gratefully acknowledge the financial support provided by the Science Foundation of China Postdoctoral Grants (No 20110490183), the Science Foundation of the Ministry of Housing and Urban-Rural Development of the People's Republic of China (Grant No 2012-K2-6), the Education Department Science Foundation of Zhejiang Province (Grant No Y201122051), the Science Foundation of Shanghai Postdoctoral Grants (No 13R21421100) and the Science Foundation of China Postdoctoral Grants (No 2012-K2-6), the Science Foundation of the Ministry of Housing and Urban-Rural Development of the People's Republic of China (Grant No 2012-K2-6), the Education Department Science Foundation of Zhejiang Province (Grant No Y201122051), the Science Foundation of Shanghai Postdoctoral Grants (No 13R21421100).

REFERENCES
Effect of activator dosage, water-to-binder-solids ratio, temperature and duration of elevated temperature curing on the compressive strength of alkali-activated fly ash cement pastes

J Shekhovtsova, E P Kearsley, M Kovtun

In this paper the effect of sodium oxide concentration, the water-to-binder-solids ratio, temperature, and the duration of elevated temperature curing on the compressive strength of alkali-activated fly ash cement pastes was investigated. Alkali concentration varied between 3% and 15% Na₂O of fly ash mass. An increase in Na₂O from 3% to 9% greatly improved the compressive strength of the pastes from 26.1 MPa to 50.8 MPa at 28 days. A further increase in Na₂O up to 15% did not provide an increase in the strength, but a decrease was observed, as well as higher strength variation. The paste activated with 9% Na₂O had the highest strength at 28 days and a low standard deviation, and 9% Na₂O was thus considered as the best value in the present study. The temperature and the duration of elevated temperature curing were found to be critical factors affecting the compressive strength at early age, but their effect decreased significantly in the long term. The water-to-binder-solids ratio affected the compressive strength considerably. An increase in the water-to-binder-solids ratio of the pastes from 0.18 to 0.29 resulted in a decrease in the compressive strength from 49.3 MPa to 21.3 MPa.

INTRODUCTION

It is known that concrete is one of the most widely used construction materials, and ordinary Portland cement has generally been used as a binder component. The production of cement requires high-energy efforts which have a significant impact on the global emission of greenhouse gases. During the production of 1 ton of cement, between 0.73 and 0.99 tons of CO₂ are released into the atmosphere (Hasanbeigi et al 2012; Turner & Collins 2013). One of the major cement manufacturers in South Africa, Pretoria Portland Cement (PPC), reported that in 2011 its carbon footprint for cement was 892 kg CO₂ per ton of cement, which is an increase of 2.6% compared to 2010 (PPC Integrated Annual Report 2011).

Various ways of reducing CO₂ emission by using alternative binders are being investigated all over the world. Partial replacement of plain Portland cement by fly ash or slag was found to reduce concrete CO₂ emission by between 13% and 22% (Flower & Sanjayan 2007).

Alkali-activated binders and geopolymers are potential alternatives to plain Portland cement. Geopolymers are inorganic materials with three-dimensional silico-aluminate structures resulting from poly-condensation. Davidovits called the reaction which takes place as a result of alkali activation of aluminosilicates at low temperatures geopolymisation (Davidovits 1988). According to data published in literature, carbon emissions from geopolymers can be 80% less than that from traditional cements (Van Deventer et al 2010), and greenhouse gas emissions can be reduced by between 44% and 64% (McLellan et al 2011). Rich sources of aluminosilicates, most widely used for alkali activation, are fly ashes, metakaolin and blast furnace slag (Shi et al 2006; Palomo et al 1999a, 1999b; Van Jaarsveld & Van Deventer 1999; Bakharev et al 1999; Duxson et al 2007a). The process of alkali activation of fly ash starts from the dissolution of the aluminosilicate source in the alkaline solution during which the breakdown of the Si–O–Si and Al–O–Al covalent bonds in the amorphous phase of the initial material occurs, and ions...
The suitability of South African fly ash for use in alkali-activated binders was studied. This paper presents the results of an investigation into the effects of alkali content, water-to-binder-solids ratio, as well as temperature and duration of elevated temperature curing on the compressive strength of alkali-activated fly ash cement pastes containing low-calcium fly ash and sodium hydroxide solution as an aluminosilicate source and activator, respectively. The study was performed on pastes in order to eliminate the effect of additional variables, such as aggregate, with the aim to obtain main trends which would also be relevant for alkali-activated fly ash cement concretes.

### MATERIALS AND METHODS

Classified low-calcium fly ash (class F) from Lethabo power station in South Africa was used for the preparation of alkali-activated fly ash cement pastes. The chemical composition of the fly ash is given in Table 1. The mineralogical composition of thefly ash was studied by means of X-ray diffraction. The fly ash sample was analysed using a PANAnalytical X’Pert Pro powder diffractometer in 0–8 configuration with an X’Celerator detector and variable divergence and receiving slits with Fe filtered Co-Kα radiation. The phases were identified using X’Pert Highscore plus software, and 20% Si (Aldrich 99% pure) was added to the fly ash to determine the amorphous (glass) content. The relative phase amounts were estimated by the Rietveld method using the Autoquan – BGMN Rietveld Program, employing the fundamental parameter approach. Fly ash consists mainly of the amorphous phase (59.7% wt), with crystalline inclusion of mullite (29.9% wt) and quartz (10.2% wt). The laser diffraction method analysis of particle size distribution (Malvern Mastersizer 2000) showed that more than 80% of the fly ash particles were smaller than 45 μm.

Commercially available sodium hydroxide flakes (98.5% purity) were used to prepare the activator solutions. The Na₂O content was calculated as a percentage of the fly ash mass. It is essential to note that a noticeable decrease in the consistency of the pastes was observed when hot activator solution was mixed with the fly ash. Therefore, activator solutions were prepared in advance and cooled down to room temperature before mixing with the fly ash.

The experimental program was divided into three stages:

1. **Investigation into the effect of Na₂O content on the compressive strength and microstructure of alkali-activated fly ash cement pastes at constant water-to-binder-solids ratio, temperature and duration of elevated temperature curing.**

2. **Study of the effect of temperature and duration of elevated temperature curing on the compressive strength of the pastes at a constant activator content and water-to-binder-solids ratio.**

3. **Investigation into the effect of the water-to-binder-solids ratio ((mass of water from sodium hydroxide solution + mass of added water) / (mass of fly ash + sodium oxide in sodium hydroxide)) on the compressive strength of alkali-activated fly ash cement pastes at constant sodium oxide content, temperature and duration of elevated temperature curing.**

The combined effect of the variables on the compressive strength of the alkali-activated fly ash cement pastes was not studied.

Alkali-activated fly ash cement pastes were prepared by mixing the fly ash with the activator solution for three minutes in a pan mixer. The pastes were cast into prismatic 40×40×160 mm moulds immediately after mixing, covered with a film, and placed without any delay into an oven pre-heated to the desired temperature. During the first stage of the study five levels of Na₂O content, varying from 3% to 15%, at a constant water-to-binder-solids ratio of 0.2 were investigated. Samples were oven-cured at an elevated temperature of 60°C for 24 hours. After elevated temperature curing, samples were taken out of the oven without gradual cooling, de-moulded, cooled down to room temperature, and tested. Samples, to be tested at other ages, were kept in a room at constant temperature and humidity (25°C and 55% RH). Compressive strength testing was carried out on halves of broken prisms, and the test results of the experiment are expressed in this paper as the arithmetic mean of the six compressive strength values obtained in compliance with SANS 50196-1 on a set of three prisms. Flexural strength results are not presented in this paper, due to an extremely high strength variation and the absence of any significant trend.

During the second stage of the study, the Na₂O content that provided the highest compressive strength at 28 days in the first stage was used to investigate the effect of temperature and duration of elevated temperature curing on the strength development.

---

### Table 1 Chemical composition of Lethabo fly ash, wt%

<table>
<thead>
<tr>
<th>Element</th>
<th>SiO₂</th>
<th>TiO₂</th>
<th>Al₂O₃</th>
<th>Fe₂O₃</th>
<th>MnO</th>
<th>MgO</th>
<th>CaO</th>
<th>Na₂O</th>
<th>K₂O</th>
<th>P₂O₅</th>
<th>SO₃</th>
<th>SrO</th>
<th>LOI</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass</td>
<td>55.22</td>
<td>1.57</td>
<td>31.93</td>
<td>3.36</td>
<td>0.02</td>
<td>0.73</td>
<td>4.57</td>
<td>0.01</td>
<td>0.87</td>
<td>0.42</td>
<td>0.23</td>
<td>0.12</td>
<td>0.80</td>
<td>0.15</td>
</tr>
</tbody>
</table>

(silicon and aluminium) pass into the solution with the formation of Si–OH and Al–OH groups, followed by the polymerisation and reorganisation of species. Thereafter the solidification of aluminosilicate gel occurs with the formation of three-dimensional low-ordered structures (Van Deventer *et al.* 2007; Fernandez-Jiménez *et al.* 2005b). This gel has been found to be responsible for the cementitious properties of the final material (Criado *et al.* 2008), and the quantity affects the mechanical strength of the final product (Zhang *et al.* 2013).

Utilisation of ash is a serious problem for South Africa, as energy production relies mostly on coal. Sasol (petrochemical) and Eskom (power utility) are the largest coal processing companies in South Africa. More than 30 million tons of coal are consumed annually by Sasol, and about 8 million tons of gasification ash are produced (Matjie *et al.* 2005). Eskom consumes more than 100 million tons of coal per annum (Ash Management in Eskom 2013), and more than 35 million tons of ash (fly and bottom) were produced in 2011 (Eskom Integrated Report 2011). Almost 90% of the ash produced by Eskom is fly ash. The morphological features of fly ash result in improved workability of cement systems, while the pozzolanic activity makes fly ash suitable for use as a cement extender (Kruger & Krueger 2005). However, in South Africa only about 5% of all fly ash produced by Eskom is beneficially used (Ash Management in Eskom 2013; Bada & Potgieter-Verraak 2008) and millions of tons of ash are being stored and disposed of in ash dams and landfills annually, creating the risk of toxic elements present in fly ash being released into the soil and groundwater (Carlson & Adriano 1993).

Developing an alternative application for fly ash as a raw material to produce an environmentally friendly construction material, contributes towards finding a solution for fly ash utilisation. This would expand the raw-material base of the building industry for fly ash utilisation. This would expand the raw-material base of the building industry for fly ash utilisation.
of alkali-activated fly ash cement pastes. Four different temperatures (25°C, 40°C, 60°C and 80°C) and six durations of elevated temperature curing (ranging from 4 to 24 hours with 4-hour intervals) were studied in order to find efficient curing conditions. The water-to-binder-solids ratio was kept constant at 0.2.

The effect of the water-to-binder-solids ratio (in the range of 0.18 to 0.29) on the strength development of the alkali-activated fly ash cement pastes was investigated during the third stage of the study. The Na₂O content, provided by the highest 28-days compressive strength, was used for paste preparation during the third stage. The pastes were cured at 60°C for 24 hours. At low water-to-binder-solids ratios (0.2 and below) the pastes were dry and lumpy, but the consistency of the pastes was improved significantly when external energy was applied. Vibration for an appropriate time resulted in proper compaction of the alkali-activated fly ash cement pastes.

The microstructure of fresh fractured surfaces of alkali-activated fly ash cement pastes, coated with carbon, was investigated with a scanning electron microscope (SEM) at 20 kV (JEOL JSM 5800).

**RESULTS AND DISCUSSION**

**The effect of Na₂O content on the compressive strength and microstructure**

**Compressive strength**

Compressive strength and standard deviation values of alkali-activated fly ash cement pastes with different sodium oxide contents are shown in Table 2. The compressive strength of alkali-activated fly ash is strongly affected by the concentration of alkaline solution (Fernandez-Jimenez & Palomo 2005a) and this is confirmed by the present study. An increase in Na₂O from 3% to 9% of fly ash mass resulted in a significant increase in the compressive strength of the pastes tested after elevated temperature curing. A further increase in alkali content to 12% and 15%, however, resulted in a decrease in the compressive strength. This is in good agreement with the results published by Stevenson and Sagoe-Crentsil (2005) and Gou et al (2010) who noticed an increase in compressive strength with increased alkali content up to 10%, but a decrease in strength at 12% and 15%. Compressive strengths from 23.0 MPa to 25.5 MPa were obtained by Somna et al (2011), with NaOH concentrations varying from 9.5 M (7.3% of Na₂O) to 14 M (10.1%). They reported a decrease in compressive strength for a concentration of 16.5 M (11.3%) due to early precipitation of aluminosilicate products. Palomo et al (1999a) indicated that the strength decrease of the alkali cement was caused by an excess of OH⁻ concentration in the system. Rattanasak and Chindaprasirt (2009) found that higher concentrations of Si⁴⁺ and Al³⁺ were obtained with a 10 M NaOH (7.7%) compared to 15 M (10.7%). A high concentration of OH⁻ resulted in a decrease in the ability of Al and Si to participate in the geopolymerisation. In this study, a gain in strength of up to 30% and more was observed after 28 days of curing for all samples (see Table 2). A noticeable change in compressive strength with time was observed by other researchers (De Vargas et al 2011; Somna et al 2011). This confirms that age has a significant effect on the strength of alkali-activated fly ash cement pastes, and the structural development does not stop with the discontinuation of elevated temperature curing. The strength increase probably relates to the transformation of aluminosilicate gel to more stable structures (Škvara et al 2009; Ravikumar et al 2010). At the age of 28 days the highest compressive strength value was again obtained for the 9% Na₂O paste. The paste with 12% Na₂O yielded a higher compressive strength value after 91 and 182 days, but after one year of curing, the strength of the Na12 paste was lower than that of Na9. The difference in compressive strength values between 9% and 12% at later ages is negligible, and therefore using lower alkali content is more economical, as the alkali is the most expensive component in the alkali-activated fly ash cement paste composition. Therefore, 9% of Na₂O can be considered as the most suitable alkali concentration to produce alkali-activated fly ash cement paste containing the South African fly ash studied in this research.

It should be mentioned that the greatest gain in strength was observed for alkali-activated fly ash cement pastes with high alkali content. Similar findings were reported by De Vargas et al (2011). The authors explained that the gain in strength with time was caused by additional reaction products generated during the reaction between alkali activator and microspheres, which packed inside plerospheres. The alkaline solution first partially dissolved the external layer of the plerosphere and only then attacked the microspheres.

As can be seen from the standard deviation values in Table 2, the observed compressive strength of the 15% Na₂O paste samples varied significantly. It can be assumed that the high concentration of alkali in the pastes results in inconsistent compressive strength. Additional tests were performed on larger

---

**Table 2 Compressive strength and standard deviation of alkali-activated fly ash cement pastes with different Na₂O content**

<table>
<thead>
<tr>
<th>Paste</th>
<th>Na₂O, wt %</th>
<th>after elevated temperature curing</th>
<th>7 days</th>
<th>28 days</th>
<th>91 days</th>
<th>182 days</th>
<th>364 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na3</td>
<td>3</td>
<td>17.7 (1.15)</td>
<td>19.8 (0.87)</td>
<td>26.1 (0.98)</td>
<td>25.4 (1.45)</td>
<td>23.7 (1.43)</td>
<td>25.7 (2.07)</td>
</tr>
<tr>
<td>Na6</td>
<td>6</td>
<td>36.0 (1.23)</td>
<td>36.6 (1.37)</td>
<td>43.7 (2.01)</td>
<td>49.1 (2.36)</td>
<td>44.4 (1.90)</td>
<td>51.7 (2.86)</td>
</tr>
<tr>
<td>Na9</td>
<td>9</td>
<td>43.1 (1.24)</td>
<td>41.6 (2.76)</td>
<td>50.8 (2.00)</td>
<td>60.0 (2.90)</td>
<td>63.6 (1.52)</td>
<td>67.9 (5.19)</td>
</tr>
<tr>
<td>Na12</td>
<td>12</td>
<td>36.1 (1.07)</td>
<td>39.8 (1.96)</td>
<td>48.6 (1.77)</td>
<td>62.8 (3.96)</td>
<td>68.9 (3.85)</td>
<td>64.9 (1.36)</td>
</tr>
<tr>
<td>Na15</td>
<td>15</td>
<td>36.6 (2.75)</td>
<td>38.8 (3.98)</td>
<td>49.8 (2.36)</td>
<td>55.4 (6.06)*</td>
<td>60.8 (8.62)*</td>
<td>62.7 (4.48)</td>
</tr>
</tbody>
</table>

* These values include more than one statistical outlier, which varied by more than ± 10% from the mean and would normally be excluded according to SANS 50196-1.

---

**Table 3 Compressive strength, standard deviation and coefficient of variation at 28 days**

<table>
<thead>
<tr>
<th>Na₂O, wt %</th>
<th>Average strength, MPa</th>
<th>Standard deviation, MPa</th>
<th>Coefficient of variation, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>51.4</td>
<td>4.52</td>
<td>8.8</td>
</tr>
<tr>
<td>12</td>
<td>41.4</td>
<td>3.33</td>
<td>8.1</td>
</tr>
<tr>
<td>15</td>
<td>35.6</td>
<td>5.79</td>
<td>16.3</td>
</tr>
</tbody>
</table>
sample populations for Na9, Na12 and Na15 pastes to establish the distribution of the compressive strength, and these results are shown in Table 3. Curing conditions were the same as in the first test. Samples Na3 and Na6 were not studied in the second test due to their relatively low compressive strength compared to samples with higher alkali dosage and satisfactory standard deviation. At least thirty-six compressive strength determinations at the age of 28 days were performed for each of these pastes. For the paste with 9% of Na2O, six results within 36 determinations of the compression strength varied by more than ± 10% from the mean, which was 16.7% of the total sample population. For the paste with 12% of Na2O, seven strength values out of 42 varied by more than ± 10% from the mean, which was also 16.7%. The paste with 15% of Na2O had 18 strength values out of 36 varying by more than ± 10% from the mean, which represented 50% of the total sample population. The strength deviation was thus greater for alkali-activated fly ash cement pastes containing 15% Na2O.

Results show that the coefficient of strength variation within the large batch of Na15 paste was much higher than that of Na9 and Na12 pastes (16.3%, 8.8% and 8.1% respectively). It is interesting to note that high alkali content affects not only the deviation of compressive strength within a batch, but also between two batches of the same composition. The average 28-days strength of the Na9 paste for the second (large) batch was almost the same as for the first batch (51.4 MPa and 50.8 MPa respectively) while the average 28-days strength of Na15 paste for the second batch was much lower than for the first batch (55.6 MPa and 49.8 MPa respectively). The same trend could be observed for Na12 paste where the average 28-days strength for the first batch was 48.6 MPa while it reached only 41.4 MPa for the second batch. Limited strength deviation should be expected because of the variability of the fly ash composition due to variations in coal composition and burning conditions. Nevertheless, the strength deviation between batches was much greater with higher alkali concentrations.

Another important observation was the formation of efflorescence on the surfaces of the Na12 and Na15 paste samples after 28 days of hardening. The formation of efflorescence is an indirect indicator of excess Na2O content in the alkali-activated fly ash cement pastes with 12% and 15% of Na2O. The excess alkali migrates with moisture to the surface of the samples and produces salts that appear as white efflorescence. The presence of excess alkali could be one of the causes of the lower-strength, high-standard deviation, and variations in the strength between the batches of alkali-activated fly ash cement pastes.

Microstructure

SEM images of alkali-activated fly ash cement pastes with different Na2O content at the age of 40 days are presented in Figures 1 and 2.

Figure 1 shows the microstructure of the Na3, Na6 and Na9 pastes. The surface of the particles is covered by shell-shaped reaction products with smooth fly ash particles under the shell. It was found in previous research that the main reaction product of the alkali activation of fly ash was disordered aluminosilicate gel, also known as geopolymeric gel (Palomo et al 1999a; Duxson et al 2007b; Provis & Van Deventer 2009). It can be seen that fly ash particles are glued to one another by reaction products. Voids observed between fly ash particles indicate that not enough aluminosilicate gel has been formed to fill the space, resulting in the friable appearance of the microstructure. The amount of unreacted fly ash particles decreases with increasing concentration of the alkali, and the matrix appears more continuous, which indicates a formation of increased amounts of gel (see Figure 2). As it was assumed that the aluminosilicates gel was responsible for the mechanical properties of the final product (Zhang et al 2013), it was expected that the more continuous appearance of the matrix was related to the greater amount of gel, resulting in the higher compressive strength. This is in good correlation with the compressive strength results, with the exception of the Na15 paste. The matrix of the Na15 paste appears continuous and the microstructure looks the most solid, but the compressive strength tends to be less than that of the Na9 and Na12 pastes (see Table 2).

One of the reasons for the strength drop with increasing alkali content over 9% Na2O could possibly be micro-cracks which can be observed in images of the microstructure of alkali-activated fly ash cement pastes containing 12% and 15% Na2O (see Figure 2).
The nature of these cracks is not yet clear. De Vargas et al (2011) discovered similar cracks in alkali-activated fly-ash-based geopolymers. The authors reported that micro-cracks could be found easier in the samples with higher amounts of alkali (16.4% Na₂O compared to 8.2% Na₂O), but they did not observe a reduction in compressive strength. At the same time Fernandez-Jimenez and Palomo (2005a) linked cracks with elevated temperature curing during the activation process, or mechanical damage during sample preparation for SEM observation. However, all the samples presented in the micrographs in this study had the same regime of elevated temperature curing and sample preparation for the SEM study, but cracks were only observed in the Na12 and Na15 pastes. The appearance of the small cracks differs from long, wider cracks which most probably are the result of sample preparation. The small cracks appear within the continuous matrix and do not go out beyond its boundaries. A detailed study of these cracks and their origin should be conducted.

The effect of temperature and duration of elevated temperature curing on the compressive strength

Based on the results of the first stage of this study, the Na9 paste was used to investigate the effect of temperature and duration of elevated temperature curing on the compressive strength of alkali-activated fly ash cement pastes. Compressive strength results and standard deviation values for the second stage of the study are shown in Table 4.

It was expected that an elevation of curing temperature would accelerate the dissolution of the glass phase of the fly ash and, as a result, the strength development of the alkali-activated fly ash cement pastes. Previous studies showed that temperature accelerated the alkali activation of metakaolin (Alonso & Palomo 2001), as well as slag pastes (Bakharev et al 1999). The accelerating effect of elevated temperature also applies to fly ash pastes (Katz 1998). The results presented in this paper confirm the trend (see Table 4).

The importance of elevated temperature curing can clearly be seen in Figure 3a. The error bars on all figures represent deviation of the compressive strength from the mean for six values. Samples cured at 25°C did not set after 24 hours and could hardly be de-moulded after seven days, when the compressive strength was 1.1 MPa. Even such a low strength was an indicator of chemical reaction between the fly ash and the activator solution. The strength of the paste cured at 60°C for 24 hours, and tested immediately after the elevated temperature curing

<table>
<thead>
<tr>
<th>Pastes ID (T for temperature in °C; D for duration in hours)</th>
<th>Average compressive strength and standard deviation (in brackets) at different testing ages, MPa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>after elevated temperature curing</td>
</tr>
<tr>
<td>T65D4</td>
<td>3.1 (0.20)</td>
</tr>
<tr>
<td>T70D4</td>
<td>15.5 (0.86)</td>
</tr>
<tr>
<td>T75D4</td>
<td>31.4 (1.60)</td>
</tr>
<tr>
<td>T25D24</td>
<td>–</td>
</tr>
<tr>
<td>T40D24</td>
<td>1.6 (0.14)</td>
</tr>
<tr>
<td>T60D4</td>
<td>2.1 (0.11)</td>
</tr>
<tr>
<td>T60D8</td>
<td>15.2 (1.36)</td>
</tr>
<tr>
<td>T60D12</td>
<td>28.5 (0.89)</td>
</tr>
<tr>
<td>T60D16</td>
<td>43.3 (1.47)</td>
</tr>
<tr>
<td>T60D20</td>
<td>49.8 (2.49)</td>
</tr>
<tr>
<td>T60D24</td>
<td>49.4 (1.50)</td>
</tr>
<tr>
<td>T80D4</td>
<td>36.0 (1.23)</td>
</tr>
<tr>
<td>T80D8</td>
<td>40.2 (2.62)</td>
</tr>
<tr>
<td>T80D12</td>
<td>42.9 (1.98)</td>
</tr>
<tr>
<td>T80D16</td>
<td>51.0 (3.32)</td>
</tr>
<tr>
<td>T80D20</td>
<td>50.9 (3.73)</td>
</tr>
<tr>
<td>T80D24</td>
<td>53.0 (3.69)</td>
</tr>
</tbody>
</table>

* These values include more than one statistical outlier, which varied by more than ± 10 % from the mean and would normally be excluded as in SANS 50196-1.
exceeded the strength of the paste, cured at 25°C for 91 days. Curing of the alkali-activated fly ash cement paste at 25°C is not practical, due to slow strength development, intensive efflorescence formation, relatively low strength and high standard deviation (see Table 4). Thus, elevated temperature curing is necessary to provide faster strength development and lower standard deviation of the strength. An increase in temperature from 25°C to 60°C produced significant acceleration in the strength development. Immediately after elevated temperature curing, the paste cured at 40°C, 60°C and 80°C for 24 hours had a compressive strength of 1.6 MPa, 49.4 MPa and 53.0 MPa respectively. The increase in the compressive strength of the paste was a result of an increase in the degree of polymerisation caused by the elevation of the curing temperature (Povnanik 2010). It is important to note that when the paste was cured at elevated temperature for 24 hours, an increase in the temperature to above 60°C did not result in a considerable gain in compressive strength. The use of 60°C instead of 80°C for the prolonged elevated temperature curing of alkali-activated fly ash cement pastes will require less energy, and will thus be more economical.

After 28 days of curing, the compressive strength of samples T40D24, T60D24 and T80D24 increased from 1.6 MPa, 49.4 MPa and 53.0 MPa, to 25.5 MPa, 59.1 MPa and 63.5 MPa respectively. Sample T25D24 could not be de-moulded at day one, but it still gained 10.9 MPa after 28 days of hardening. The compressive strength of the alkali-activated fly ash cement pastes increased with the reaction time, and the gain was greater in those cases where the pastes were cured at low temperatures (25°C and 40°C). At the age of 91 days, the paste that had been cured at 80°C had a lower compressive strength compared to the paste cured at 60°C. This strength decrease was probably caused by the contraction of aluminosilicate gel, due to dehydration and excessive shrinkage occurring during the curing at high temperatures for 24 hours (Van Jaarsveld et al 2002). Chindaprasirt et al (2007) drew conclusions regarding the effect of elevated temperature curing on strength, considering seven-day strength results, but the results of this study and others (De Vargas et al 2011; Arioz et al 2012) show that the difference in strength of alkali-activated fly ash cement pastes cured at low and high temperatures decreases with age, and even more, lower temperatures could lead to higher strength in the long term.

Another important observation is the long-term gain in compressive strength of the alkali-activated fly ash cement pastes cured at elevated temperatures for different periods of time. After 28 days of hardening, the compressive strength of sample T60D4 raised by 20 times (from 2.1 MPa to 41.5 MPa), while the compressive strength of sample T80D4 increased only by 1.4 times (from 36.0 MPa to 50.8 MPa) (see Table 4). For samples T60D24 and T80D24, the compressive strength increased from 49.4 MPa to 59.1 MPa, and from 53.0 MPa to 63.5 MPa respectively, showing a 1.2 times gain for both samples. These results show that the ageing of the alkali-activated fly ash cement pastes has a greater effect on the strength development of the pastes cured at the lower temperature for a short period of time.

It was shown that the speed of reaction between alkali and fly ash depends on curing temperature, especially in the first few hours (Palomo et al 1999a). A relatively small increase in the temperature for elevated temperature curing resulted in a significant gain in the compressive strength of the pastes cured for a short period of time (Figure 3b). Increasing the temperature from 65°C to 70°C produced a significant increase in the strength from 3.2 MPa to 15.5 MPa. A further increase (by 5°C) doubled the strength from 15.5 MPa to 30.5 MPa. Subsequently, increasing the temperature to 80°C did not lead to a significant increase in the strength in comparison to 75°C. After 28 and 91 days of hardening, the difference in the strength of the alkali-activated fly ash cement pastes, cured at elevated temperatures in the range of 60°C to 80°C, was not significant. Thus, despite the significant effect of the increase in the curing temperature (in the range from 60°C to 80°C) on the early compressive strength of the pastes, the difference in the curing temperature had a limited effect on the long-term strength of the pastes when elevated temperature curing was applied for a short period of time (four hours).
Figure 4 shows the effect of the duration of elevated temperature curing on the compressive strength of the alkali-activated fly ash cement pastes cured at 60°C and 80°C.

After four hours of elevated temperature curing at 60°C (see Figure 4a), the compressive strength of the paste tested immediately after the curing was 2.1 MPa. With other parameters being equal, the compressive strength of the paste cured at 80°C was 36 MPa (see Figure 4b). This once again confirms that temperature during the initial curing plays a key role in the strength development of alkali-activated fly ash cement pastes. For longer durations of elevated temperature curing, the difference in the strength of the pastes cured at 60°C and 80°C becomes less prominent. The compressive strength of sample T60D24 was 49.4 MPa, while that for sample T80D24 was 53.0 MPa.

It is important to note that a significant increase in the compressive strength, tested immediately after elevated temperature curing, of the paste cured at 60°C was observed only when the duration of the curing was increased to 16 hours (see Figure 4a). Subsequent increases in the duration of elevated temperature curing did not result in a substantial increase in strength. A similar trend was observed when the paste was cured at 80°C (see Figure 4b). Insignificant strength gain took place when the duration of elevated temperature curing exceeded 16 hours. In the long term, curing at 80°C for more than 16 hours negatively affected the compressive strength. At 28 and 91 days, the pastes cured in an oven for 20 and 24 hours had lower strengths in comparison to the paste cured for 16 hours. Van Jaarsveld et al (2002) reported that curing for longer periods of time at elevated temperature appeared to weaken the microstructure, suggesting that small amounts of structural water needed to be retained in order to reduce cracking and maintain structural integrity. Despite the fact that a few researchers (Swanepoel & Strydom 2002; Chindaprasirt et al 2007) commented alkali concentration is 9% Na2O results in a decrease in strength, high water-to-binder-solids ratio should be achieved at the lowest possible water-to-binder-solids ratio.

The effect of the water-to-binder-solids ratio on compressive strength

Na9 paste cured at 60°C for 24 hours was studied during the third stage. The water-to-binder-solids ratio varied in the range of 0.18 to 0.29. The effect of the water-to-binder-solids ratio on the compressive strength of the alkali-activated fly ash cement pastes is shown in Figure 5.

Increasing the water-to-binder-solids ratio of the pastes from 0.18 to 0.29 resulted in a decrease in the compressive strength from 49.3 MPa to 21.3 MPa, but the consistency of the pastes improved. Fernandez-Jimenez and Palomo (2005a) admitted the importance of the water-to-binder ratio. According to Davidovits (1988), no water combined within the geopolymer. Water acts as a carrier of alkalis (Škvara et al 2009) and provides consistency to the fresh geopolymer mixture (Hardjito & Rangan 2005). Škvara et al (2009) reported that about 65% of all water in geopolymers was in "free" condition as it was evaporable at 180°C, with 30% presumed to come from nano-pores in the geopolymer gel. The water introduced to the paste could evaporate from unsealed samples during elevated temperature curing, thus negatively affecting the final structure of material. Samples should therefore preferably be sealed to prevent extensive moisture evaporation.

During the design of alkali-activated fly ash cement concretes, a required workability should be achieved at the lowest possible water-to-binder-solids ratio.

CONCLUSIONS

The trends observed for the alkali-activated fly ash cement pastes should be relevant for concretes containing alkali-activated fly ash cement.

The alkali content plays an important role in the development of the compressive strength of alkali-activated fly ash cement pastes and their microstructure. The recommended alkali concentration is 9% Na2O of fly ash mass, which provides the highest compressive strength and low standard deviation at 28 days. Excessive alkali content (> 9% Na2O) results in a decrease in strength, high standard deviation and high coefficient of strength variation between different batches.
Alkali content higher than 9% Na₂O also causes efflorescence formation and, possibly, the formation of micro-cracks in the microstructure of the pastes. Therefore, the amount of alkali must be strictly controlled during the production of alkali-activated fly ash cement.

The temperature and duration of curing of alkali-activated fly ash cement pastes affect the compressive strength significantly. Curing at 25°C is possible, but it is not practical, due to delayed setting, intensive efflorescence formation, very slow strength development, relatively low strength at 28 days and large strength deviation. Therefore, it is important to provide elevated temperature curing, thus accelerating the strength development of the alkali-activated fly ash cement pastes.

Elevated curing temperature has a greater effect on the early strength than the long-term strength, especially of pastes cured for a short period of time. An increase in temperature over 60°C did not noticeably affect the 28-day and 91-day compressive strength of alkali-activated fly ash cement pastes cured for four and 24 hours. A decrease in the 91-day compressive strength of the paste cured at 80°C for 24 hours was observed, in comparison to the paste cured at 60°C for the same period of time. The duration of elevated temperature curing has a more prominent effect on the early strength. The 28-day and 91-day compressive strengths are less affected by the duration of elevated temperature curing. There is no significant increase in the early compressive strength when the duration of elevated temperature curing exceeds 16 hours. Elevated temperature curing at 60°C for 16 hours is recommended for curing of alkali-activated fly ash cement pastes.

The compressive strength of alkali-activated fly ash cement pastes is significantly affected by the water-to-binder-solids ratio. The compressive strength decreases with an increase in water-to-binder-solids ratio. Therefore, alkali-activated fly ash cement concretes should be designed to achieve a required workability at the lowest possible water-to-binder-solids ratio.

Alkali-activated fly ash cement pastes were produced with a compressive strength between 10 MPa and 60 MPa at 28 days. These results indicate that the Lethabo fly ash can be used as a source of alumino-silicates in alkali-activated fly ash cement formulation to produce a concrete with good mechanical properties. The short-term and long-term properties of alkali-activated fly ash cement concrete are currently under investigation at the University of Pretoria.

**ACKNOWLEDGEMENTS**

The authors gratefully acknowledge Dr R A Kruger and Ash resources (Pty) Ltd for supplying the fly ash for these experiments. The authors would like to thank Wiebke Grote and Jeanette Dykstra for XRD and XRF analyses, and the staff of the Microscopy and Microanalysis Laboratory of the University of Pretoria for providing access to the microscope, and assistance during the investigation.

**REFERENCES**


Directional wave spectra provide detailed information about wave climates some of which can be important for coastal vulnerability assessments and design applications. Spectral data is also important for calibrating and validating spectral wave models that are widely used in coastal engineering. This paper discusses directional spectra derived from measurements on the east coast of South Africa (16-year data set). A variety of spreading functions are applied and compared. The Cosine-2s and Gaussian distributions produce similar results and seem to give an appropriate representation of directional spreading at the case study location. The spectra show some seasonal variations, with the broadest directional spreading in summer and the narrowest (from the southeast) in winter. The winter season also has the highest wave energy of the seasons. The spectral data has no clear evidence of cyclone activity contributing significant northeasterly wave energy as has often been conjectured for this location. The occurrence of wave energies above a threshold that causes significant coastal erosion varies seasonally, and mainly comprises long period swells linked to distinct weather systems. The analysis and results reported here provide insights for modelling coastal vulnerability and designing coastal infrastructure.

INTRODUCTION

Summary wave statistics (e.g. wave height, period and direction) are an incomplete and often insufficient means of characterising complex wave fields. Parameters such as significant wave height, peak wave period and direction are only partial descriptors of the full wave energy spectrum and neglect significant spatio-temporal information. Understanding wave directional spectra, including their variability and seasonality, has important applications in coastal vulnerability assessment and engineering design. For example, some beach types may only be vulnerable to specific wave directions, and this information may not be evident in basic wave statistics.

Spectral wave models have become an integral part of coastal engineering design and assessments. Detailed calibration and validation of these models can be improved by using measured directional wave spectra. Therefore, understanding site-specific features of directional wave spectra is an increasingly important component of all coastal and marine work. The measurement and analysis of directional wave spectra has captured global interest for a variety of applications, but mainly for the application of spectral wave models. For example, Alves and Melo (1999) analysed Brazilian wave data with the intention of creating input data for numerical models of wave evolution. Lucas et al. (2011) analysed directional wave spectra from Portuguese coastal waters and North Atlantic locations in an attempt to create a statistical description of the wave climates, while Naffaa (1995) collected and analysed wave spectra along the Nile Delta coast to estimate longshore sediment transport and to design coastal protection works.

The east coast of South Africa (Figure 1) has two relatively long records of wave data. Basic wave statistics derived from the data have been reported by Corbella and Stretch (2012d), Rossoouw (2001) and Rossoouw (1984), but no detailed descriptions of the measured directional wave spectra have been published to date. The region has an energetic wave climate and some chronic erosion problems are occurring on the KwaZulu-Natal coastline (Corbella and Stretch, 2012a, c).

The principal forcing mechanisms for storm waves on the KwaZulu-Natal coastline are generally thought to be associated with three weather systems, namely (1) cold fronts and coastal lows, (2) extra-tropical cyclones and cut-off low pressure systems, and (3) tropical cyclones. Detailed descriptions of South African weather conditions can be found in Hunter (1987), Preston-Whyte and Tyson (1993), and Taljaard (1995). Cold fronts or extra tropical cyclones move from west to east and are associated with large intense wave fields. The effect of these systems on the KwaZulu-Natal coastline is relatively small compared to the effects on the west coast and southern east coast. Cold fronts and coastal lows generally exist closer to the coast than cut-off lows or cyclones, and are therefore typically associated with smaller wave heights and shorter periods from a southerly wave direction. Cut-off lows are associated with a southeasterly wave direction, but form further offshore.
and produce longer period waves. Tropical cyclones are sometimes cited as an important destructive force along the KwaZulu-Natal coast. However, recent research has suggested that these events may not be as significant as previously conjectured (e.g. Kruger et al. 2010; Corbella & Stretch 2012b). Tropical cyclones produce long period waves typically from the north-east or east-north-east.

The aims of this paper are (1) to compare various directional spreading functions for the directional wave spectra, (2) to investigate seasonal variations in wave spectra, (3) to elucidate the spectra typically associated with erosive storm events, and (4) to evaluate the difference in spectral characteristics of wave events driven by distinct weather systems such as tropical cyclones near Madagascar.

We start by describing the available data and the methods used for the analysis, followed by a presentation of the results, in turn followed by a discussion.

**METHODS**

**Case study**

The east coast of South Africa (Figure 1) has two relatively long records of wave data measured at Durban and Richards Bay. Basic wave statistics derived from the data have been reported by Corbella and Stretch (2012d), Rossouw (2001) and Rossouw (1984). A summary of these wave statistics is given in Table 1, and wave roses are shown in Figure 2. The largest waves tend to occur during the autumn and winter seasons, followed by spring and summer. All directions are stated in the nautical convention.

**Spectral data**

Spectral data for the east coast of South Africa is limited to Durban and Richards Bay (Figure 1). The collection of wave data at both sites has been primarily for port operations. Although Richards Bay has 34 years of spectral wave data, this study was solely concerned with directional wave spectra that are only available for 16 years. Durban has only six years of directional wave spectra data. Durban’s six-year data set is from a single 0.9 m diameter Datawell Directional Waverider buoy located at constant coordinates and a water depth of 30 m (Table 2). The Richards Bay data sets are made up of two different instruments, but located at the same coordinates and a water depth of 22 m (Table 2). The directional wave recording buoys that have been used on the east coast of South Africa have been predominately Datawell Waverider buoys with the exception of a CSIR-developed buoy (called a GPS 3D buoy) using similar principles. The Waverider buoys sample at a rate of 1.28 Hz, and therefore collect a total of 256 heave samples every 200 seconds. The available data sets have spectral data available at 3-hour, 1-hour and 0.5-hour intervals.

It should be noted that the waves recorded by the wave recording buoys do not accurately represent deep-water waves. To explore the effects of refraction we simulate waves of different periods and directions with constant wave heights in a calibrated SWAN model.
(Booij et al. 1999). The SWAN model consists of three curvilinear grids. The coarse grid has an average cell size of 500 m × 500 m, the medium grid is 120 m × 120 m and the fine grid is 20 m × 20 m.

The recorded data consists of frequency, power spectral density (PSD), mean wave direction and directional spread for each frequency interval. The directional spread is the standard deviation of the wave directions in the recording interval, and is defined as:

$$\sigma(\omega) = \sqrt{2(1 - m_1(\omega))}$$  \hspace{1cm} (1)

where

$$m_1(\omega) = \sqrt{a_1^2(\omega) + b_1^2(\omega)}$$ \hspace{1cm} (2)

is the amplitude of first harmonic and $a_1$ and $b_1$ are angular Fourier coefficients in an expansion of the form:

![Wave roses for Durban and Richards Bay wave data for both combined and seasonally grouped cases](image-url)

Figure 2: Wave roses for Durban and Richards Bay wave data for both combined and seasonally grouped cases (reproduced from Corbella & Stretch 2012d)
Cosine-2s frequency a spreading function can be used. A uniform or Top-hat distribution as coefficients can be derived directly from the Fourier (e.g. Goda 2008). Directional distributions remains limited knowledge on the subject inherently difficult to measure, so there Directional distributions of sea states are considered four directional spreading functions: Gaussian, uniform (Top-hat), Cosine-squared, Cosine-2s and the Gaussian theory. Taking the directional spread of waves squared and Cosine-2s distributions.

Directional spreading
Directional wave spectra are a means of describing superimposed sea states, where each sea state has a unique wave number, frequency and direction of propagation. The directional wave spectrum therefore represents the distribution of wave energy in the frequency and direction domain, and is generally expressed as:

\[
E(\omega, \theta) = S(\omega)D(\omega, \theta)
\]

where \(E(\omega, \theta)\) is the directional wave spectrum, \(S(\omega)\) is the frequency spectrum and \(D(\omega, \theta)\) is the angular distribution or directional spreading function. The directional spreading function is dimensionless and is normalised so that:

\[
\int_{-\pi}^{\pi} D(\omega, \theta) \, d\theta = 1
\]

Directional distributions of sea states are inherently difficult to measure, so there remains limited knowledge on the subject (e.g. Goda 2008). Directional distributions can be derived directly from the Fourier coefficients \(a_1, a_2, b_1\) and \(b_2\) without requiring the use of a model (Kuik et al. 1988; Datawell BV 2010). The available data for the present study does not contain the Fourier coefficients, and so models of the directional distributions are applied. In this study we considered four directional spreading functions: Gaussian, uniform (Top-hat), Cosine-squared and Cosine-2s distributions.

Motivation for the use of a Gaussian distribution can be found in the central limit theory. Taking the directional spread of waves at frequency \(\omega\) as the standard deviation of the wave directions \(\sigma(\omega)\), the Gaussian directional spreading function can be expressed as:

\[
D(\omega, \theta) = \frac{1}{\sigma(\omega) \sqrt{2\pi}} e^{-\frac{(\theta - \theta_0(\omega))^2}{2(\sigma(\omega))^2}}
\]

where \(\theta_0(\omega)\) is the mean wave direction for frequency \(\omega\).

A uniform or Top-hat distribution as a spreading function can be used to demonstrate the sensitivity of the directional spectrum to the directional distribution. The Top-hat distribution is defined by:

\[
D(\omega, \theta) = \begin{cases} \frac{1}{\sigma(\omega)} & \text{when } |\theta - \theta_0(\omega)| < \frac{\sigma(\omega)}{2} \\ 0 & \text{otherwise} \end{cases}
\]

The Cosine-squared distribution is perhaps the simplest of the idealised directional distributions. The Cosine-squared distribution is defined as:

\[
D(\omega, \theta) = \frac{1}{\pi} \cos^2(\theta - \theta_0(\omega)) \quad \text{when } |\theta - \theta_0(\omega)| < \frac{\pi}{2}
\]

The Cosine-2s model is widely used due to its simplicity and general effectiveness (Kumara et al. 1999). It was proposed by Longuet-Higgins et al. (1963), and is expressed as:

\[
D(\omega, \theta) = \frac{2}{\Gamma(2s(\omega) + 1)} \frac{\cos^2(\theta - \theta_0(\omega))}{2s(\omega) + 1}
\]

where 
\(G(\omega) = \frac{2^s(\omega)\Gamma^2(s(\omega) + 1)}{2\Gamma^2(2s(\omega) + 1)}\)

and where \(\Gamma(.)\) is the gamma function, and \(s\) is the spreading parameter.

Methods for estimating the spreading parameter \(s\) have been proposed by numerous authors (e.g. Mitsuyasu et al. 1975; Hasselmann et al. 1973; Cartwright 1963; Wang 1992; Goda 2008). In this study we avoid the need for wind data by estimating \(s\) from the first Fourier harmonic, an approach that is widely used (Zhang & Zhang 2006) and yields the result:

\[
s(\omega) = \frac{m_1(\omega)}{1 - m_1(\omega)}
\]

Figure 3 shows plots of the four different directional spreading functions with the directional spread taken as \(\sigma = 30^\circ\). A directional spread of \(30^\circ\) is representative of the average directional spreading for the dominant wave energy frequencies (refer to Figure 10). The plot illustrates that the Cosine-squared, Cosine-2s and the Gaussian distribution are very similar.

**Seasonal distribution of directional wave spectra**

The three data sets were compared both with and without grouping the data into seasons defined using the meteorological convention (refer Table 3). The seasons were compared using difference plots and by computing Pearson and Spearman correlation coefficients.

**Exceedance percentiles of erosive storm events**

Corbella and Stretch (2012a, b) previously found that significant wave heights exceeding 3.5 m were typically associated with major coastal erosion, and they

---

**Table 3 Months associated with each season**

<table>
<thead>
<tr>
<th>Season</th>
<th>Months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td>December – February</td>
</tr>
<tr>
<td>Autumn</td>
<td>March – May</td>
</tr>
<tr>
<td>Winter</td>
<td>January – August</td>
</tr>
<tr>
<td>Spring</td>
<td>September – November</td>
</tr>
</tbody>
</table>
Figure 4 Comparison of the averaged Durban Waverider directional spectra using Gaussian, Top-hat, Cosine-squared and Cosine-2s spreading distributions with difference plots relative to the Gaussian distribution.
therefore used this as a threshold criterion to define a storm event. In a similar fashion the peak storm energies were related to the 3.5 m significant wave height threshold. The associated peak energy threshold was then used as an indication of significant erosive events. In the present study, the energy exceeding this threshold was used to investigate exceedance percentiles in \((\omega, \theta)\) space. These distributions provide valuable information about the energy distribution of erosive storm events, and therefore can have important applications in coastal vulnerability assessments and in the design of coastal works.

**Results**

In this section we start by investigating the effects of refraction between deep water waves and waves recorded by the wave recording buoys. We then compare the different directional spreading distributions, identifying an appropriate one. This is then used to compare the three data sets. The seasonality of the energy spectra is then presented. Finally the \((\omega, \theta)\) distribution of wave energy for erosive storms is presented.

**Refraction of deep water waves**

Since the Durban data is recorded at a depth of 30 m and the Richards Bay data is recorded at a depth of 22 m, there can be significant refraction of the deep-water waves. This is particularly true for the long period waves. A calibrated SWAN model was used to estimate the effects of refraction. The SWAN model parameters are detailed in Table 4. A deep-water significant wave height of 2.5 m was used for all the simulations, while varying the wave period and the wave direction from 10–20 seconds and 30°–180° respectively. The simulation results are limited to the Durban wave recording instruments and are shown in Table 5. For the 21 simulations in Table 5 on average the waves refract by 6° and a maximum of 16°. The largest refraction occurs for deep-water waves that approach from a 30° wave direction. Since Richards Bay’s wave recorder is in shallower water it would not be unreasonable to expect refractions as large as 20 degrees.

![Figure 5 Simulations of a sea state (1 km²) in Richards Bay on 10 March 2004, using the (a) Gaussian, (b) Top-hat, (c) Cosine-squared and (d) Cosine-2s spreading functions. A (e) satellite image (1 km²) is shown alongside the (f) directional wave spectrum using a Gaussian spreading distribution](image)

**Table 4 Parameters used in the SWAN model**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth-induced breaking</td>
<td>Battjes and Janssen (1978)</td>
</tr>
<tr>
<td>Alfa</td>
<td>1.00</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.80</td>
</tr>
<tr>
<td>Bottom friction</td>
<td>Madsen et al (1988)</td>
</tr>
<tr>
<td>Friction coefficient</td>
<td>0.02</td>
</tr>
<tr>
<td>Non-linear triad interaction</td>
<td>Deactivated</td>
</tr>
<tr>
<td>Wind growth</td>
<td>Deactivated</td>
</tr>
<tr>
<td>White-capping</td>
<td>Active (Komen et al 1994)</td>
</tr>
<tr>
<td>Quadruplets</td>
<td>Deactivated</td>
</tr>
<tr>
<td>Refraction</td>
<td>Activated</td>
</tr>
<tr>
<td>Frequency shift</td>
<td>Activated</td>
</tr>
<tr>
<td>Wave forcing</td>
<td>Dissipation rate</td>
</tr>
<tr>
<td>Accuracy</td>
<td>0.02 in 99% of grid points</td>
</tr>
<tr>
<td>Maximum iterations</td>
<td>25.00</td>
</tr>
</tbody>
</table>

**Table 5 Refraction of deep water waves to the location of the Durban wave recording instruments**

<table>
<thead>
<tr>
<th>Direction/Period</th>
<th>30</th>
<th>60</th>
<th>90</th>
<th>120</th>
<th>150</th>
<th>180</th>
<th>210</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>42</td>
<td>66</td>
<td>92</td>
<td>118</td>
<td>150</td>
<td>174</td>
<td>200</td>
</tr>
<tr>
<td>15</td>
<td>46</td>
<td>71</td>
<td>96</td>
<td>118</td>
<td>148</td>
<td>169</td>
<td>199</td>
</tr>
<tr>
<td>20</td>
<td>46</td>
<td>73</td>
<td>98</td>
<td>119</td>
<td>147</td>
<td>167</td>
<td>196</td>
</tr>
</tbody>
</table>
Figure 6 Comparison of the averaged Durban and Richards Bay data and their difference for (a) – (c) the entire data set, (d) – (f) summer, (g) – (i) autumn, (j) – (l) winter, (m) – (o) spring
Comparison of spreading functions
Figure 4 shows the averaged directional spectra for the Durban Waverider data using the Gaussian, Top-hat, Cosine-squared and Cosine-2s directional spreading distributions. Figure 4 also includes the percentage difference plots relative to the Gaussian distribution. The plots demonstrate a very similar distribution, and comparison of all the spectra give Spearman rank correlations near unity. On visual inspection the Gaussian and the Cosine-2s distributions produce the most similar results. They have Spearman and Pearson correlations of 0.99. The difference plots show an absolute difference of 3.7%. The other directional distributions differ from 20% to 60%. The major difference between the distributions is in their tails, while the main body of the distributions are generally similar.

Figure 5 shows simulations of Gaussian waves using the WAFO Matlab tool-box (WAFO Group 2011). The simulations are of recorded data from 10 March 2004 using the four spreading functions (a) Gaussian, (b) Top-hat, (c) Cosine-squared and (d) Cosine-2s. The simulations are compared to (e) a Google Earth satellite image. The Gaussian, Cosine-squared and Cosine-2s simulations are almost indistinguishable and look similar to the satellite image, while the Top-hat distribution does not replicate the appearance of a real sea state. We conclude that the Gaussian, Cosine-squared and Cosine-2s distributions are all able to model a real sea state. There is no generally accepted standard functional form for directional spreading in the literature. The results using the Gaussian distribution are very similar to those of the widely used Cosine-2s distribution (Figure 3). Since the statistical theory underpinning the Gaussian distribution is widely familiar, we adopted it for our evaluation of the directional wave spectra characteristics.

Comparison of Durban and Richards Bay
Corbella and Stretch (2012a) claimed that the Durban and Richards Bay data was sufficiently similar to be used interchangeably. However, they did not consider the directional information since only six years of data is available for comparison. Furthermore, comparisons based on averaged statistics from short data sets are not definitive, because it has been found that a minimum of 50 years of data is required to accurately estimate a wave climate and obtain a representative average (WASA Group 1998).

Averaged directional spectra for both Richards Bay and Durban are shown in Figure 6. Both locations have similar energy distributions between 90° and 180° and similar peak energies between 135° and 180°. However, it is evident from the difference plots that Durban has stronger northeast-erly energy components. Richards Bay is approximately 160 km away from Durban.
and any meteorological forcing between the two locations will result in a northerly component for Durban and a southerly component for Richards Bay. The existence of such events may explain why Durban has a stronger northerly component and Richards Bay has a higher peak energy in the south. This may also be a function of the coast’s orientation that allows a longer northerly fetch at Durban. Much of the northeasterly energy is contained around relatively high frequencies (greater than 0.1 Hz). These frequencies are usually associated with locally generated wind waves. Figure 7 shows wind roses for Durban and Richards Bay. Durban winds are more frequent and stronger from the northeast than the Richards Bay winds. The additional northeasterly wave energy in Durban can therefore be attributed to local wind conditions. Corbella and Stretch (2012a) correctly stated that the Durban and Richards Bay data is interchangeable in terms of significant wave height and wave period. However, considering directional energy spectra, Richards Bay may not provide enough northeasterly wave energy to use in place of the Durban data for design applications and coastal vulnerability assessments.

Comparing data from different instruments

Richards Bay has five and a half years of directional data from a directional GPS 3D buoy, and more recently ten and a half years from a directional Waverider buoy (Table 2). These two recording instruments do not overlap, and so it is not possible to do a direct comparison. Figure 8 shows a comparison of the averaged directional spectra of the two wave recording instruments. Since these averages show similar directional spectra, and the recording instruments were in the same location, it is acceptable to combine the data sets into a 16-year record.

Richards Bay seasonality

All of Richards Bay’s directional spectral data were combined into a 16-year record to produce the average seasonal spectra plotted in Figure 9. In addition we present the average directional spread per frequency (Figure 10). There is little seasonality in the Richards Bay data. Figure 10 shows that the seasons have different magnitudes of variability and may be placed in descending order of variability as summer, spring, autumn and winter. The energy is spread over the widest range of directions during spring and summer. This is due to the absence of large swell waves in these months, and these months also being the windiest months of the year. These two factors result in the average distributions being influenced by the predominant south-westerly and northeasterly winds. Autumn and winter do not show as much variability, but have more energy due to the presence of southeasterly storms during these months.
The determination of wave origins can only be made with the aid of directional wave spectra. These exceedance percentiles for potentially erosive wave energy are clearly important for coastal vulnerability assessments.

**DISCUSSION**

The available directional data in the case study region only totals 16 years, and therefore cannot be considered as fully representative of the wave climate (WASA Group 1998). Nevertheless, while cognisant of the limitations in the short data record, several key features are already clearly evident.

Firstly there is no evidence of significant northeasterly wave energy, with the exception of locally generated wind waves in Durban. It therefore seems apparent that tropical cyclones contribute very little to northeasterly storm waves outside of summer. It is expected that only cyclones entering the Mozambique Channel are capable of producing storm waves from the northeast and will generally only contribute to east-south-easterly wave energy, as is evident from Figure 13. It must be remembered that, due to wave refraction, many of the waves measured as east-south-east may be east-north-east deep-water waves. These waves, particularly in summer and autumn, may be evidence of tropical cyclones. The majority of waves, including the largest energy contributions, come from the southeast and seem to be associated with a combination of cold fronts and cut-off lows.

The determination of wave origins can only...
be achieved by investigating the forcing mechanisms.

Secondly, the wave energy spectra show some seasonal variations. Summer storm waves are produced in the south-southeast and the east-southeast. The occurrence of waves are produced in the south-southeast some seasonal variations. Summer storm mechanisms.

Figure 13

Tropical cyclone tracks (eye of cyclones) in the western Indian Ocean between 1985 and 2005 (adapted from Wikimedia Commons)

In this paper we have presented a formal analysis of directional wave spectra data from the east coast of South Africa. Measurements from two locations about 160 km apart were analysed and found to be generally similar, but with some differences in the directional energy distribution – the Durban data has more northeasterly energy than Richards Bay. Various spreading functions have been investigated, and the Cosine-2s and Gaussian distributions produce similar results and are representative of directional spreading along the east coast of South Africa. There is evidence of seasonality in the directional wave spectra. Summer and spring have the least energy and the most directional variability, while winter has the most energy with the majority of it focused from the southeast. There is no clear evidence of tropical cyclones contributing significant northeasterly wave energy. For design and vulnerability considerations the southeasterly wave energy is the most important.

ACKNOWLEDGEMENTS

The authors thank the eThekwini Municipality, the National Ports Authority and the CSIR for access to the directional spectral wave data. Author Corbella is supported by a post-doctoral fellowship from the Nelson Endowment Fund in the UKZN School of Engineering, which is acknowledged with thanks. Author Stretch is grateful to the eThekwini Municipality for their support of a Sponsored Chair in Civil Engineering at UKZN.

REFERENCES


**BIBLIOGRAPHY**

Review of pump suction reducer selection: Eccentric or concentric reducers

R M Mahaffey, S J van Vuuren

Eccentric reducers are traditionally recommended for the pump suction reducer fitting to allow for transportation of air through the fitting to the pump. The ability of a concentric reducer to provide an improved approach flow to the pump while still allowing air to be transported through the fitting is investigated. Computational fluid dynamics (CFD) were utilised to analyse six concentric and six eccentric reducer geometries at four different inlet velocities to determine the flow velocity distribution at the inlet to the pump. It was found that eccentric reducers with angles greater or equal to 15° and concentric reducers with an angle greater or equal to 20° did not pass the assessment criteria related to the inlet conditions. Air could be hydraulically transported through all of the concentric reducers modelled except for the 20° concentric reducer. A correctly designed concentric reducer will not only provide a more uniform velocity distribution in comparison to an eccentric reducer, but will allow for the hydraulic transportation of air through the reducer.

INTRODUCTION

The design of the pump inlet pipework defines the resulting hydraulic conditions experienced at the pump inlet/impeller. The importance of the pump inlet condition is often overlooked and, according to Jones et al (2008) is likely to be the single reason mostly responsible for the failure of pumps. Failure of the pump inlet design to produce a uniform velocity distribution at the pump inlet can lead to noisy operation, random axial load oscillations, premature bearing or seal failure, cavitation damage to the impeller and inlet portions of the casing, and occasional damage on the discharge side due to liquid separation (ANSI 2009). Typical components of pump inlet pipework (bends, valves, spool pieces, reducers, etc) create flow conditions that are prone to a non-uniform velocity distribution.

Reducer fittings used in pump inlet pipework are divided into two types: concentric reducers and eccentric reducers. Design guidelines, pump operating manuals and design standards prescribe which type of reducer to use for various conditions. An eccentric reducer with the flat side on top is typically the prescribed reducer to be used. This reducer prevents air to accumulate at the upstream end of the reducer. In a case study presented by Van der Westhuizen (2011) the incorrect reducer selection was noted as one of the causes of a reported pump failure; in this case study the reducers installed in the pump inlet pipework were eccentric reducers. In the failure investigation, field pressures in the installed eccentric reducer fitting were measured at eight circumferential locations towards the downstream end of the fitting. These pressure measurements presented a non-uniform pressure distribution within the reducer, with the greatest difference being 11 kPa (65% of the largest pressure reading). This contributed to the reported pump failure’s pressure distribution is caused by an acceleration in velocity along the bottom edge (with the flat side on top) of the eccentric reducer as the flow path narrows from below.

The geometry of an eccentric reducer is asymmetrical, and asymmetrical flow conditions will result, introducing non-uniform velocity and pressure distributions at the pump inlet. These flow conditions are contradictory to the recommended pump inlet designs, which recommend symmetrical flow conditions with a uniform velocity and pressure distribution. The flow conditions resulting from a concentric reducer will be more uniform compared to the conditions resulting from an eccentric reducer. Prescribing eccentric reducers is based on the perception of the transportability of air.

Various relationships are used to evaluate the hydraulic transportability of air. The capacity for a pipeline to hydraulically transport air is a function of the velocity of the liquid, the diameter of the pipeline and the slope of the pipeline (Van Vuurn et al 2004). If air can be hydraulically transported in a pipeline, why will air not have the ability to be transported through a concentric reducer?

Keywords: pump station design, eccentric and concentric reducers, computational fluid dynamics (CFD), numerical modelling, air transport in pipelines
The velocity distributions resulting from various eccentric and concentric reducer designs were studied with the use of computational fluid dynamics (CFD) under various flow velocities. The velocity distributions were assessed according to criteria prescribed by ANSI (1998) and the guidelines provided by Wallingford (2001) in Sinotech CC (2005), and Van Vuuren (2011). With the CFD analyses, together with the theoretical assessment of the hydraulic transportability of air through the reducers, a design recommendation on the selection of reducer geometry can be made.

An introduction to pump inlet design and a theoretical overview into the hydraulic transportation of air are provided in the next section of this paper. The setup of the CFD models and the selection of the reducer geometries are then presented. The paper concludes with a brief presentation of the results and recommendations.

**PUMP INLET DESIGN – SELECTION OF REDUCER TYPE**

Reducers are typically placed immediately upstream of a pump in order to operate the suction pipework at acceptable energy losses (NPSH available). Figure 1 shows a schematic representation of the inlet side of typical parallel centrifugal pumps.

The selection of the type of reducer to be installed on the inlet side of a pump has an influence on the hydraulic transportation of air within the suction pipework and the flow conditions exiting the reducer and entering the pump. Numerous references prescribe the type of reducer to be used in a suction pipework – a summary of these design requirements is presented in Table 1. From this table it is observed that the eccentric reducer with the flat section on top is prescribed for all cases where the suction pipework enters in a horizontal plane. This prescription is to allow for the hydraulic transportation of air through the reducer by ensuring that an air pocket does not accumulate at the upstream end of the reducer, as illustrated in Figure 2. However, in the event of no potential for air accumulation or vertical inlet pipes, a concentric reducer is prescribed by ANSI (2009).

---

**Table 1 Summary of reducer design requirements**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reference type</th>
<th>Reducer selection</th>
<th>Conditions for selection</th>
<th>Reason for reducer selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSI (2009)</td>
<td>National standard</td>
<td>Concentric reducer</td>
<td>Vertical inlet pipes</td>
<td>Ensures no air accumulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Installations where there is no potential for vapour accumulation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eccentric reducer</td>
<td>Horizontal installations where air accumulation is possible</td>
<td>Allows for transportation of air</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flat section on top</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minimum straight lengths of pipe are required</td>
<td>Ensures disturbance-free flow</td>
</tr>
<tr>
<td>Bloch (2010)</td>
<td>Journal article</td>
<td>Eccentric reducer</td>
<td>Flat section on top for suction pipe entering in a horizontal plane</td>
<td>Allows for transportation of air</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flat section on top for suction pipe entering from below</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flat section at the bottom for suction pipe entering from above</td>
<td>Ensures no air accumulation</td>
</tr>
<tr>
<td>Mackay (2004)</td>
<td>Pump station handbook</td>
<td>Eccentric reducer</td>
<td>Flat section on top</td>
<td>Eliminates the potential problem in eddy currents in a high point in the suction line that might travel into the impeller eye</td>
</tr>
<tr>
<td>Bloch &amp; Burdis (2010)</td>
<td>Pump station handbook</td>
<td>Eccentric reducer</td>
<td>Flat section on top</td>
<td>Allows for transportation of air</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Minimum straight pipe and reducer length required</td>
<td>Ensures disturbance-free flow</td>
</tr>
<tr>
<td>KSB (2012)</td>
<td>Pump manufacturer operating instructions</td>
<td>Eccentric reducer</td>
<td>Flat section on top</td>
<td>Allows for transportation of air</td>
</tr>
<tr>
<td>Jones et al. (2008)</td>
<td>Pump station handbook</td>
<td>Eccentric reducer</td>
<td>Flat section on top</td>
<td></td>
</tr>
<tr>
<td>SAPMA (2002)</td>
<td>Pump station handbook</td>
<td>Eccentric reducer</td>
<td>Flat section on top</td>
<td></td>
</tr>
</tbody>
</table>

![Figure 1 Typical pump inlet design components](image1)

![Figure 2 Design requirements allowing for air transport](image2)
The hydraulic capacity of the system

The origin of air at this position in the

aspects are not addressed:

the reducer. However, the following three

top will aid in the transport of air through

of an eccentric reducer with the flat side on

providing by ANSI/HI.

A reason for this preference is, however, not

sent in the system, the air may be released

after maintenance where air may be pre-

start-up conditions after construction or

configuration will produce asymmetrical flow

metrical, and flow through a fitting of this

the pump (fully open, non-flow disturbing valves, vaned elbows, long-radius elbows and

reduces are not considered as flow-disturbing fittings)

The suction diameter is to be at least one size larger than the suction fitting on the

pump, in such cases an eccentric or concentric reducer fitting is fitted

Table 2 A summary of recommended maximum velocities for pump inlet piping

<table>
<thead>
<tr>
<th>Reference</th>
<th>Reference type</th>
<th>Recommended maximum velocity (m/s)</th>
<th>Conditions for recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANSI/HI 9.6.6 (2009)</td>
<td>National standard</td>
<td>2.4</td>
<td>Suction pipe at least as large as the pump suction connection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Valves and other flow-disturbing fittings are to be at least one pipe size larger than the pump suction nozzle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Excludes slurries</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Values greater than 2.4 m/s are to be evaluated with respect to flow distribution, erosion, NPSH, noise, water hammer and the manufacturer’s recommendations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Maximum velocity applies to any point in the suction piping</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For fluids close to the vapour, pressure velocities are to be kept low enough to avoid flashing of the liquid in the piping</td>
</tr>
<tr>
<td>ANSI (1998)</td>
<td>National standard</td>
<td>2.4</td>
<td>Velocities may be increased at the pump suction flange by the use of a gradual reducer</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Higher velocities are acceptable, provided the piping design delivers a smooth inlet flow to pump suction (velocity distributions are specified)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The velocity in the pump suction piping should be constant or increasing as the flow approaches the pump</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The velocity is to be large enough to prevent sedimentation in horizontal piping</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>There shall be no flow-disturbing fittings closer than five suction pipe diameters from the pump (fully open, non-flow disturbing valves, vaned elbows, long-radius elbows and reducers are not considered as flow-disturbing fittings)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The suction diameter is to be at least one size larger than the suction fitting on the pump, in such cases an eccentric or concentric reducer fitting is fitted</td>
</tr>
<tr>
<td>Mackay (2004)</td>
<td>Pump station handbook</td>
<td></td>
<td>The suction diameter is to be at least one size larger than the suction fitting on the pump</td>
</tr>
<tr>
<td>SAPMA (2002)</td>
<td>Pump station handbook</td>
<td>0.75</td>
<td>The suction pipe should never be smaller than the suction connection of the pump and in most cases should be at least one size larger</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pipe velocities should be in the range of 0.5 m/s to 0.75 m/s unless suction conditions are exceptionally good</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Elbows or tees located adjacent to the pump suction flange will result in uneven flow patterns which prevent the liquid from filling the impeller evenly</td>
</tr>
<tr>
<td>Bloch &amp; Burdis (2010)</td>
<td>Pump station handbook</td>
<td>2.5</td>
<td>Maximum velocity applies to any point in the suction piping</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Suction pipe at least as large as the pump suction nozzle</td>
</tr>
<tr>
<td>KSB (2012)</td>
<td>Pump manufacturer</td>
<td>2</td>
<td>For an individual suction lift pipe</td>
</tr>
<tr>
<td></td>
<td>operating instructions</td>
<td></td>
<td>If it is for a suction manifold, the flow velocity is to be kept as low as practically possible</td>
</tr>
<tr>
<td>Jones et al (2008)</td>
<td>Pump station handbook</td>
<td>1.8 ~ 2.4</td>
<td>The conditions in ANSI 2009 and ANSI 1998 are to be followed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The velocity is not applicable to suction manifolds where the recommendation is 1.8 m/s</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The velocity applies to constant speed pumps</td>
</tr>
</tbody>
</table>

A reason for this preference is, however, not provided by ANSI/HI.

The philosophy of prescribing the use of an eccentric reducer with the flat side on top will aid in the transport of air through the reducer. However, the following three aspects are not addressed:

- **The origin of air at this position in the pumping system.** If a pump station sump, fore bay, intake bay, suction pipework and reservoir outlets are correctly designed it is impossible for air to accumulate. Pumps typically operate under positive suction pressure heads (when started) and therefore air should not be able to enter the system on the suction side. Under first-time start-up conditions after construction or after maintenance where air may be present in the system, the air may be released with a manual bleed-off valve.

- **The hydraulic capacity of the system to transport air.** Trapped air in pipelines can be hydraulically transported along the length of a pipeline through valves and fittings depending on the pipeline slope, flow rate and fitting geometry. A concentric reducer with the correct flow rate and geometry will allow air to be hydraulically transported through it.

- **The non-uniform velocity distribution on the impeller created by the flow through the eccentric reducer.** The geometry of an eccentric reducer is asymmetrical, and flow through a fitting of this nature will produce asymmetrical flow conditions, such as non-uniform velocity and pressure distributions. These flow conditions are contradictory to the recommended pump inlet designs, which recommend symmetrical flow conditions with a uniform velocity and pressure distribution.

**PUMP INLET PIPING REQUIREMENTS**

Pump inlet piping requirements are typically prescribed according to a maximum allowable velocity in the suction pipework and valves, and an allowable velocity distribution at the pump inlet. A summary of recommended maximum velocities for pump inlet piping as identified in a literature review is provided in Table 2. It can be observed that acceptable velocities for pump suction pipework are typically in the range of 0.75 m/s to 2.5 m/s, and the suction pipework is to be at least one size larger than the pump inlet nozzle, thereby defining the use of a reducer.

The requirements for maximum velocity distributions are to limit the radial thrust on the pump’s shaft and couplings. The phenomenon of radial thrust is illustrated in Figure 3, where the pressure/velocity differential between positions P1 and P2 causes an additional imbalance of forces on the pump’s impeller. The bearings, shafts and casings of a pump are designed to resist the anticipated radial load, but pumps have been destroyed within days, or even a few hours, through excessive wear in bearings, seals, shaft sleeves and wearing rings caused by excessive shaft deflections resulting from radial thrust (Jones et al 2008). The ideal is therefore that...
the approach flow to a pump should be an undisturbed flow – free from unequal velocity distributions, unequal pressure distributions, entrained air or gas bubbles, vortices and excessive pre-swirl at the pump inlet – to minimise the unbalanced forces directed onto the impeller in the axial direction (ANSI 1998; Sulzer 2010; Jones et al 2008).

**HYDRAULIC TRANSPORTABILITY OF AIR**

**Source of air in pipelines**

Air can enter a pipeline, and specifically on the suction end of a pump, in the following manners (Van Vuuren et al 2004):
- Air present in the pipes due to first time start-up or start-up after maintenance
- Air released from solution at sufficiently low pressures
- Air that enters the pipework due to insufficient seals and/or faulty connections
- Air entering the pipe from a free surface due to an incorrect design and/or incorrect operating conditions.

This air present in the pipeline negatively affects pumps and pumping systems, and these negative effects include:
- A reduction in the water cross-sectional area (Jones et al 2008, ANSI 2009)
- An increase to the flow resistance (Jones et al 2008, ANSI 2009)
- Possible loss of prime to the pump if a slug of air is swept into the pump case during a restart, causing a partial or complete loss of prime (ANSI 2009)
- The creation of an environment conducive to corrosion (Jones et al 2008)
- Creation of turbulence and air entrainment (Mackay 2004)
- Implosion of entrained air due to increasing pressure at the eye of the impeller causing damage identical to that of cavitation (Mackay 2004)
- Noisy operation (Bloch & Burdis 2010)
- Unbalanced axial loads (Bloch & Burdis 2010).

It is therefore imperative to hydraulically transport air through the inlet pipework and pump, and not allow an accumulation of air in the upstream pipework of the pumping system.

**Fundamental equations for the hydraulic transport of air**

The hydraulic transportability of air refers to the ability of a fluid flowing within a conduit to transport free air in the direction of the flow. This free air is then normally transported within the conduit where the fluid cannot transport the free air any further. The minimum fluid velocity required to transport free air within a conduit is referred to as the critical velocity (Van Vuuren et al 2004). The hydraulic transportability of air within a pipeline is a function of the pipe slope, the amount of accumulated air and the flow characteristics of the fluid within the pipeline (i.e. Froude number, flow velocity and flow rate) (Pothof & Clemens 2011; Van Vuuren et al 2004).

Various relationships for the assessment of the hydraulic transportability of air have been developed by various researchers. According to Van Vuuren et al (2004), the two most widely used relationships to calculate the critical velocity are those presented by Kalinske and Bliss (1943) (Equation 1), and by Wisner et al (1975) (Equation 2). Van Vuuren et al (2004) also derived their own relationship to calculate the critical velocity (Equation 3).

\[
\frac{Q_c^2}{gD^5} = 0.707\sin\theta
\]

\[
\frac{V_c}{\sqrt{g\theta d}} = 0.25\sqrt{\sin\theta} + 0.825
\]

\[
V_c = a\sqrt{gD\theta^b}
\]

Where:
- \( Q_c \) = Critical flow rate (m³/s)
- \( V_c \) = Critical velocity (m/s)
- \( g \) = Gravitational acceleration (m/s²)
- \( D \) = Diameter of the pipe (m)
- \( \theta \) = Slope of the pipeline (°)
- \( a \) = Constant (0.2178)
- \( b \) = Constant (0.4007)

These relationships can be utilised to evaluate the capacity of a pumping system to transport air through a concentric reducer, as it will represent flow conditions similar to that of a sloped pipe. The slope of the pipeline utilised in Equations 1 to 3 for a concentric reducer was calculated with the methodology described in Figure 4.

**NUMERICAL MODEL (CFD SETUP)**

**Introduction**

Computational fluid dynamics (CFD) is a defined set of methods that are used to numerically solve the governing laws of fluid motion in or around a material system, where its geometry is also modelled (Hirsch 2007). CFD has recently experienced remarkable growth in its application, and has been specifically elevated with the growth of the computational power of parallel computer processors, so that it now has a defined position in the parallel experimental investigations of fluid dynamic problems and research (Strum 2010).

The commercially available CFD software produced by CD-adapco: STAR-CCM+ (Version 8.04.007) was utilised to study the velocity distributions resulting from four velocity inputs in six concentric and six eccentric reducer geometries. Each of the simulations was run for 1 000 iterations, which were found to provide sufficient convergence and stable results.

**Geometry and flow selection**

The application of a range of constant velocities as inputs in the CFD calculations

---

**Figure 3 Development of radial thrust in a centrifugal pump**

**Figure 4 Definition of the reducer geometries**
allows for a single reducer dimension to be modelled, i.e. the flow patterns resulting from a 1 m/s flow rate through a DN 200 to DN 150 reducer will be representative of the flow patterns through a DN 2000 to DN 1500 reducer if the reducer angles are equal. The DN 200 to DN 150 reducer was selected as the representative size reducer to be modelled with varying reducer angles (reducer transition lengths). With an increasing reducer angle (shorter reducer transition length) the uniformity of the pressure and velocity distributions will decrease and the hydraulic transportability of air through the concentric reducer will decrease.

The reducer angles were defined to represent the typical range of reducer geometries as defined by local fitting manufacturers and the guidelines presented in AWWA C208 (2008). The definition of the reducer geometries is illustrated in Figure 4, and the geometries modelled are summarised in Table 3. In order to provide flow information resulting from the application of additional downstream straight lengths of pipe requirements, as described by ANSI/HI (2009), and to model the effect of the discontinuities at the upstream and downstream ends of the reducer, a straight pipe length of 600 mm (3 x D1) was added upstream and downstream to all the reducer geometries. The 3D reducer geometries were created in Autodesk’s Inventor and imported into STAR-CCM+ as Parasolid text files (*.x_t).

The maximum velocity in the pump inlet pipework is typically recommended to be 2.4 m/s (see Table 2). By using the maximum velocity in the study, the non-uniform flow profiles will be exaggerated while the hydraulic transportability of air will be at its maximum. A velocity of 1 m/s is often used in the industry and this velocity will therefore provide velocity distributions typically experienced, and the capacity of flow to hydraulically transport air through the fitting will be lower with this velocity than with a velocity of 2.4 m/s. Other flow velocities were selected ranging from 1 m/s to the maximum velocity of 2.4 m/s (1 m/s, 1.5 m/s, 2 m/s and 2.4 m/s) to provide a complete set of results.

**CFD input parameters**

**Model selection**

The CR 10 model was selected to be used as the model where the CFD input parameters, such as meshing models and physics values, were tested to produce a numerical model that accurately represents the fluid flow problem and that could be solved within an acceptable period of time with adequate solution stability. After an iterative process the CR 10 model provided an acceptable solution, and the remaining 11 reducer geometries were set up with the final CR 10 model parameters. The significant inputs for the numerical models are described below (the numerous input parameters that were not changed, are not discussed).

**Meshing models**

STAR-CCM+ provides various meshing models that can be selected to suit the specific problem. The following meshing models were selected in this instance:

- **Surface remesher.** The surface remesher re-triangulates a closed starting surface to obtain a higher quality surface (CD-adapco 2012).
- **Polyhedral mesher.** Polyhedral type cells (14 cell faces on average) produce a more accurate solution when compared to a tetrahedral mesh (4 cell faces) (CD-adapco 2012).

![Figure 5 CR 10 polyhedral and prism layer mesh representation](Image)

**Table 4 Input parameters for the meshing models**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polyhedral base size</td>
<td>5 mm</td>
</tr>
<tr>
<td>Tet/Poly density</td>
<td>1.0 (default value)</td>
</tr>
<tr>
<td>Polyhedral growth factor</td>
<td>1.0 (default value)</td>
</tr>
<tr>
<td>Number of prism layers</td>
<td>10</td>
</tr>
<tr>
<td>Prism layer stretching</td>
<td>1.5 (default)</td>
</tr>
<tr>
<td>Relative size of the prism layer thickness</td>
<td>33.3% of base size (default)</td>
</tr>
</tbody>
</table>

**Physics input parameters**

The physics models and input parameters were selected to best suit the flow problem. Brief descriptions of the Star-CCM+ physics input parameters follow:

- **Space.** The space model selected is the three-dimensional space model.
- **Time.** The time physics model controls the iteration and/or time stepping of the solver. The steady time model was selected.
- **Material.** The material model controls the definition of the substances being analysed, liquid in the case of this study.
- **Flow.** Segregated flow was selected to represent the incompressible water flow.
- **Viscous regime.** The flow being modelled is viscous and is in the turbulent zone. Therefore the turbulent viscous regime was selected. By selecting the turbulent viscous regime, the Reynolds-Averaged Navier-Stokes (RANS) turbulence model was auto-selected.
- **Equation of state.** The equation of state model is used to compute the density and density derivatives relating to temperature and pressure. Temperature does not form part of this study, and the change density of the water will be insignificant due to the pressure loss experienced in the pump inlet.

![Table 3 Geometries of the reducers modelled](Image)

<table>
<thead>
<tr>
<th>Geometry name</th>
<th>Reducer type</th>
<th>D1 (mm)</th>
<th>D2 (mm)</th>
<th>θ (°)</th>
<th>L (mm)</th>
<th>t (mm)</th>
<th>ID1 (mm)</th>
<th>ID2 (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER 30</td>
<td>Eccentric reducer</td>
<td>30.0</td>
<td>94</td>
<td>4.5</td>
<td>210.1</td>
<td>156.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER 20</td>
<td>Eccentric reducer</td>
<td>20.0</td>
<td>148</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER 15</td>
<td>Eccentric reducer</td>
<td>15.0</td>
<td>202</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER 10</td>
<td>Eccentric reducer</td>
<td>10.0</td>
<td>306</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER 5</td>
<td>Eccentric reducer</td>
<td>5.0</td>
<td>617</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER 2.5</td>
<td>Eccentric reducer</td>
<td>2.5</td>
<td>1237</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 20</td>
<td>Concentric reducer</td>
<td>20.0</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 15</td>
<td>Concentric reducer</td>
<td>15.0</td>
<td>101</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 10</td>
<td>Concentric reducer</td>
<td>10.0</td>
<td>153</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 5</td>
<td>Concentric reducer</td>
<td>5.0</td>
<td>309</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 2.5</td>
<td>Concentric reducer</td>
<td>2.5</td>
<td>618</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CR 2</td>
<td>Concentric reducer</td>
<td>2.0</td>
<td>773</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
piping. The constant density model was selected to model the equation of state.

- Reynolds-Averaged turbulence. A model needs to be selected to provide closure to the RANS equations. The K-Epsilon model was selected for this study, as it provided a lower computational cost and was well suited to the flow properties of the study.

- K-Epsilon model. STAR-CCM+ provides a choice of seven different K-Epsilon models. The selection of the K-Epsilon turbulence model is determined by the wall treatment (near-wall modelling assumptions), the Reynolds number and the coarseness or the fineness of the mesh. The standard two-layer K-Epsilon model and the realisable two-layer K-Epsilon model offer the most mesh flexibility. They can be utilised on a wide range of $y^+$ values (all – $y^+$). The two-layer all $y^+$ wall treatment and realizable K-Epsilon two-layer were automatically selected on selection of the K-Epsilon turbulence model. This selection corresponds to the guidelines described above and was not changed in the study.

**Boundary conditions**

Appropriate boundary conditions that represent the real physical fluid flow conditions are required to be selected in the CFD software. The boundary types selected to represent the fluid flow problem are velocity input, pressure outlet and fluid wall. The locations of these boundary conditions are illustrated in Figure 6. Descriptions and input parameters for the selected boundary types are provided below:

- **Velocity inlet.** This boundary type was selected for the inflow boundary (upstream end) of the model and represents normal flow conditions entering the pipe perpendicularly, with a constant velocity that is supplied as an input parameter.

- **Pressure outlet.** This boundary type was selected for the outflow boundary (downstream end) and represents a uniform constant pressure at the outlet that is supplied as an input parameter. The pressure was set to 150 000 Pa (150 kPa) to provide a positive pressure for the model.

- **Wall.** A wall boundary represents an impermeable surface and is applicable as an impermeable boundary for inviscid flows and as an impermeable, non-slip boundary for viscous flow simulations (CD-adapco 2012). A roughness value (0.06 mm) was added to the fluid wall to enable near-wall turbulence and thereby apply friction to the model in order to accurately model the pipework. The roughness value utilised was obtained from Wallingford and Barr (2006) for an epoxy-lined steel pipe in a normal condition (typically the type of lining used for fittings in water pump stations in South Africa).

**CFD visualisation and reporting**

In order to visualise the magnitude of the velocity distribution through the pump inlet pipework, two scalar scenes were created. The first scene was created on the X-Y plane at the downstream end of the reducer – this scene allows for the visualisation of the simulation velocity results. The second scene was created on the Y-Z plane to visualise the propagation of the velocity differential through the reducer. In addition to the scalar scenes, X-Y plots of the velocity were generated along both the x-axis and y-axis of the pipe at four different positions (probe positions). The four probe positions are defined with Probe Position 1 starting at the downstream end of the reducer and repeated at intervals of 1 x $D_2$ (downstream diameter), ending at Probe Position 4. XY velocity plots along two circles with diameters of 0.8 x $ID_2$ and 0.6 x $ID_2$ at Probe Position 1 were also
created. The orientation of the scalar scenes, the locations of the four probe positions and the orientation of the x-axis, y-axis and z-axis are illustrated in Figure 7.

Limitations
In the analyses, uniform flow conditions were provided at the inlet of the geometry; in engineering applications this may seldom be the case. These non-uniform flow conditions may influence the velocity distribution in such a way that a concentric or eccentric reducer within the recommended range may produce a velocity distribution that is outside of the acceptance criteria, which needs to be quantified when designing the inlet pipework to a pump.

Assessment Criteria
Further to the velocity and approach flow requirements provided in Table 2, assessment criteria were defined in order to compare the performance of the various reducer geometries. ANSI (1998) and Jones et al. (2008) specified the requirement for time-averaged velocities at the pump suction to be within 10% of the cross-sectional area average velocity. Wallingford (2001) in Sinotech CC (2005), and Van Vuuren (2011) provided the following guidelines for velocity distribution variations in suction pipework:

- Velocity variation along line ABCD is less than 10% of the average velocity (see Figure 8).
- Maximum velocity variation along a circle AEDF is ± 5% of the average velocity (see Figure 8).

The requirement described by Wallingford (2001) in Sinotech CC (2005), and Van Vuuren (2011) did not indicate the location of the circle AEDF. Two circles were therefore defined (0.6 × D2 and 0.8 × D2) to assess the sensitivity of the location of the circle. The assessment criteria were labelled and the final assessment criteria used are summarised in Table 5.

Results and Discussion
The results that were obtained are reflected under the following headings:

- Scalar scene results
- Results – Criterion 1 and Criterion 3 y-axis
- Results – Criterion 2
- Results – Criterion 3 x-axis
- Results – Hydraulic transportability of air
- Results matrix

Scalar scene results
The scalar scenes generated through the study allow for the development of the velocity distribution through the reducer to be visualised.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Definition of average velocity</th>
<th>Acceptance criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 1</td>
<td>Average velocity along line ABCD</td>
<td>Velocity variation along line ABCD is less than 10% of the defined average velocity</td>
</tr>
<tr>
<td>Criterion 2 HRW 1</td>
<td>Average velocity along circle AEDF</td>
<td>Velocity variation along circle AEDF is less than 5% of the defined average velocity</td>
</tr>
<tr>
<td>Criterion 2 HRW 2</td>
<td>Average velocity along circle BGCH</td>
<td>Velocity variation along circle BGCH is less than 5% of the defined average velocity</td>
</tr>
<tr>
<td>Criterion 3 x-axis</td>
<td>Cross-sectional area average velocity</td>
<td>Velocity variation along line EGHF (x-axis) is less than 10% of the defined average velocity</td>
</tr>
<tr>
<td>Criterion 3 y-axis</td>
<td>Cross-sectional area average velocity</td>
<td>Velocity variation along line ABCD (y-axis) is less than 10% of the defined average velocity</td>
</tr>
</tbody>
</table>
Four scalar scenes for an input velocity of 1.5 m/s are provided as examples in Figure 9 and Figure 10 to illustrate the difference in velocity distributions resulting from the different reducer types and geometries. The main observations resulting from the scalar scenes are:

- The downstream flow distribution evens out rapidly and presents a typical uniform distribution by the time it reaches a distance of 1 x D2 from the reducer (Probe Position 2).
- The velocity distribution for the concentric reducers produce velocity distributions that are symmetrical around the centre of the pipe, i.e. the velocity contours are circular in shape.
- The eccentric reducers produce areas of high velocities towards the bottom (sloped) side of the reducer, which result in non-uniform velocity distribution at the downstream end of the reducer (Probe Position 1).
- The magnitude of variance in the velocity distributions increases with an increase in reducer angle.
- In the case of the reducers with larger angles, the resulting velocities may be of such a nature that they will fall outside of the acceptance criteria.

**Results – Criterion 1 and Criterion 3 y-axes**

Criterion 1 requires the velocity variation along line ABCD to be less than 10% of the average velocity along line ABCD. Criterion 3 y-axis requires the velocity variation along line ABCD (y-axis) to be less than 10% of the cross-sectional average velocity. The average velocities at the downstream end of the reducer (at ID2) were calculated and the average velocities along line ABCD (y-axis) at Probe Position 1 were calculated for all the simulations. The average velocities along line ABCD are within 0.9% (0.4 m/s) of the average velocity. Therefore the Criterion 1 and Criterion 3 y-axes were assumed to be equal for the assessment, and the assessment was conducted on the cross-sectional area average velocity. The results for Probe Positions 2, 3 and 4 are not provided due to the uniformity in the velocity distribution at these positions, as illustrated in Figure 9.

The results for all four velocities presented similar outcomes. The upper end of the eccentric reducers (ER 15, ER 20 and ER 30) and the CR 20 model concentric extend past the upper limit of the acceptance envelope and therefore do not pass acceptance Criterion 1 and Criterion 3 y-axes. The remaining concentric reducers (CR2, CR 2.5, CR 5, CR 10 and CR 15) and eccentric reducers (ER 2.5, CR 5 and ER 10) all fall...
The results for the velocity distributions along the y-axis (line ABCD) for the 1.5 m/s velocity are provided as an example of the results obtained for Criterion 1 and Criterion 3 y-axes in Figure 11.

### Results – Criterion 2

Criterion 2 HRW1 requires the velocity variation along circle AEDF to be less than 5% of the average velocity along circle AEDF, and Criterion 2 HRW2 requires the velocity variation along circle BGCH to be less than 5% of the average velocity along circle BGCH.

The velocities along the circumference of the two circles are plotted on the y-axis and an angle relating to the position of the cylinder is plotted on the x-axis. This angle is a component of the coordinate system illustrated in Figure 12. The result for the CR 20 Model at 1.5 m/s is provided in Figure 13 as example of the results obtained for Criterion 2. Models CR 20 and CR 30 fall outside both the HRW 1 and HRW 2 criteria, and the CR 15 model falls outside the HRW 1 criteria. Therefore CR 15, CR 20 and CR 30 do not pass acceptance Criterion 2. The remainder of the eccentric reducers (CR 2.5, CR 5 and CR 10) are within the acceptance envelope for Criterion 2. This highlights that the positions of the Criterion 2 circles have an effect on the outcomes of the results.

### Results – Criterion 3 x-axis

Criterion 3 x-axis requires the velocity variation along line EGHF (x-axis) to be less than 10% of the cross-sectional average velocity. The results for Probe Positions 2, 3 and 4 are not provided, due to the uniformity in the velocity distribution at these positions, as illustrated in Figure 9. The concentric reducers were not assessed according to Criterion 2, as their velocity distributions are symmetrical around the centre of the pipe, i.e. the velocity contours are circular in shape and will produce results that will fall within the acceptance envelope.

The velocities along the circumference of the two circles are plotted on the y-axis and an angle relating to the position of the cylinder is plotted on the x-axis. This angle is a component of the coordinate system illustrated in Figure 12. The result for the CR 15 Model at 1.5 m/s is provided in Figure 13 as example of the results obtained for Criterion 2. Models CR 20 and CR 30 fall outside both the HRW 1 and HRW 2 criteria, and the CR 15 model falls outside the HRW 1 criteria. Therefore CR 15, CR 20 and CR 30 do not pass acceptance Criterion 2. The remainder of the eccentric reducers (CR 2.5, CR 5 and CR 10) are within the acceptance envelope for Criterion 2. This highlights that the positions of the Criterion 2 circles have an effect on the outcomes of the results.

### Results – Criterion 3 x-axis

Criterion 3 x-axis requires the velocity variation along line EGHF (x-axis) to be less than 10% of the cross-sectional average velocity. The results for Probe Positions 2, 3 and 4 are not provided, due to the uniformity in the velocity distribution at these positions, as illustrated in Figure 9. The velocities along the circumference of the two circles are plotted on the y-axis and an angle relating to the position of the cylinder is plotted on the x-axis. This angle is a component of the coordinate system illustrated in Figure 12. The result for the CR 15 Model at 1.5 m/s is provided in Figure 13 as example of the results obtained for Criterion 2. Models CR 20 and CR 30 fall outside both the HRW 1 and HRW 2 criteria, and the CR 15 model falls outside the HRW 1 criteria. Therefore CR 15, CR 20 and CR 30 do not pass acceptance Criterion 2. The remainder of the eccentric reducers (CR 2.5, CR 5 and CR 10) are within the acceptance envelope for Criterion 2. This highlights that the positions of the Criterion 2 circles have an effect on the outcomes of the results.

### Results – Criterion 3 x-axis

Criterion 3 x-axis requires the velocity variation along line EGHF (x-axis) to be less than 10% of the cross-sectional average velocity. The results for Probe Positions 2, 3 and 4 are not provided, due to the uniformity in the velocity distribution at these positions, as illustrated in Figure 9. The results for all four velocities presented similar results. The CR 20 model falls outside of acceptance Criterion 3 x-axis for all four flow velocities, and the CR 30 model falls outside of acceptance Criterion 3 x-axis for the 1.5 m/s, 2 m/s and 2.4 m/s velocities, and falls within the criteria for the
1 m/s velocity. The remainder of the models (CR 2, CR 2.5, CR 5, CR 10, CR 15, ER 2.5, ER 5, ER 10, ER 15 and ER 20) all fall within the Criterion 3 x-axis acceptance envelopes.

The results of the velocity distributions along the x-axis (line EGHF) used in Criterion 3 x-axis for the 1.5 m/s velocity are provided as an example of the results obtained for Criterion 3 x-axis in Figure 14.

### Comparison – Hydraulic transportability of air

The critical velocity for each of the concentric reducers was calculated with the upstream (larger) diameter (D1) as the input diameter in the calculation, as the velocity increases through the reducer, thereby increasing the capacity to hydraulically transport the air. The calculated critical velocities are provided in Table 6. The values calculated with Equation 1 and Equation 3 are representative of one another. However, the values calculated with Equation 2 are not representative (this could be due to the addition of a constant of 0.825 x (gD)$^{0.5}$ that affects the equation at low velocities). Only the values obtained with Equation 1 and Equation 3 were utilised to assess the hydraulic transportation of air through the models. The only model where air could not be hydraulically transported through it is the CR 20 model at 1 m/s. This model, however, also fails Criterion 1 and Criterion 3. It is noted that the critical velocity is a function of the pipe diameter (Equations 1, 2 and 3), and other results will differ from the results presented for the DN 200 upstream diameter utilised in this study. The hydraulic transportability of air will need to be assessed during the design of the pump station pipework.

### Results matrix

The final assessment of the 48 models simulated with CFD according to the four criteria, is provided in Table 7. Concentric reducers CR 2, CR 2.5, CR 5, CR 10 and CR 15, and eccentric reducers ER 2, ER 5 and ER 10 all produce velocity distributions that fall within all four of the acceptance criteria at all four flow rates modelled. The CR 20 concentric reducer and eccentric reducers ER 15, ER 20 and ER 30 all fail one or more of the assessment criteria.

The range of angles modelled is representative of a wide range of reducers, separated by large increments in order to keep the sample group manageable (i.e. ER 2.5, ER 2, ER 10 and not ER 1, ER 2, ER 3, etc). The reducers that have been listed for the maximum allowable slopes (CR 15 and ER 10) may be conservative. For example, a concentric reducer of

---

<table>
<thead>
<tr>
<th>Model</th>
<th>Slope (°)</th>
<th>Diameter (m)</th>
<th>Critical velocity (Eq 1)</th>
<th>Critical velocity (Eq 2)</th>
<th>Critical velocity (Eq 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR 2</td>
<td>2.0</td>
<td>0.2191</td>
<td>0.293</td>
<td>1.278</td>
<td>0.422</td>
</tr>
<tr>
<td>CR 2.5</td>
<td>2.5</td>
<td>0.2100</td>
<td>0.328</td>
<td>1.259</td>
<td>0.451</td>
</tr>
<tr>
<td>CR 5</td>
<td>5.0</td>
<td>0.2100</td>
<td>0.463</td>
<td>1.290</td>
<td>0.596</td>
</tr>
<tr>
<td>CR 10</td>
<td>10.0</td>
<td>0.2100</td>
<td>0.654</td>
<td>1.334</td>
<td>0.787</td>
</tr>
<tr>
<td>CR 15</td>
<td>15.0</td>
<td>0.2100</td>
<td>0.798</td>
<td>1.367</td>
<td>0.925</td>
</tr>
<tr>
<td>CR 20</td>
<td>20.0</td>
<td>0.2100</td>
<td>0.918</td>
<td>1.394</td>
<td>1.038</td>
</tr>
</tbody>
</table>

---

**Figure 14** Criterion 3 x-axis 1.5 m/s velocity distribution

---

**Figure 15** Concentric reducers standard sizes and acceptable slope of reducer
**CONCLUSIONS**

Figures 15 and 16 present charts illustrating the standard reducer sizes (from AWWA C208 and from the fitting manufacturers) and maximum slope of the acceptable size reducer, as defined in Table 7. From these charts it can be observed that various standard size reducer fittings fall outside of the acceptance criteria for the eccentric reducers. For the eccentric reducers this includes AWWA C208 standard eccentric reducer (angle of 14°). The addition of a straight length pipe may be used on the downstream end of the selected reducer to produce a more uniform velocity distribution at the pump inlet. This addition, however, decreases the NPSH available and increases the pump station’s footprint and cost.

**Table 7 Final results matrix**

<table>
<thead>
<tr>
<th>Acceptance</th>
<th>CR 2</th>
<th>CR 2.5</th>
<th>CR 5</th>
<th>CR 10</th>
<th>CR 15</th>
<th>CR 20</th>
<th>ER 2.5</th>
<th>ER 5</th>
<th>ER 10</th>
<th>ER 15</th>
<th>ER 20</th>
<th>ER 30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Criterion 2 HRW 1</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Criterion 2 HRW 2</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Criterion 3 X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Criterion 3 Y</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Air transport</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

**Figure 16 Eccentric reducers standard sizes and acceptable slope of reducer**

16° may still provide acceptable results, as the slope is less than the CR 20 reducer that did not pass the various criteria.
RECOMMENDATIONS

Review of pump inlet criteria
Based on the results obtained from this research it is recommended that concentric reducers be considered to improve the inlet conditions at pumps and that the following criteria should be considered during the assessment of pump suction pipe work:

- Criterion 1: Velocity variation along line ABCD is less than ±10% of the average velocity along line ABCD.
- Criterion 2: Velocity variation along line EGHF is less than ±10% of the average velocity along line EGHF.
- Criterion 3: Maximum velocity variation along a circle AEDF is ±5% of the average velocity along circle AEDF; the diameter of circle AEDF is 0.8 x ID.
- Criterion 4: Maximum velocity variation along a circle BGCH is ±5% of the average velocity along circle BHCG; the diameter of circle BGCH is 0.6 x ID.
- Criterion 5: Time-averaged velocity at any point is to be within ±10% of the cross-sectional area average velocity.
- Criterion 6: Air must be able to be hydraulically transported through the reducer.

Intended further research
It is recommended that:

- Additional examples with various bends, elbows and sweep tees typically found in pump inlet pipework that could possibly create swirl or non-uniform flow conditions entering into the suction reducers, are to be modelled. With this information, guidelines on typical suction manifold and suction pipework can be developed.
- Additional CFD modelling and/or physical model studies should be performed, where the hydraulic transport of air through the concentric reducer is explored, to assess which one of the hydraulic transport of air equations best suits this problem.

REFERENCES


Dynamic behaviour of underspanned suspension road bridges under traffic loads

J Oliva, J M Goicolea, P Antolin, M Astiz

Underspanned suspension bridges are structures with important economical and aesthetic advantages, due to their high structural efficiency. However, road bridges of this typology are still uncommon because of limited knowledge about this structural system. In particular, there remains some uncertainty over the dynamic behaviour of these bridges, due to their extreme lightness. The vibrations produced by vehicles crossing the viaduct are one of the main concerns.

In this work, traffic-induced dynamic effects on this kind of viaduct are addressed by means of vehicle–bridge dynamic interaction models. A finite element method is used for the structure, and multibody dynamic models for the vehicles, while interaction is represented by means of the penalty method. Road roughness is included in this model in such a way that the fact that profiles under left and right tyres are different, but not independent, is taken into account. In addition, free software (PRPgenerator) to generate these profiles is presented in this paper.

The structural dynamic sensitivity of underspanned suspension bridges was found to be considerable, as well as the dynamic amplification factors and deck accelerations. It was also found that vehicle speed has a relevant influence on the results. In addition, the impact of bridge deformation on vehicle vibration was addressed, and the effect on the comfort of vehicle users was shown to be negligible.

INTRODUCTION

Structural system

In underspanned suspension bridges the tensioned cable system is located below the deck, and tendons are deflected by struts or braces that are connected with the girder and thus transmit the upward cable deviation forces that help to support it. Cables are anchored in the deck over abutments and piers, if piers exist. This system is more efficient in single-span or simply-supported structures; in continuous bridges the under-deck cable system contribution in the negative bending regions (over piers) is very small, and tendons should be placed above the girder in those parts (Ruiz-Terán & Aparicio 2007b). This paper deals only with simply-supported bridges.

Cable forces are transferred into the deck as anchorage reactions that introduce axial compression stresses, and only vertical reactions appear at bearings. Cables are prestressed to neutralise permanent loads. Therefore, tendons are tensioned, the deck is compressed and bending is reduced due to the upward deviation forces in the deck.

This typology has significant advantages. In the first place, bridge cost is reduced due to the high structural efficiency. This efficiency in the use of materials arises from the important contribution of the axial response compared with bending. Furthermore, because of this bending response reduction, it is possible to build more slender decks, so the advantages are not only economical, but also aesthetic. It is important to remark that, although deck height is reduced, the overall structural depth is increased due to the location of the cable system. Hence, substantial vertical clearance is required below the road for this type of bridge to be built.

Simply-supported underspanned suspension viaducts can span medium distances of around 80 m that would otherwise be bridged with a multi-span solution (for example a three-span bridge with lengths 24+32+24 m and a deck height of 1.5 m). Hence this solution avoids the erection of piers, with the subsequent reduction of construction time and other complications. In the case of a conventional solution with one single span, the resulting structure would be much heavier. Depending on the total weight of the bridge, an underspanned suspension structure could even be lifted into place by a crane, which would significantly reduce the disturbances produced during construction.

On the other hand, their inherent lightness makes these bridges more sensitive to certain dynamic actions.

Keywords: underspanned suspension bridges, road roughness, vehicle–bridge dynamic interaction, deck vibrations, penalty method
Background

First examples of bridges with the tensioned system located below the girder are found in the nineteenth century. In the 1830s G H Dufour built the Ile aux Barques Bridge in Geneva, Switzerland, with a main span of 33.5 m (Peters 1987), and J Smith erected the Micklewood Bridge (Figure 1), a 30 m single-span structure, in Scotland (Drewry 1832). In both structures the girder rests on chains that are anchored into the deck. Chains and deck are connected to each other by vertical struts that are in compression, in contrast with the vertical tensioned elements of traditional suspension bridges.

In the last quarter of the twentieth century this kind of structure started to reappear. Some authors ascribe this reappearance as having been initiated by Fritz Leonhardt’s design for the Neckar Valley Bridge in the late 1970s. Leonhardt used this solution in order to avoid pier foundations in the hillside due to soil-related problems. End piers were removed in the design stage, and first and last spans are supported by cable systems located below the girder. These cables introduce vertical reactions in the deck by means of vertical elements and thus replace the eliminated piers; Figure 2 shows one of these systems. Ruiz-Terán and Aparicio (2007b) presented a wide and exhaustive state-of-the-art study of these bridges, considering the Neckar Valley viaduct as the birth of the typology (the term under-deck cable-stayed bridges is used by these authors), and therefore nineteenth century structures are not included in their work. Other researchers conceive these bridges as an evolution from externally prestressed bridges in which bending behaviour is enhanced by locating tendons outside cross-section bounds. These authors call these structures bridges with highly eccentric external tendons (Mutsuyoshi et al 2010).

Examples of simply-supported underspanned suspension bridges can be found all over the world, for instance the Tobu and Inachus footbridges in Japan, or the Truc de la Fare road bridge in France which spans 53 m. A very exhaustive review can be found in Ruiz-Terán and Aparicio (2007b), and several examples of footbridges are explained in Strasky (2005).

In spite of the important advantages pointed out above, underspanned suspension bridges are still unusual structures, and authorities remain reluctant to build them. This is due to limited knowledge about these viaducts. One of the first concerns is the shear capacity, as the girder depth is substantially reduced, but it has been proved experimentally that the shear resistance is even higher than in conventional girders (Mutsuyoshi et al 2010). The structural behaviour of this kind of bridge was studied both analytically and experimentally in Witchukreangkrai et al (2000), Aravinthan et al (2001), and Ruiz-Terán and Aparicio (2007a). Some design criteria were proposed by Ruiz-Terán and Aparicio (2008) by using simply-supported 80 m span bridges.

Nowadays one of the main concerns is the lack of knowledge regarding the dynamic response of these bridges to different excitations. Ruiz-Terán and Aparicio (2009) dealt with the sudden breakage of stay cables due to a truck impacting with them, which was the principal reason why this solution for the
In this work, the dynamic response of under-spanned suspension bridges under the action of running vehicles is analysed by means of a vehicle–bridge interaction (VBI) model. Road roughness is considered in the simulations, and differences between excitations under left and right tyres are taken into account by means of an approach presented by the authors in Oliva et al. (2013a; 2013b).

The main objective of this work is to shed some light on the dynamic behaviour of these viaducts under traffic loads, as this knowledge is still limited. Important traffic-induced dynamic excitation will happen due to the extreme lightness and flexibility of this structural solution. In fact, accelerations in these decks are higher than in conventional bridges.

In this study one bridge of medium length is employed as a representative example, and only one set of support conditions is considered. Double bearings are set at both abutments, and therefore torsional rotation is prevented at those points. The use of sliding supports will not have a significant influence on the results, as vertical and torsional modes will not be affected. Another solution could be the use of only one support at one abutment. Structural torsional stiffness would decrease significantly and hence torsional frequencies of vibration would also be reduced. However, the low torsional rigidity inherent in this structural system makes this option inadvisable, and hence this solution is not considered in this work.

**VEHICLE-BRIDGE INTERACTION MODELS**

**Vehicle**

A two-axle truck is considered in this work. This vehicle model has also been used by other authors, e.g., Law and Li (2010), and is based on the H20-44 truck design loadings included in the AASHTO (1998) specifications. In this study a seat is added to the truck model in order to assess the dynamic effects on the driver.

The complete model consists of three rigid bodies that represent the box and both axles, plus one DOF mass that reproduces the driver seat. The vehicle model and its eight DOFs are depicted in Figure 3. The driver seat mass \( m_d = 80 \text{ kg} \) is connected to the vehicle body by a vertical spring \( k_d = 10507 \text{ N/m} \) and dashpot \( c_d = 876 \text{ N·s/m} \). Driver seat properties are taken from Zuo and Nayfeh (2007). Truck model mechanical properties can be found in Zhu and Law (2002).

**Structure**

An 80 m long simply-supported under-spanned suspension viaduct is considered in this work (Figure 4), as described in Ruiz-Terán and Aparicio (2009). The cable system is anchored in the deck over the two abutments and is deflected by two slightly inclined steel struts. The deck is a voided concrete girder with a depth of 1 m and a total width of 13.2 m (Figure 5). Double supports are set at both abutments, hence torsional rotation is not permitted at those points.

This bridge is modelled by means of the finite element method – shells are employed to represent the deck and beams for the struts, and both kinds of elements consider shear deformation and adopt reduced integration. Cables are represented by means of truss elements. Forces and displacements induced in the bridge by the crossing vehicles are small enough to adopt linear elastic behaviour in the structure. Nonstructural masses representing other dead loads, such as pavement and safety walls, are included in the model. The damping matrix is built following the Rayleigh method by setting a damping ratio of 1% for the first mode \( (0.78 \text{ Hz}) \) and also for 50.0 Hz. Figure 6 shows the first bending mode of vibration \( (1^{st} \text{ global mode}) \) and the first torsion mode of vibration \( (3^{rd} \text{ global mode}) \).
Vehicle–bridge dynamic interaction

Vehicle–bridge vertical interaction is gathered through a node-to-surface contact at each vehicle tyre, and those contacts are implemented by means of the penalty method. This method does not introduce additional variables to the problem, and the computation is therefore faster. On the other hand, the constraint equation is only fulfilled approximately and some penetration will be unavoidable. This penetration will depend on the penalty parameter $\varepsilon$. When there is no contact, no forces are added and separation is reproduced in the numerical model. The whole system is set out as a fully coupled system of equations and it is solved by direct time integration with the HHT-\(\alpha\) method (Hilber et al 1977). This methodology has two main advantages over other fully coupled methods (Deng & Cai 2010; Yin et al 2010; Neves et al 2012): (1) modal superposition is not used for the bridge subsystem and hence structural nonlinearities could be considered, and (2) vehicle tyres can lose contact with the deck surface. This separation capability is of interest in certain situations, as will be shown below.

As stated before, the constraint equation is fulfilled approximately when the penalty method is employed. Hence the solution is only an approximation of the correct enforcement of the constraint condition obtained with the Lagrange multiplier method (Wriggers 2002). In Figure 7(a) penetration at the right rear wheel when the vehicle crosses the bridge at 110 km/h is shown when Penalty and Lagrange multiplier methods are employed; road roughness is included. As can be seen, some penetration takes place with the penalty method. However, the contact force is the same in both cases. Figure 7(b) shows the vertical reaction under the same wheel during a short period of time, in order to facilitate visualisation. It can be concluded that the penalty method leads to correct results, although some penetration is inevitable.

Figure 7 Penalty method versus Lagrange multiplier method: (a) penetration at right rear wheel, (b) contact force at right rear wheel

ROAD SURFACE DESCRIPTION

Road roughness is generally the most important source of dynamic excitation in road traffic, hence a correct definition of the road surface is a key point in vehicle–bridge interaction problems. When the actual profile of a particular road stretch is not needed, but a set of profiles that are representative of a certain sort of road, stochastic definitions for the generation of synthetic profiles have to be used, as for example in Deng and Cai (2010). The fact that profiles under left and right tyres are different, but not independent, is seldom considered in this kind of simulation; therefore it is assumed that the road profile is constant across the deck width. In this paper those differences are considered by means of a procedure developed by the authors and described in Oliva et al (2013a). The influence of left–right dissimilarity on road vehicle–bridge interaction dynamics has been shown in Oliva et al (2013b).

In order to facilitate the generation and use of parallel profiles for other researchers, the authors have implemented a free-to-download program (http://w3.mecanica.upm.es/prpgenerator/index.php) that creates pairs of profiles with the mentioned procedure. This simple application named PRPgenerator was developed in MATLAB^{\textregistered} and it is introduced for the first time in this paper. A brief description can be found in Appendix A.

The procedure employed in this work for considering that fact assumes hypotheses of road surface isotropy and homogeneity. In a homogeneous and isotropic road surface every straight profile has the same statistical characteristics, independent of its direction or position. Thus, parallel profiles along the road share statistical properties, but are not the same. Given that, the cross-Power Spectral Density ($G_x$) of the pair of parallel profiles at a certain distance is obtained from the direct Power Spectral Density or PSD ($G$). The methodology is not explained here, for the sake of briefness, but it is detailed in Oliva et al (2013a; 2013b). The ISO-8608 (1995) PSD definition must be slightly modified in order to fulfil the

Figure 8 Parallel road profiles at 2.05 m
isotropy admissibility conditions (Kamash & Robson 1977). It must be set constant and equal to \( G(n_a) \) at all wave numbers below a limiting wave number \( n_a \). Thus it becomes acceptable, as its integral is bounded and the function remains monotonically non-increasing. This limit is set as \( n_a = 0.01 \) m\(^{-1}\), which is the minimum spatial frequency considered in roads (ISO-8608 1995). The PSD becomes:

\[
G(n) = \begin{cases} 
G(n_a) \left( \frac{n_a}{n} \right)^{-2} & \text{for } n \leq n_a \\
G(n_a) \left( \frac{n}{n_0} \right)^{-2} & \text{for } n > n_a
\end{cases}
\]  

where \( G(n) \) is the one-sided power spectral density for the spatial frequency or wave number \( n \) and \( G(n_a) \) is the one-sided power spectral density for the reference spatial frequency \( n_a = 0.1 \) m\(^{-1}\). The value for \( G(n_a) \) is prescribed by ISO-8608 (1995) as a function of the road class.

Two parallel road profiles are generated as follows (Sayers 1998):

\[
y_1(x) = \sum_{i=1}^{N} \sqrt{2}G(n_i)\Delta n \cos(2\pi n_i x + \phi_i) \quad (2)
\]

\[
y_2(x) = \sum_{i=1}^{N} \sqrt{2}G(n_i)\Delta n \cos(2\pi n_i x + \phi_i) + \sqrt{2}(G(n_i) - G(n_a))\Delta n \cos(2\pi n_i x + \theta_i) \quad (3)
\]

where \( \phi_i \) and \( \theta_i \) are two sets of random phase angles which are uniformly distributed from 0 to \( 2\pi \). An example of two class A parallel profiles is depicted in Figure 8.

**NUMERICAL STUDIES**

For the truck and bridge presented in this section numerical simulations are used. Geometric nonlinearities can easily be considered in the proposed framework, but nonlinear effects in the considered scenarios have been shown to be negligible. Therefore, linear models are employed in this work because of the shorter computation time required. Six different running speeds were employed (30, 50, 60, 80, 110 and 120 km/h) in order to consider typical urban and highway vehicle speeds in different countries around the world. The road surface is assumed to be very good (class A, \( G(n_a) = 16 \times 10^{-6} \) m\(^3\)) (ISO-8608 1995), 2000 spatial frequencies between 0.01 m\(^{-1}\) and 10.0 m\(^{-1}\) are used in the profile generation, and road irregularities are sampled every 2 cm in the longitudinal direction. Ten road surfaces (A01, A02, ..., A10) are generated so that the results are statistically significant. Thus, 60 different simulations had to be performed for this part of the work (6 speeds x 10 profiles). The H20-44 vehicle crosses the structure in the right lane with its right wheels at a distance of 2.55 m from the right edge of the bridge (see Figure 9). Driving on the right-hand side of the road is assumed in this work as it is the most common option in the world. The outcome of the study with left-hand traffic would be the same.

Results are obtained at every node on the bridge surface. These nodes are gathered in longitudinal lines whose transversal position and name are shown in Figure 9. The bridge surface is discretised with 14 elements in width and 81 elements in length, so that there are 15 longitudinal lines with 82 nodes per line.

With regard to the Serviceability Limit State of vibration in road bridges, Eurocode EN1990 (EN1990:2002/A1+AC 2010) is very vague, and no indication is given with respect to the comfort of passengers or pedestrians. With respect to footbridges it is stated that pedestrian comfort criteria for serviceability should be defined in terms of maximum acceptable acceleration of any part of the deck. The recommended maximum value for vertical acceleration is set as 0.7 m/s\(^2\). This limit is adopted in this work as a limitation to pedestrian comfort on the sidewalks of road bridges.

Maximum acceleration at every bridge surface node is obtained with the ten different road surfaces (A01, ..., A10). The mean value of those ten maxima is computed in order to get representative accelerations of the road class, instead of an absolute maximum that would represent a critical situation. Figure 10 shows the mean value of maximum vertical acceleration on the whole bridge surface for vehicle speeds of 110 km/h and 30 km/h. In both cases high accelerations are found right under the vehicle path due to local vibrations. At high speeds these accelerations are significantly higher than those of the rest of the surface. At low speeds the local apexes are not so markedly protruding. Clear peaks can be noticed in the two cases near the first abutment, located precisely under the truck (between lines A7 and V4); the maximum value is in line V4, because it belongs to the cross-sectional cantilever. These first abutment high values are caused by the sudden entrance of the vehicle.
of the node located in line V4 at 1.0 m from the abutment, with \( V = 50 \text{ km/h} \), is depicted in Figure 11 by the label \textit{Raw}; high accelerations appear when each truck axle enters the bridge.

In Figure 12 maximum acceleration of some significant lines (V1, V4, V6 and A5) is depicted for 120 and 60 km/h. A peak at the beginning of V4 is easily noticeable in both cases. For 120 km/h accelerations in nodes belonging to V4 are clearly the highest in the bridge, for 60 km/h local vibrations are less important and V4 accelerations are similar to those found in lines V1 and V6. Acceleration in the bridge longitudinal midline (A5) is always significantly lower than in the cantilevers, which fact indicates the relevance of torsion effects.

The Eurocode EN1990 limitation is generally fulfilled at 60 km/h, and is only violated near the first abutment because of the very local effects of the vehicle entrance, as explained before. At high speed (120 km/h) the maximum acceleration criterion is not satisfied in the vehicle side cantilever (V4–V6) all along the bridge length. It is remarkable that vertical acceleration along the bridge length is uniform for every vehicle speed, with the exception of the vehicle entrance region.

Two different standpoints exist regarding the acceleration effects on human comfort. The first considers that people are affected most by the largest peaks, and therefore maximum acceleration is limited as in Eurocode EN1990. According to the second school of thought it is assumed that the degree to which vibration may be noticed or tolerated is determined by some averaged effect over a period of time. International standard ISO-2631-1 (1997) specifies a method of evaluation of the effect of vibration on human beings by means of the weighted root-mean-square (RMS) acceleration which is defined as:

\[
aw = \sqrt[2]{\frac{1}{T} \int_0^T a^2_w(t) dt}
\]  

(4)

where \( a_w(t) \) is the weighted acceleration as a function of time and \( T \) is the duration of the vibration.

Human response to vibration is a function of frequency and therefore data must be weighted in order to give greater prominence to frequencies where humans are most sensitive. Analogue transfer functions that determine the frequency weighting are specified in ISO-2631-1 (1997). Those functions are transformed into digital filters by means of bilinear transformations.

For a standing person subjected to vertical accelerations underneath their feet, as is the case with a pedestrian on a bridge, \( W_k \) weighting is applicable. Its parameters are given in ISO-2631-1 (1997). As an example, Figure 1 shows vertical weighted acceleration \( a_w(t) \) in line V4 at 1.0 m from the first abutment; high-frequency peaks are eliminated by the frequency weighting. Reaction to vibration depends on many factors (frequency, duration of vibration, activity, age, etc), and therefore absolute limits cannot be established, only indications of likely human reactions. Some limits are given by the Department of Environment and Conservation of New South Wales in Australia (DEC 2006), but no specific values are provided for bridges. The highest admissible value in buildings is set as 0.08 m/s². A weighted RMS acceleration of 0.08 m/s² is also adopted in British Standard BS 6472:1992 (1992) as a value below which there exists a low probability of adverse comment from users of residential buildings.
during the day, with a time of exposure to vibration lower than 225 seconds (the bridge in this study is crossed by a person approximately one minute). Building limits are not directly applicable to bridges. The same value (0.08 m/s²) is defined in the German guideline VDI 2057 (2002) as a limit above which vibration is strongly perceptible by a human being. Higher values would probably be tolerated by pedestrians on a road bridge, but it seems reasonable to adopt 0.08 m/s² as a guidance value for human comfort on viaduct decks, although presumably conservative.

As for maximum acceleration, weighted RMS acceleration is computed for every bridge surface node with the ten different road surfaces considered, and the mean value is calculated. Figure 13 shows the mean value of the weighted RMS acceleration all over the bridge surface. First the abutment peaks and high values on the vehicle trajectory disappear, and the highest values are now obtained in the bridge edges (V1 and V6), both with high and low speeds. In Figure 14, values in some significant lines (V1, V4, V6 and A5) are depicted for 120 and 60 km/h.

Weighted RMS remains at acceptable levels on the whole deck surface for all vehicle speeds. It is remarkable that differences between the left and right bridge edges, which were evident when using maximum acceleration values, almost vanish if weighted RMS is employed.

With respect to the Dynamic Amplification Factor (DAF) of vertical displacement, the mean value of the ten cases is depicted in Figure 15 for the bridge midline (A5) considering different vehicle speeds. High DAF values appear and they are highly influenced by vehicle speed. DAF values as high as 1.54 are reached with V = 120 km/h near the first strut. These high values indicate the huge relevance of dynamic effects on this kind of structure.

Regarding vehicle vibration, Figure 16 shows vertical acceleration at the driver seat when the vehicle runs at 120 km/h with
profile A07. The vertical response of the vehicle on the bridge and on a rigid road is compared under the same road surface conditions in order to assess the bridge flexibility influence. The results in Figure 16 show that bridge deflection has an effect on the driver seat. 

For vertical vibrations on the seat surface, $W_v$ weighting is also used (ISO-2631-1 1997). Figure 17 shows the mean value of weighted RMS acceleration on the driver seat considering the ten surfaces – results on the bridge and on a rigid road are compared. The increment is very small (maximum of 4%), and it can therefore be concluded that the bridge flexibility influence on driver comfort is of no significant relevance.

**Effects of pothole presence**

The eventual presence of a pothole of considerable size is considered by adding a 50 cm long and 5 cm deep defect on road surface A01, located in the midspan and affecting both vehicle sides. Tyre contact is lost in the pothole and high reactions appear when it is regained (Figure 18). The wheel–bridge separation capabilities of the interaction model are necessary for a correct simulation of this scenario. This behaviour produces very high accelerations in the bridge surface (Figure 19); even at low speeds, higher values are beyond 6.5 m/s² when $V = 50$ km/h.

**Road quality influence**

Profiles of different road classes (A or Very Good, B or Good, and C or Medium) are employed in order to assess road class influence in the dynamic behaviour of the deck. Profiles of different road classes are related by a scale factor:

$$y_B(x) = y_A(x) \frac{\sqrt{G(n_0)}_B}{\sqrt{G(n_0)}_A}$$

(5)

Thus B and C surfaces can be obtained by multiplying A pairs of profiles by 2 and 4 respectively. Profile A10 is employed in this section and the vehicle speed is set as 110 km/h. Maximum vertical acceleration at lines V1 and A5 is depicted in Figure 20. As can be seen, road quality has significant influence on the deck vibration, and the increase in maximum vertical acceleration is of relevance. The Eurocode limit is not reached in the midline (A5), but is exceeded in line V1 with road level B.

Road class also has influence in the weighted RMS acceleration (Figure 21) – values increase when road quality declines, and exceed 0.08 m/s² manifestly in line V1 with road classes B and C. In the midline (A5) weighted RMS acceleration also increases, but values remain acceptable.
CONCLUSIONS

Vehicle-induced dynamics in underspanned suspension bridges were studied in this paper. Analyses were performed by means of a fully coupled vehicle–bridge dynamic interaction model. The vehicle was represented through a multibody system (MBS); the bridge was modelled with the finite element method (FEM), and interaction was gathered by means of contacts with a penalty formulation. The fully coupled system equations were solved by direct integration in time.

Penetration inherent in the penalty method was shown to have no effect on the relevant results against the Lagrange multiplier method, where restriction is perfectly satisfied and no penetration takes place.

Road roughness was considered in such a way that the fact that profiles in the left and right tyres were different, but not independent, were taken into account. In order to facilitate their consideration, a parallel road profiles generation programme (PRPgenerator) was developed by the authors, and this is available on the web for free download as a standalone application.

A very good road class (A) was considered in the study (ISO-868 1995). Vertical acceleration peaks appeared directly under the vehicle path due to local vibrations. Absolute maximum values arose near the first abutment because of sudden vehicle entrance. The Eurocode EN1990:2002/A1+AC (2010) maximum acceleration criterion was generally fulfilled on the deck surface during truck runs at urban speeds (50–60 km/h). With highway speeds (110–120 km/h) this limitation was not satisfied in the vehicle proximity, but on the rest of the deck it was. Frequency weighting
(ISO-2631-1 1997) reduced the high frequency peaks on acceleration histories, and the largest weighted RMS values were not found in the vehicle path but on both bridge edges. In addition, differences between left and right viaduct edges were negligible. Weighted RMS values were low enough to be tolerated on the whole deck surface for every vehicle speed.

Significant differences were found between the bridge longitudinal midline and edges. Accelerations were clearly higher on the edges due to the excitation of the deck torsion. Therefore it would be convenient to locate sidewalks along the deck longitudinal midline, if possible, or set vehicle lanes near that line in order to reduce the excitation of this rotation.

Road class has an evident influence on pedestrian comfort. Hence, good construction and maintenance are of great importance in this kind of viaduct due to its dynamic sensitivity. The presence of a significant pothole could induce very high deck accelerations. Therefore, careful conservation is important. With respect to vehicle behaviour, bridge flexibility causes additional vibration in the driver seat. However, the effect of the additional acceleration on driver comfort is very small – the maximum increase of the weighted RMS acceleration on the driver position was only 4%.

ACKNOWLEDGEMENTS

This work was motivated by lectures given by Dr Ana María Ruiz-Terán (Imperial College London) at the Technical University of Madrid. The authors wish to acknowledge this source of inspiration. The aid of the Ministerio de Ciencia e Innovación of the Spanish government through sub-programme INNPACTO and project VIADINTEGRA (Ref IPT-370000-2010-012) is greatly appreciated. The authors are grateful also for support provided by the Technical University of Madrid, Spain.

APPENDIX A (PRPgenerator)

PRPgenerator is a simple-to-use and free application that generates pairs of parallel road profiles. It has been implemented in Matlab® and is available as stand-alone software. The program uses Power Spectral Density definition from ISO-8608 (1995) and assumes the homogeneity and isotropy of the road surface. PRPgenerator has been developed by the Computational Mechanics Group of the School of Civil Engineering at the Technical University of Madrid (UPM). Its theoretical background can be found in Oliva et al (2013a; 2013b), and the application is presented for the first time in this paper. It computes cross-Power Spectral Density and coherence function for parallel profiles and generates profiles with the total length, sample distance and number of spatial frequencies specified by the user. Different road classes can be considered and any vehicle width may be used. Figure 22 shows the main and only window of the application.

REFERENCES


Guidelines for the preparation of papers and technical notes

Authors should comply with the following guidelines when preparing papers for publication in the journal.

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 X2080, 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 verelene@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 reviewing 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.
- Tables, figures and equations should be numbered consecutively and should appear in the text directly after they have been referred to for the first time.
- Figures should be produced using computer graphics. They should be submitted as separate files in JPEG (jpg) format and not embedded in the text document. Lettering on figures should be equivalent to a Times New Roman 9 pt font or slightly larger (up to 12 pt) if desired. Lettering smaller than 9 pt is not acceptable. Hand-drawn figures will not be accepted.
- Mathematical expressions: Equations should be presented in a clear form which can be easily read by non-mathematicians. Symbols should be typed using the Times New Roman symbol set. Equations should be in italics. Each equation should appear on a separate line and be numbered consecutively.
- Symbols should preferably be defined in the text, but if this is not feasible, a list of notations may be provided for inclusion at the end of the paper. Symbols should conform to accepted international usage. The SI system of units should be used. Decimal commas may be used, but decimal points are preferred.
- Since mathematical formulae are generally scanned in during the production process, they should be set out as they are to appear in the printed paper.
- Headings: Sections and paragraphs should not be numbered. The following hierarchy of headings should be followed:

HEADING OF MAIN SECTION
Heading of subsection
Heading of sub-subsection

References: References should follow the Harvard system. The format of text citations should be as follows: ‘Jones (1999) discovered that …’ or ‘recent results (Brown & Carter 1985; Green et al 1999) indicated that …’
- References cited in the text should be listed in alphabetical order at the end of the paper. References by the same author should be in chronological order.
- The following are examples of a journal article, a book and a conference paper:


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.