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

\[ k_x \frac{\partial^2 h}{\partial x^2} + k_z \frac{\partial^2 h}{\partial z^2} = 0 \]
2 Modelling of the fracture process zone to improve the crack propagation criterion in concrete
S Shahbazpanahi, A A A Ali, F N Aznieta, A Kamgar, N Farzadnia

10 Friction behaviour of fresh concrete in the vicinity of formwork
S Bouharoun

18 The integration of railway asset management information to ensure maintenance effectiveness
N J van der Westhuizen, P J Gräbe

30 Laser-based assessment of road aggregate particle shape and texture properties with the aim of deriving comparative models
J J Brydenbach, J K Anochie-Boateng, P Paige-Green, J L van Rooy

36 Incorporating rainfall uncertainty into catchment modelling
J G Ndiritu

47 Rocking shear wall foundations in regions of moderate seismicity
J E van der Merwe, J A Wium

57 A field study of the in situ moisture regime during active hydraulic tailings deposition
C J Madroburt, G E Blyght

69 The use of ground-penetrating radar to develop a track substructure characterisation model
D J Vorster, P J Gräbe

79 Causes of delay in road construction projects in Malawi
M J Kamanga, W J v d M Steyn

86 Critical normal traffic loading for flexure of skew bridges according to TMH7
A D Malan, G C van Rooyen

96 Structurally efficient housing incorporating natural forms
M Gohnert, A Fitchett, J Bulovic, N Bhikhoo

103 Exploring the flow number (FN) index as a means to characterise the HMA permanent deformation response under FN testing
L F Walubita, J Zhang, A E Alvarez, X Hu
INTRODUCTION

The mechanical behaviour of ductile materials is different from that of quasi-brittle materials. Cracks grow in ductile materials such as metals due to the intersection and coalescence of micro-voids, while in quasi-brittle materials such as concrete, cracks propagate when the aggregates interlock or when micro-crack bridging occurs (Yang & Liu 2008).

In fracture mechanics a crack is assumed to start when there is a notch or a stress concentration in the tension zone. Linear elastic fracture mechanics (LEFM) was first used to study crack propagation during World War II (Esfahani 2007). Later some studies used LEFM to analyse crack propagation in concrete. However, Kaplan (1961) found that LEFM could not be applied to crack problems in normal concrete sizes.

The first model based on nonlinear fracture mechanics in concrete was proposed by Hillerborg et al (1976). It was shown that there is a region called the fracture process zone (FPZ) in front of the real crack tip, which is responsible for crack closure (see Figure 1). This significant and relatively large zone contains micro-cracks in matrix-aggregate, gel pores, shrinkage cracks, bridging and branches of cracks that are located ahead of the macro-cracks. Since a significant amount of energy is stored in the FPZ, a crack can have stable growth before the peak load. In addition, the existence of the FPZ accounts for the strain softening behaviour in the stress-crack opening curve that is observed after peak load. In this region interlocking crack surfaces contribute to a gradual decline in stress and prevent sudden failure (Esfahani 2007). The dimension of the FPZ depends on the size of the structure and the length of the initial crack, as well as on the loading and material properties of the concrete. The length of the FPZ is of special interest as compared with its width. The effective modulus of elasticity is reduced when the crack moves from undamaged regions into the FPZ.

The so-called Griffith energy approach can be used to describe the crack propagation criterion in the fracture process at the crack tip. This approach states that the energy release rate, defined as the amount of the energy stored in the FPZ which is required to form the crack, must be sufficiently larger...
than the critical fracture energy. Hence to study the crack state, the crack propagation criterion can be defined in terms of the energy release rate.

To simulate the FPZ, Hillerborg et al. (1976) used cohesive stress, which is a function of crack opening. In the FPZ cohesive stress reaches its maximum at the tip of the crack, which is equal to the tensile strength \( f_c \) and in the critical opening of the crack \( w_c \) it declines to zero. The area under the stress-opening curve is defined as the energy release rate. The cohesive zone model (CZM) was applied by Hillerborg to model the FPZ in normal-size structures using either the nodal force release rate or the technique of interface element with zero initial thickness (Ingraffea et al. 1984). Since the length of the FPZ, \( l_{eq} \), is of special interest compared to its width, the interface element with zero thickness is preferred (Yang & Liu 2008).

To model the FPZ, Bazant and Oh (1983) used dummy bands in which micro-cracks were uniformly distributed. This model, called the crack band model (CBM), was simulated using a layer continuum element in the finite element method. As the CBM depends on the width of the element, it was suggested to model only Mode I fracture (Rots & Brost 1987).

The non-local continuum approach is another method that uses the width and the length of the interface element for modelling the FPZ. This approach, however, uses too many degrees of freedom, and for this very reason it is not computationally affordable (Ingraffea et al. 1984).

From a finite element point of view, to model the FPZ the stiffness of the element should be properly chosen. In practice, compared to undamaged zones, the FPZ has a different stiffness due to micro-cracking, bridging and branching processes, which provide the energy required for the crack growth. In the softening zone, although some resistance is observed, the stress drops dramatically. Parallel use of the constitutive model of the CBM and the cohesive model may lead to a better simulation of the FPZ behaviour in terms of elastic branching and the softening zone. In addition, in order to estimate the nodal force, an accurate constitutive model for normal and shear stresses is needed.

To estimate the crack propagation state in the finite element model, either of two methods are used: the strength-based or the energy-based approach. To achieve higher accuracy, the energy-based approach is often used to simulate the CZM. The energy approach criterion depends on the stiffness matrix, the displacement and the crack geometry (Wua et al. 2011). Therefore the crack propagation criterion should be modified based on the new stiffness matrix in the FPZ.

In addition, when the FPZ has fully propagated and reaches its maximum length, a stress-free length appears in front of the notch (or macro-crack) behind the FPZ (Wua et al. 2011). This stress-free length has not been considered in previous research (Bocca et al. 1991; Xie & Gersle 1995; Prasad & Krishnamoorthy 2002; Yang & Liu 2008; Shi 2009; Ooi & Yang 2011; Guo et al. 2012).

When modelling cracks, another issue to deal with is the direction of the crack. The initial direction of the propagation is mostly unknown. Numerous reports have proposed the use of approximate re-meshing algorithms as the crack starts to grow (Xie & Gersle 1995). In these algorithms, a significant number of nodes are created for re-meshing and a large stiffness matrix is created. In this approach, the computational complexity is relatively high. An alternative method is the so-called inter-element boundaries technique, which identifies the crack path (Alfaiate et al. 1997).

This paper proposes a new constitutive model for the stiffness matrix to model the FPZ. Based on the new stiffness element, Griffith differential energy is improved to predict the propagation criterion. This work uses the method presented by Alfaiate et al. (1997) instead of re-meshing. In this method, the crack propagation direction is identified by following the inter-element boundaries. The proposed approach is capable of modelling the mixed mode of the crack, which is described by the stress versus the crack opening displacement curve. Three examples are modelled to validate the criterion: a plain concrete beam with a notch shear crack, a notched reinforced concrete beam and a reinforced concrete beam with simple support are analysed and compared to recent experimental results and previous modelling.

## Numerical Model

### Stiffness of Interface Element

As mentioned earlier, the FPZ has a softening behaviour due to the interlock of aggregates and micro-cracks. Thus different stiffness characteristics are used in the finite element method to model the FPZ. A single four-node interface element with linear variation of the crack opening displacement (COD) is shown in Figure 2. To model the mixed mode, the relationship between the stress and the displacement is represented by matrix \( D \), which is given as:

\[
D = \begin{bmatrix}
D_{nn} & D_{nt} \\
D_{nt} & D_{tt}
\end{bmatrix}
\]

(1)

where \( D_{nn} \) and \( D_{tt} \) are the normal and tangential stiffness respectively, and \( D_{nt} \) represents the additional stiffness due to the interaction between the shear and the normal stiffness in the fracture mechanism.

To estimate the normal stiffness, the normal stress versus COD curve was used. Figure 3 illustrates the concrete COD curve due to normal stress.

The total opening can be separated into two components:

\[
dw = dw_e + dw_t
\]

(2)

where \( dw_e \) and \( dw_t \) are the opening elastic and the opening softening, respectively.

The softening parameter is defined as:

\[
S = \frac{da}{dw_t}
\]

(3)

This parameter is in fact the slope of stress in the softening portion of the curve and its value is negative. If \( E_s \) and \( E_t \) are the slope of the elastic zone and the slope of the softening zone at the \( i \)th iteration of the nonlinear solution respectively, we can write:

\[
S = \frac{da}{dw_t} = \frac{E_t}{1 + E_t E_s}
\]

(4)

Eq (4) implies that the softening parameter changes due to changes in \( E_t \). On the
other hand, the normal stiffness can be expressed as:

$$D_{nn} = \frac{L}{S}$$  \hspace{1cm} (5)

where the $L$ is the interface element length. From Eq (4) and (5) we obtain the following:

$$D_{nn} = \frac{1 + \frac{E_i'}{E_i}}{E_i} E_{tt}$$  \hspace{1cm} (6)

From the study done by Yoshikawa et al (1989), shear stiffness and the additional stiffness due to the interaction between the shear and normal stiffnesses are represented as a linear function of the normal stiffness as:

$$D_{tt} = \mu \beta D_{nn}$$  \hspace{1cm} (7)

$$D_{nt} = \beta D_{tt}$$  \hspace{1cm} (8)

where $\mu$ and $\beta$ are a frictional coefficient and dilatancy factor respectively. Although these two parameters are functions of the normal stiffness, in this paper, for the sake of simplicity, they are assumed to be constant. The stiffness of the interface element $K_i$ in the softening zone is given by:

$$K_i = B^T D B dA$$  \hspace{1cm} (9)

where $B$ is the strain-displacement matrix used to model the interface element (Desai et al 1984), $dA$ is the differential element of the crack surface area, and the $T$ is transpose. In material nonlinearity, the matrix $D_{ij}$ changes with different levels of loading, particularly after the peak load, where the stiffness matrix is of interest. The method of nonlinearity is an incremental-iterative technique. In this method, the stiffness matrix is updated at the beginning of each load increment. Small displacements are assumed to implement the nonlinear solution (Prasad & Krishnamoorthy 2002). Based on small displacements, at each step of the loading the tangential stiffness method is used at the $j$th iteration ($i = j + 1$) of the nonlinear solution by using the $j$th displacements of the interface element. To estimate the slope of the softening zone, an initial displacement, $w_{ij}$ and the elastic slope, $E_{ij}$, are used (Gerstle & Xie 1992). Corresponding to the initial displacement, the residual force is calculated linearly, and then the correction of the displacement to the trial value is calculated. The stiffness matrix is updated by accumulating the corrections of the displacements. This iterative procedure is continued until the residual force converges to zero. The stress distribution at all nodes is automatically implemented in the finite element program.

The vector of the cohesive forces in the nodes is given by:

$$F = B^T \sigma_d dA$$  \hspace{1cm} (10)

where $\sigma_d = [\sigma_T]^T$ is the stress vector due to the proposed constitutive model for normal and shear stresses in the FPZ. By using Gaussian integration, employing linear shape functions, and setting the determinant of the Jacobian matrix equal to $1/2$, the stiffness of the interface element and nodal cohesive forces can be obtained.

The normal stress, $\sigma$, is obtained from Peterson’s constitutive law (1981) according to the on the interface element. For shear stress, $\tau$, Bazant and Gambarova (1980) assumed that opening displacement takes place prior to the occurrence of the slip. However, in the present study, based on the shear stress versus the crack sliding displacement (CSD) curve obtained by Yoshikawa et al (1989), it is proposed that when the normal stiffness is in the elastic zone, the tangential stiffness value is very large and constant. However, when the normal stress in the softening part decreases, the tangential stiffness starts to reduce (Figure 4). It seems more logical to assume that sliding starts when the normal stress begins to decrease.

To estimate the maximum shear stress, $\tau_{max}$, the initial slip, $S_0$, and the critical slip $S_c$ shown in Figure 4, the following method is used. Previous experimental results have shown that the relationship between the slide and opening and displacements can be approximated as a linear relationship (Paulay & Loebel 1974). The relationship between COD and CSD is assumed to be given by:

$$\delta_s = a(\delta_n - \delta_n^{*})$$  \hspace{1cm} (11)

where $\delta_s$ and $\delta_n$ are the slide and opening displacements respectively, the parameter $a$ is greater than one, and $\delta_n^{*}$ is a constant. Based on Eq (11), the values of $S_0$ and $S_c$ (Figure 4) are assumed as $a(w_0 - \delta_n^{*})$ and $a(w_0 - \delta_n^{*})$ respectively. In addition, from Eq (7) the maximum shear stress, $\tau_{max}$ is given by:

$$\tau_{max} = \frac{a(w_0 - \delta_n^{*})}{\mu \beta D_{nn}}$$  \hspace{1cm} (12)

 Crack propagation criterion

The Griffith criterion for crack propagation is:

$$\frac{dU}{dA} - \frac{d}{dA} (U - W + E_j) > 0, \text{ Crack does not propagate}$$

$$\frac{dU}{dA} - \frac{d}{dA} (U - W + E_j) = 0, \text{ Critical condition}$$

$$\frac{dU}{dA} - \frac{d}{dA} (U - W + E_j) < 0, \text{ Crack propagates}$$  \hspace{1cm} (13)

where $U, W, E_j$ are the total potential energy, the strain energy, the work done by applying the load, and the surface dissipated energy respectively. Note that $\frac{dW}{dA}$ is given by (Xie & Gerstle 1995):

$$\frac{dW}{dA} = \int \left( \frac{\delta u_T^N M_T}{dA} + \frac{\delta u_T^N N_T}{dA} \right) dS + \int \left( \frac{\delta u_T^N M_T}{dA} + \frac{\delta u_T^N N_T}{dA} \right) dV$$  \hspace{1cm} (14)

Here $u$ and $N$ are the nodal displacement vector and the shape function matrix respectively, the parameter $t$ is the surface load on the surface $S_t$ and $b$ is the body force in the volume $V$. We can also write $\frac{dW}{dA}$ as:

$$\frac{dW}{dA} = \int \left( \frac{\delta u_T^N M_T}{dA} + \frac{\delta u_T^N N_T}{dA} \right) \sigma_i dA$$  \hspace{1cm} (15)

where matrix $M$ is the linear shape function matrix which depends on the crack opening displacement and is related to the geometry of the crack (Xie & Gerstle 1995).

In this study, the effect of the softening strain energy is considered in order to determine the strain energy release rate. The strain energy release rate is assumed to be the same in both mixed mode and Mode I (Bocca et al 1991). The rate of change in the strain energy in the interface element with respect to area is expressed as:

$$\frac{dU}{dA} = \frac{d}{dA} \int \sigma^T dV = \int \frac{d}{dA} \int \sigma^T dV$$  \hspace{1cm} (16)

where $\sigma$ and $\epsilon$ are normal stress and strain respectively. On the other hand, the strain energy density $W_{ij}$ is written as:

$$W_{ij} = \int \sigma^T dV = \int \sigma^T \left( \frac{S}{E_{ij}} + d\epsilon \right)$$

$$= \int \sigma^T \left( \frac{S}{E_{ij}} d\epsilon + d\epsilon \right)$$  \hspace{1cm} (17)
where $\varepsilon'$ and $\varepsilon''$ are the elastic strain and the softening strain respectively. The rate of change in the strain energy density is:

$$\delta W_d = \frac{\delta \varepsilon'}{\delta A} \int \left[ 1 - \frac{S}{E} \right] \sigma^2 \, dA$$

$$= \left[ 1 - \frac{S}{E} \right] \frac{\delta \varepsilon'}{\delta A} \int \sigma^2 \, dA$$

Substituting Eq (18) into Eq (16) yields:

$$\frac{\delta U}{\delta A} = \left[ 1 - \frac{S}{E} \right] \int \sigma^2 \delta \varepsilon' \, dV$$

Using the chain rule in partial differentiation, we obtain:

$$\frac{\delta \varepsilon'}{\delta A} = \frac{\delta (Bu)}{\delta A} \frac{\delta B}{\delta A} + \frac{\delta u}{\delta A} B$$

Thus, using Eq (20), we write Eq (19) as:

$$\frac{\delta U}{\delta A} = \left[ 1 - \frac{S}{E} \right] \int \left( \sigma^2 \frac{\delta u}{\delta A} B + \sigma^2 B \frac{\delta u}{\delta A} \right) \, dV$$

Substituting Eq (14), Eq (15) and Eq (22) into Eq (13) and rearranging the terms, we obtain the following:

$$\frac{\delta U}{\delta A} = \frac{\delta u T}{\delta A} \left[ \left( 1 - \frac{S}{E} \right) B^T D_B \frac{\delta u}{\delta A} dV + \int \sigma^2 \frac{\delta u}{\delta A} \, dA \right]$$

$$+ \int \delta N^T t \, dS + \int \sigma^2 \frac{\delta u}{\delta A} \, dV$$

$$= \frac{\delta u T}{\delta A} \left[ \left( 1 - \frac{S}{E} \right) B^T D_B \frac{\delta u}{\delta A} dV + \int \sigma^2 \frac{\delta u}{\delta A} \, dA \right]$$

$$+ \int \sigma^2 \frac{\delta u}{\delta A} \, dA$$

The first bracket is the equilibrium equation and is equal to zero. The body forces and surface load are small and can be ignored (second bracket). Hence, using the fact that (Xie & Gerstle 1995):

$$\int \frac{\delta u T}{\delta A} B^T D_B \frac{\delta u}{\delta A} \, dV = \frac{1}{2} \int \sigma^2 \frac{\delta u}{\delta A} \, dA$$

Eq (23) will be:

$$\frac{\delta U}{\delta A} = \frac{1}{2} \int \left( 1 - \frac{S}{E} \right) B^T D_B \frac{\delta u}{\delta A} - u^T \sigma^2 dA$$

In the finite difference method, in order to estimate the first part of Eq (25), only crack-tip stiffness elements can be used by applying crack extension (Yang & Liu 2008). The second part can be evaluated by using a Gaussian integration method (Xie & Gerstle 1995). Eq (25) can be applied to mixed-mode and multiple-crack fracture problems.

Furthermore, a stress-free region appears in front of the initial notch or macro-cracks when the FPZ length is fully propagated (Wua et al 2011). It has been shown that if the crack-opening displacement reaches 3.6 $G_c f$, the stress-free region appears in front of the initial notch. $G_c$ is the area under the curve (fracture energy) in Figure 1. Thus the stress-free region length is 0.08 times the ligament length, approximately $\bar{h} - \alpha_0$, where $\bar{h}$ and $\alpha_0$ are the depth of the beam and the length of the initial notch respectively. In finite element methods the length of the stress-free region is formulated by:

$$a_{n=0} = n \times L$$

where $n$ is the number of elements that fail behind the crack. When the FPZ is fully propagated, $n$ elements will be set to zero behind the crack and the crack will grow along the respective elements upon appropriate identification of the cracking direction in each step, as described in the next section.

**Crack propagation direction**

One of the most important aspects in discrete cracking is the direction of propagation. In this investigation, the crack is assumed to follow the existing inter-element boundaries and no re-meshing algorithm is needed to decrease the computational complexity (Alfaiate et al 1997). This method has a simple algorithm. It is assumed that the directions of maximum principal tensile stresses, which were automatically known in the program at each step, were perpendicular to the crack propagation direction (Xie & Gerstle 1995). Thus the angle ($\gamma$) is recognised (Figure 5). Crack propagation follows one of the inter-elements (AB) or (AC). It is assumed that the crack will not stop or intersect the main element and that the direction of propagation is perpendicular to the maximum tensile principal stress (Figure 5). There are two possible cases for the crack path: if the orientation angle ($\gamma$) is less than 45°, the path of growth is (AC), otherwise it will be (AB). The stiffness matrix, nodal forces and displacements of the interface element are changed from the local system to the global system by using the transformation matrix. Although the crack paths are non-smooth, the ones found with this method are in good agreement with the correct crack path.

The FEAPpv program code was developed to analyse cracks in concrete (Taylor 2009). Four-node isoparametric elements are used for bulk concrete, for which the material behaviour is considered to be linear elastic. Figure 6 shows the major steps used to solve fracture in the beam in the present numerical model. To model the post-peak curve of the structure, displacement was assumed to be incremented rather than the load, i.e. displacement-controlled numerical analysis was applied. The displacements were classically those for nodes at the crack mouth on the modelling of crack tip behaviour. The non-linear dynamic relaxation method is implemented to find the load-displacement curve. This method is preferred to other methods such as Newton (Gerstle & Xie 1992).

**NUMERICAL EXAMPLES**

Three benchmark test specimens were analysed to validate the model. Figure 7...
illustrates the previously tested plain beam used to simulate mixed-mode fracture (Arrea & Ingraffea 1982). The boundary conditions and material properties are indicated in Figure 7.

The Young’s modulus, Poisson’s ratio, and tensile strength of concrete were assumed to be 24 800, 0.18 and 4 MPa respectively. The thickness of the beam was 152 mm and the length of the initial notch was 82 mm. The parameter values of fracture were $G_c = 150$ N/m, $w_c = 0.135$ mm and $w_0 = 0.0001$ mm. Four-node isoparametric elements were used for the bulk concrete with linear elastic behaviour, and plane stress was considered as the analysis condition. The values of parameters $\mu$, $\beta$, $\alpha$ and $\delta_n^*$ were chosen as 1.16, 1.64, 0.5 and 0.1 respectively. The initial mesh (c) is illustrated in Figure 8.

Figure 9 shows the result of the load versus the crack mouth sliding displacement (CMSDD) curve for the beam, the experiment of Arrea and Ingraffea (1982) and the numerical model of Xie et al (1995).

The round dots represent the experimental envelope by Arrea and Ingraffea (1982), the results of Xie and Gerstle (1995) are shown in black dashed lines and the results of the proposed model are shown by coloured lines. As seen in Figure 9, the results of the proposed model show good agreement with those of the experimental method. Mesh (a) has 110 interface elements, mesh (b) has 225 interface elements, and mesh (c) has 306 interface elements. Approximate matching of the three curves demonstrates independence of the model from the mesh size and shows fast convergence of the proposed model. It can be seen from Figure 9 that the peak loads are close to each other, although the mesh size changes.

In the elastic part, the results lay close to the midpoint of the experimental results obtained previously. However, the peak load obtained by the numerical method is slightly shifted below the upper limit of the envelope. The difference between the data of the proposed model and the experimental data is inevitable since the behaviour of concrete is assumed to be linear elastic in fracture mechanics, but in fact it is nonlinear plastic; compression crushing has also been ignored. The peak load obtained by the proposed method differs by almost 7% from that of the experimental method. The peak load in the numerical model is considerably different from that of the experimental peak load.

It is seen that after the peak load, the curves in the softening zone (up to 57 kN) are closer to the experimental data than to the numerical model (Xie et al 1995), which is slightly more brittle thereafter. In the softening zone after 60 kN, the proposed model shows more agreement with the experimental data in terms of ductility. This may be because the stress-free zone in the tip of the notch was not considered in previous models.

Figure 10 shows the predicted crack path in mesh (c) compared with the experimental data. The FPZ propagation elements are shown in red lines, while the stress-free elements are displayed in black. It should be noted that the crack path is a smooth
curve, although in this study the crack path is illustrated by straight lines. It can be seen that the predicted crack path in the mesh (c) is very close to the experimental result.

To further validate the proposed numerical method in mode I cracking, the experimental data reported by Prasada and Krishnamoorthy (2002) were chosen. The test arrangement, the boundary conditions and the geometry of the RC beam are illustrated in Figure 11. The Young’s modulus, compressive strength, tensile strength and fracture energy were 29 270 MPa, 30.1 MPa, 4.11 MPa, 100 N/m respectively. The yield strength of the steel was 395 MPa. Other parameters were assumed to be the same as those used in the previous example. The bond between the bars and the concrete was assumed to be perfect, i.e. no bond-slip was considered.

The data from the numerical analysis are compared with experimental records and the numerical model information of Prasada and Krishnamoorthy (2002) (Figure 12). The numerical results are in good agreement with the upper limit of the experimental envelop. This could be due to the assumptions in fracture mechanics such as the linear elastic behaviour of concrete and tension cracks which are not available in practice, especially for low loads. In this study, the bond-slip of the steel bars was ignored, unlike in the experimental test and the numerical model by Prasada and Krishnamoorthy (2002). This is why there is a slight over-estimation in the numerical model for a load of about 13.2 kN.

The FPZ appeared in front of the notch tip at a loading of about 2.5 kN, while deflection at mid-span was 0.012 mm. The FPZ began to grow with increasing deflection at mid-span as the load increased. Since the crack-opening displacement was smaller than \( 3.6 \frac{G_c}{f_t} \), the FPZ did not fully propagate. In the initial stages, the load was sustained by the FPZ and reinforcement bars were not involved yet. When the load was 13.2 kN, the reinforcement bars arrested the crack. As shown in Figure 12, at a load of 19 kN the stiffness of the beam was slightly reduced. This may be because the stress in the reinforcement bars reached the yield stress. At a load of 23.4 kN, the crack-opening displacement was equal to \( 3.6 \frac{G_c}{f_t} \). Thus the FPZ completely propagated and a
stress-free crack was created in front of the notch tip.

Figure 13 shows the FPZ length, the length of the stress-free region and the COMD at the final load. The CMOD is 0.468 mm, and the deflection at mid-span is 0.576 mm at a loading of 28.1 kN. The FPZ propagation reaches to almost three-quarters of the beam depth after the thirteenth loading step. At the tenth loading step the FPZ fully propagates and the stress-free region length appears.

In addition, a reinforced concrete beam with simple supports (Figure 14), which was tested by Bresler and Scordelis (1963), was analysed using the proposed model. The RC beam was 4 572 mm long and 305.8 mm thick. The modulus of elasticity and the Poisson’s ratio of concrete are 24 000 MPa and 0.18 MPa respectively. The modulus of elasticity , the Poisson’s ratio, the cross-sectional area and the yield strength of steel are 200 GPa, 0.3 MPa, 3 290 mm² and 552 MPa respectively. The tensile strength of concrete, plastic deformation and the compression crushing, the nonlinear behaviour of concrete, plastic deformation and the bond-slip of bar concrete.

Initially, a few flexural cracks appeared near the mid-span perpendicular to the longitudinal axis when the loading reached about 55 kN. The length of the biggest crack for this load was 190 mm. The width of the flexural cracks was greater than the shear crack width, which occurred near the supports. The first flexural crack reached 245 mm in length and 0.168 mm in width at a loading of 100 kN. In one-quarter of the span to mid-span the cracks tended to grow towards the loading point. When the load was about 200 kN, the first crack propagated with a length of about 320 mm and a width of 0.187 mm. A shear crack was observed in the vicinity of the support and its width was greater than that of the flexural crack. Figure 16(a) shows the crack patterns in the experimental study of Bresler and Scordelis (1963) and Figure 16(b) illustrates the crack paths in this study at a loading of around 285 kN. As can be seen, both shear and flexural crack formation resemble the experimental data. Initially, the cracks grow straight and then slowly propagate towards the loading point. The initiation and location of some of these cracks may change due to the mesh size.

**CONCLUSIONS**

In this study an alternative stiffness matrix was applied to model the FPZ. The relationship between the normal stress and the shear transfer was considered in the implementation of the finite element method. A new constitutive model for shear stress was proposed in a four-node interface element. The energy-based crack propagation criterion was then improved by using a new stiffness matrix. The load-deflection curve in this numerical model and the curve in a previous experimental study were in good agreement. The crack directions at the tensile face in three recent experimental data sets and the present study were close to one another. Several case studies were considered and global load-deflection responses computed with the proposed model being in reasonable agreement with the results found in the literature. Therefore it can be concluded that the model is applicable, as it has been verified computationally that it is sufficiently able to predict the crack pattern.
REFERENCES
Taylor, R 2009. FEAPpv source, A finite element analysis program, personal version. University of California, Berkeley, US.
Friction behaviour of fresh concrete in the vicinity of formwork

S Bouharoun

This study seeks to understand the mechanisms that occur at the concrete/formwork interface when fresh concrete is poured into the formwork. Four ordinary concretes with paste volumes of 28, 30, 32 and 34%, and four concretes at different dosages of superplasticiser were studied. The rheological properties of fresh concrete were measured using the ICAR rheometer. A detailed study of the friction at the concrete/formwork interfaces was then performed using a plane/plane tribometer. The results highlight the role of the volume of paste and the superplasticiser dosage on the tribological behaviour of fresh concrete. The evolution of friction stress is linear (Coulomb’s Law) regardless of the formulations tested in this study. In addition, these friction stresses at the concrete/formwork interface depend on the paste volume and the dosage of superplasticiser in the concrete.

INTRODUCTION

Knowledge of the rheological and tribological characteristics of fresh concrete is very important to obtain high-quality structures. Much research has been oriented towards volume in rheology to improve the casting process and to optimise the formulation of concretes. It is known that these materials are non-Newtonian, sometimes with yield stress and frequently thixotropic (Wallevik 2003; Wallevik 2005; Helnass Moussa 2009). Their behaviour in a steady state was defined using a pseudo-plastic or Herschel-Bulkley model (De Larrard et al 1998; Ferraris 1999).

However, few studies have been conducted to date on the friction behaviour of fresh concrete in the vicinity of the formwork surfaces. The friction at the concrete/formwork interface can be considered a favourable factor in that it reduces lateral pressure exerted on the formwork, but it can also be unfavourable to the quality of the concrete surface (Vanhove 2001; Vanhove et al 2004; Kwon et al 2010; Libessart 2006; De Caro et al 2007). In recent years, some studies have reported on the friction effect on formwork pressure. Prosko (2007) modelled the friction effect based on the increase in the friction coefficient over time. Vanhove (2001) also analysed the effect of formwork friction based on the friction coefficient. These researchers measured the friction by means of a tribometer specially designed for a complex medium such as fresh concrete, in which the friction coefficient is known to be dependent on the roughness of the formwork surface. Some studies have shown that there are many factors influencing the tribological behaviour at the interface, such as mixture proportions, admixtures, temperature, casting rate and height of formwork. These factors can be classified into two categories: intrinsic and extrinsic. The intrinsic factors are related only to the material characteristics, while the extrinsic factors include the external contributions such as formwork flexibility, wall roughness and external temperature.

Friction at concrete/formwork interfaces is related to the compactness of the granular skeleton and the fines content of the mixture. This phenomenon appears typically in the case of concrete mixtures with a very compact granular skeleton and a low or high binder content (Neville 1995; Bouharoun 2011). The water/binder ratio (w/b ratio) is a factor that modifies the rheological properties of the fresh concrete and affects the interface properties. The ability of the concrete to flow is improved when this ratio increases. However, at very high w/b ratios the mixtures are more exposed to the segregation phenomenon, with a possible blockage at the concrete/wall interface (Ngo et al 2010; Ngo et al 2011).

The workability of concrete is strongly linked to the paste volume of concretes. As an active element in concrete, the paste is unique in that it can fill the voids in the granular skeleton, as well as act as a lubricant (De Larrard et al 1998; Ferraris 1999). However, the movement of aggregates is limited by intergranular friction forces and by friction forces against formwork and pumping pipes (Koehler 2004; Kwon et al 2011).
Therefore, this paste facilitates the work on construction sites, offering sufficient fluidity for concrete in the formwork to flow under vibration and fill the formwork perfectly (Bouharoun 2011; Bouharoun et al. 2010).

The main purpose of the work presented in this paper is to highlight the effect of the formulation parameters of concrete on the friction behaviour at the concrete/formwork interface. In this way, the effect of paste volume and superplasticiser dosage on the evolution of the friction stress was investigated. Friction tests were carried out using a plane/plane tribometer that can reproduce the conditions on construction sites. Four concretes with paste volumes of 28, 30, 32 and 34%, and four concretes with superplasticiser were studied. The results show that the paste volume and superplasticiser dosage have a significant influence on the friction at the concrete/formwork interface.

### MATERIALS

#### Characteristics of binder

The cementitious material used in the mixtures studied was an ordinary Portland cement (CEM I 52.5 type N according to NF EN 197-1). The Blaine surface area of the cement was 4 000 cm²/g, and the density was 3.09 g/cm³. The chemical analysis of the clinkers and the phase compositions from the Bogue calculation are reported in Table 1. The limestone filler used was BETOCARB-MQ, which is essentially composed of carbonate (99.3%). This limestone filler is characterised by a Blaine fineness of 3 970 cm²/g, a density of 2.71 g/cm³ and a water content of 0.1%.

For a better comprehension of the percentage of fines in the volume of fresh concrete, Figure 1 shows the gradation curves of the cement and limestone filler as measured using a laser granulometer. The key point that emerges from the gradation results is the greater fineness of the limestone filler compared with that of the cement. The following observations can be made based on the gradation curves:

- 90% of particles have a diameter less than or equal to 37.4 μm for cement and 66 μm for limestone filler.

- 50% of particles have a diameter less than or equal to 14.3 μm for cement and 13.5 μm for limestone filler.

- 10% of particles have a diameter less than or equal to 1.6 μm for cement and 1.2 μm for limestone filler.

### Table 1 Chemical and mineralogical composition of the cements (% w/w)

<table>
<thead>
<tr>
<th>Component</th>
<th>Content (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiO₂</td>
<td>19.5</td>
</tr>
<tr>
<td>Al₂O₃</td>
<td>5.2</td>
</tr>
<tr>
<td>Fe₂O₃</td>
<td>2.3</td>
</tr>
<tr>
<td>CaO</td>
<td>64.2</td>
</tr>
<tr>
<td>MgO</td>
<td>0.9</td>
</tr>
<tr>
<td>SO₃</td>
<td>3.5</td>
</tr>
<tr>
<td>K₂O</td>
<td>1.07</td>
</tr>
<tr>
<td>Na₂O</td>
<td>0.07</td>
</tr>
<tr>
<td>C₃S Bogue</td>
<td>66</td>
</tr>
<tr>
<td>C₂S Bogue</td>
<td>13</td>
</tr>
<tr>
<td>C₃A Bogue</td>
<td>11</td>
</tr>
<tr>
<td>C₄AF Bogue</td>
<td>7</td>
</tr>
</tbody>
</table>

### Table 2 Composition of concretes

<table>
<thead>
<tr>
<th>Mixtures</th>
<th>Proportions of concretes (kg/m³)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1</td>
</tr>
<tr>
<td>Limestone filler</td>
<td>232</td>
</tr>
<tr>
<td>Sand 0/4</td>
<td>677</td>
</tr>
<tr>
<td>Crushed aggregate 4/8</td>
<td>287</td>
</tr>
<tr>
<td>Crushed aggregate 8/12.5</td>
<td>778</td>
</tr>
<tr>
<td>Water</td>
<td>176</td>
</tr>
<tr>
<td>Cement (C)</td>
<td>232</td>
</tr>
<tr>
<td>HRWRA (%Sp/C)</td>
<td>–</td>
</tr>
</tbody>
</table>

### Table 3 Properties of concretes

| Water/binder ratio | 0.57  | 0.57  | 0.57  | 0.57  | 0.5  | 0.43  | 0.5  | 0.43  |
| Crushed aggregate/sand | 1.27  | 1.27  | 1.27  | 1.27  | 1.27 | 1.27  | 1.27 | 1.27  |
| Paste volume (%) | 28   | 30   | 32   | 34   | 30   | 30   | 34   | 34   |
| Slump (mm) | 120  | 130  | 140  | 150  | 130  | 130  | 150  | 150  |
| Air content (%) | 1.9  | 1.9  | 2    | 2    | 2.6  | 3.1  | 3.4  | 4    |

For a better comprehension of the percentage of fines in the volume of fresh concrete, Figure 1 shows the gradation curves of the cement and limestone filler as measured using a laser granulometer. The key point that emerges from the gradation results is the greater fineness of the limestone filler compared with that of the cement. The following observations can be made based on the gradation curves:

- 90% of particles have a diameter less than or equal to 37.4 μm for cement and 66 μm for limestone filler.

- 50% of particles have a diameter less than or equal to 14.3 μm for cement and 13.5 μm for limestone filler.

- 10% of particles have a diameter less than or equal to 1.6 μm for cement and 1.2 μm for limestone filler.

#### Concrete compositions

The effect of paste volume was studied by formulating four ordinary concretes with 28, 30, 32 and 34% paste. These concretes were identified as C1, C2, C3 and C4 respectively. The water/binder ratio (w/b) and aggregate/
sand ratio (a/s) were held constant. To study the influence of the superplasticiser dosage, four concretes with superplasticiser were formulated (C5, C6, C7 and C8). These concretes were formulated based on ordinary concretes C2 and C4. Tables 2 and 3 show the formulation and properties of the concretes studied in this work.

The French standard NF P 18-404 (titled Concretes – Design, suitability and inspection testing – Specimen production and conservation) was used in this study. The concretes were mixed from dry materials, using the procedure shown in Figure 2.

A resting time of 10 min was allowed for the aggregates to absorb water. This precaution prevents the rheological properties of the concrete from being altered a few minutes after mixing. The mixer used to prepare the concretes was a DZ120 V DIEM. It was equipped with several blades to optimise the mixing. Slump and air content tests were conducted on each mix to check the workability of the concretes. To keep the same workability between the three concretes with superplasticiser, the water/(cement + filler) ratio was adapted to obtain the consistency class of S3 (very plastic according to EN 206-1).

**EXPERIMENTAL METHOD**

**Measurement of the rheological parameters**

The yield stress and plastic viscosity of each concrete were measured using an ICAR rheometer (Figure 3). This device is portable on the construction sites and was developed at the University of Texas (Koehler 2004). It consists of a four-bladed vane that is immersed into the concrete sample and rotated at a series of fixed speeds. The entire rheometer is approximately the size of a hand-drill and can either be operated by hand or secured into a fixed position above a standard container.

The concrete mixture was placed in a 360 mm-diameter container and filled to a height of 300 mm. The vane, which measured 127 mm in diameter and 127 mm in height, was positioned in the centre of the concrete sample, between the vane and the sidewalls, with a gap of 127 mm above and below the vane.

For each concrete mixture, the protocol tests begins with a pre-shearing test, followed by a rest period so that the concrete can be completely restructured. The concrete is then sheared at a velocity of 3.14 rad/s for 20 s to obtain a complete breakdown of the material. Then, ramping down from 3.14 rad/s to 0.314 rad/s, in steps of 0.0684 rad/s, is applied. The resulting data were analysed on the basis of the Bingham model (Equation 1), whereby a straight line was fitted to the plot of torque, T (Nm) versus rotation speed, N (rad/s):

\[ T = Y + VN \] (1)

The intercept, Y (Nm), and the slope, V (Nm.s), of this line were considered to be related to yield stress and plastic viscosity respectively. Due to the geometry of this rheometer, it is not possible to determine the shear rate analytically in fundamental units (Koehler 2004). In this model, the shear stress is a linear function of shear rate. The expression of the model is as follows:

\[ \tau = \tau_0 + \mu \dot{\gamma} \] (2)

where

- \( \tau_0 \) = yield stress
- \( \mu \) = plastic viscosity
- \( \tau \) = shear stress
- \( \dot{\gamma} \) represents the shear rate.

The calculation of yield stress (\( \tau_0 \) in Pa) and the plastic viscosity (\( \mu \) in Pa.s) was performed using the Reiner-Riwlin equation according to the two following cases:

- The material is completely sheared:
  \[ \Omega = \frac{T}{4\pi \mu \ln \left( \frac{R_2}{R_1} \right)} \]

- The material is partially sheared:
  \[ \Omega = \frac{T}{4\pi \mu \left( \frac{1}{R_1} - \frac{2\mu \tau_0}{T} \right)} \]

![Figure 2 Mixing procedure](image)

![Figure 3 ICAR rheometer](image)

![Figure 4 Plane/plane tribometer (Djelal et al 2004)](image)
with:

\[ R_2^2 = R_{2,\text{eff}} = \sqrt{\frac{T}{2nk\tau_0}} \]  

where

\( \Omega = \) rotation speed (rad/s)
\( T = \) torque (Nm)
\( h = \) vane height (m)
\( R_1 \) and \( R_2 \) represent respectively the radius of the vane and the radius of the container
\( R_{2,\text{eff}} \) is the effective radius that separates the flow zone and dead zone near the wall.

To solve this equation, a non-linear optimisation must be carried out.

**Measurement of the friction stress**

The principle of this device was inspired by the tribometer used in soil mechanics. It can reproduce the conditions encountered by manufacturers of concrete walls and precast elements. In particular, it can create sliding contacts between concrete, release agent and formwork (Djelal et al. 2004; Vanhove et al. 2004).

Two cylinders 120 mm in diameter, with the concrete inside, were placed either side of a metal plate. The sample holders were fitted with a gasket system to prevent water egress (Figure 4). The plate was set in motion using a motor coupled to an endless screw. Plate travel was 800 mm. The concrete was pressured against the plate by a jack. It is important to note that the force is transmitted to the piston by a jack rod (Figure 5).

The frictional, or tangential, stress was calculated by the equation:

\[ \tau_f = \frac{\text{Measured force} - \text{Parasitic force}}{\text{Section of sample} - \text{Holder}} \]

\[ = \frac{F_{\text{mes}} - F_{\text{par}}}{S_c} \]

In this equation \( F_{\text{par}} \) is the resultant of the parasitic frictional forces due to the watertight system against the plate. The area in contact between the concrete and the plate is calculated from the diameter of the sample holder. In our case, this area is \( S_c = 113.1 \text{ cm}^2 \).

In this case, the experimental settings were selected according to the conditions at the job site.

- A metallic plate was cut out of the wall of formwork. Several measurements were taken on both sides of the plate using a portable roughness meter with graphic display to determine the average roughness. The results obtained indicated that the height of the maximum peak-to-valley profile (\( R_t \)) ranged from 4 to 19 \( \mu \text{m} \). The arithmetic mean of the profile deviations from the mean (\( R_a \)) was between 0.7 and 1.7 \( \mu \text{m} \). The average of these parameters was \( R_a = 1 \mu \text{m} \) and \( R_t = 7 \mu \text{m} \).

- The sliding velocity was fixed at 6 m/h. It simulates the vertical displacement of the concrete after placement in real formwork.

- The magnitude of the load increment depended on the casting rate. In the friction tests, the vertical stresses applied were normally 30 ± 2, 50 ± 2, 70 ± 2 and 90 ± 2 kPa, which corresponds to the lateral pressures applied by the concrete to the formwork at heights of 1.2 and 3.6 m.

**RESULTS AND DISCUSSION**

**Rheological characteristics**

The flow curves obtained from the expression of torque as a function of the rotation speed are shown in Figures 6 and 7. The values of the yield stress and the plastic viscosity are deduced from these flow curves, as described by the Bingham model, by calculating the intercept of the linear fit for the yield stress and the slope as the plastic viscosity. The rheological properties are summarised in Table 4.

The results indicate that the yield stress and the plastic viscosity decrease when the paste volume increases. The paste improves the deformability of fresh concrete. The yield stress depends on the friction between the grains that form the granular skeleton of concrete (De Larrard 1999). The shear stress of the fresh concrete is expressed in two components: a stress related to the viscous dissipation in the paste and a stress related to the dissipation by friction between grains. The Bingham model, commonly used for predicting the rheological behaviour of the concretes, is used to connect the two components.
mentioned above. For the concretes tested, the paste volume was increased by keeping the w/b and a/s ratios constant in the mixtures. This increase contributes to reducing the intergranular friction by lubricating the aggregate/aggregate contact to decrease the yield stress and the plastic viscosity.

To keep the same workability, the superplasticiser was incorporated into the concrete formulation by reducing the w/b ratio. Decreasing the quantity of water in the mixture has a significant effect on the evolution of yield stress and plastic viscosity. Moreover, the superplasticiser reduces the yield stress and increases plastic viscosity. The evolution of the yield stress can be explained by the properties of the superplasticiser, which disperses the cement grains by steric effect and electrostatic repulsion (Nicoleau 2004; Bethmont 2008). Indeed, superplasticiser molecules are adsorbed on the surface of the cement particles giving them a negative electrostatic charge. In this way, the cement particles repel each other to deflocculate in the cementitious mixture. This phenomenon reduces the intergranular friction by freeing the amount of water trapped by the flocs and gives the fresh concrete a better fluidity. Increasing the superplasticiser dosage allows more water trapped in the cement flocs to be freed, reducing the yield stress. The increase in the plastic viscosity as a function of superplasticiser dosage is directly related to the decrease in the w/b ratio. This reduction causes an increase in solids concentration in the medium and favours an increase in the plastic viscosity of fresh concrete (De Larrard 2002).

TRIBOLOGICAL APPROACH

To understand the mechanisms at the concrete/formwork interface linked to the paste volume and the superplasticiser dosage, Figure 8 shows the recording of the tangential force as a function of time for different contact pressures. The results show that the shape of the curves is similar, regardless of the pressure. It is important to note that the sliding velocity used in this study is constant. Therefore the representation of the tangential force according to displacement or time is comparable.

These graphs can be decomposed into two zones (Vanhove 2001):

- Zone I shows a growth in friction stress due to the start-up time of the engine to catch up with gaps and the elastic response of the mechanical system.
- Zone II reflects a stationary regime. Friction is nearly constant during the test. The stresses of static and dynamic friction are similar for this concrete/formwork interface. The friction in this zone is the friction that will be taken into account.

Figure 9 shows the evolution of the friction stress as a function of contact pressure. Each point on this graph represents an average of five points in the dynamic regime (Zone II) of Figure 8. The evolution of the curve is linear and follows a friction law of the Coulomb type in the range of pressure studied.

The normal stress applied to the fresh concrete is transmitted to the granular phase (cement, limestone filler, sand and aggregate) and to the water. This pressure causes a migration of part of the liquid phase and the fines close to the wall to fill the voids left by large aggregates. The paste migrates through the network of aggregates, leading to the fines accumulating to form a layer at the interface (Figure 10). This layer, called the “boundary layer”, is composed of water and fine particles smaller than 80 μm. It is important to note that the wall effect also causes the formation of a concrete/formwork interface rich in fine particles and water from the rest of the mixture. In this case, the finest particles are pushed by each other near
In addition, the results indicate that the friction stress increases when the contact pressure increases. Normal stress is transmitted to the boundary layer by a chain of forces. This phenomenon involves grain-grain contacts which play a major role in the stress distribution at the interface and increase intergranular friction close to the wall (Silbert et al. 2002; Cambou et al. 2004). Therefore, when the contact pressure increases, the intensity of contact between the grains forming the mixture becomes more important.

The maximum peak-to-valley height of the roughness profile (Rt) of the metallic plate can reach 19 μm, allowing the trapping of fine particles (cement, filler) in its cavities. As trapped grains cannot move, they tend to force other grains up near the formwork surfaces. This phenomenon also creates a mechanical adherence between this thin boundary layer and the formwork, which increases with contact pressure. In this way, the medium becomes more rigid at the interface and a considerable shear force is required to create movement close to the formwork surfaces. During the displacement of the plate, the friction between particles acts as a resistance force. The movement of the plate leads to the breaking of the network formed by the fine particles. This configuration generates an increase in friction stresses at the concrete/formwork interface. This process indicates that the material undergoes internal shear near the formwork.

In this study, the friction between the fresh concrete and formwork surfaces was measured under pressure. The phenomenon of adhesion is created by the pressure and the wetting power of the fresh concrete close to the formwork.

**Effect of paste volume**

Figure 11 presents the evolution of the friction stress as a function of pressure for the concretes C1, C2, C3 and C4. The results show that the friction stress increases when the paste volume increases.

The cement grains in contact with water tend to agglomerate to form cement flocs. This phenomenon is linked to the presence of electrical charges on the grain surface. In the absence of a water reducer, they tend to trap a certain amount of water inside the flocs of grain (Rixom et al. 1986). This trapped water moves with the agglomerate as a solid (Coussot et al. 2002).

The amount of fines in the vicinity of the formwork is greater for concretes with a higher paste volume, regardless of the contact pressure. Indeed, when the percentage of paste increases, the number of flocs in the concrete becomes higher, leading to congestion in the boundary layer. Under the effect of contact pressure, these agglomerates have more difficulty in moving. The boundary layer becomes more rigid, which increases the friction stress.

In the concretes’ composition, the paste volume was increased by keeping the w/b ratio at 0.57 and by reducing the amount of aggregate. Therefore, the quantity of fines becomes more important in mixtures with a thicker boundary layer at the interface.
– 72.5% of the cement and the limestone filler is less than the maximum peak-to-valley height of the roughness profile (Rt = 20 μm). These phenomena also increase the probability of fine particles being trapped in the formwork roughness.

The volume of paste at the concrete/formwork interface did not act as a lubricant as in the case of rheology. In this case, the friction interface is not governed by the plastic viscosity. The contact between grains that occurs in the shear plane therefore determines the friction stress at the interface. It is important to note that the shear plane occurs between the boundary layer and the formwork surface.

Similar results were found in a study conducted by Neville (Neville 1995; Bouharoun 2011) on the fresh concrete friction against a metal plate. An increase in friction can occur in the case of mixtures with high fines contents. In this situation, the friction between solids and grains in the granular skeleton increases by direct contact and causes a blockage in the boundary layer.

**Effect of superplasticiser dosage**

Figures 13 and 14 illustrate the evolution of the friction stress as a function of contact pressure for the six concretes (C2, C4, C5, C6, C7 and C8). These curves highlight the influence of the superplasticiser dosage in concrete with constant paste volume. The results show that the friction stress decreases when the dosage of superplasticiser increases.

In the presence of the superplasticiser, the physical interpretation of the Bingham model indicates that the yield stress is generated by intergranular friction and the viscosity depends on the degree of deflocculation and the solids volumetric fraction. In this case, the parameter that can explain the evolution of the friction stress is the yield stress. Indeed, the deflocculation caused by the presence of the superplasticiser in concrete reduces the intergranular friction and thereby leads to a reduction in the yield stress. This phenomenon is reflected at the interface by a loss of rigidity of the boundary layer, which facilitates the mobility of cement grains and limestone filler in the vicinity of the formwork, which in turn facilitates deformation of the medium under shear. When the quantity of superplasticiser in concrete is sufficient, the dispersing power on the cement grains increases and causes an increase of fluidity in terms of their yield stress.

Moreover, the limestone filler particles are inert and help to fill the interstices left by the larger aggregates in concrete. Therefore, the superplasticiser acts only on the cement grains. With 69% of cement and 12% of limestone filler having a diameter less than 20 μm, the lubrication caused by the superplasticiser determines the tribological behaviour at the concrete/formwork interface. This phenomenon can also explain the role of cement grains in the presence of the superplasticiser.

**Dosage limit**

In order to better understand the combined effect of the superplasticiser and the paste volume, Figure 15 presents an example of the evolution of friction stress as a function of superplasticiser dosage at 90 kPa of contact pressure for the concretes C5, C6, C7 and C8. The other curves have the same shape regardless of the contact pressure. The results show that an increase in the superplasticiser dosage reduces the friction stress at the interface up to a limit value, beyond which the paste volume will not influence the evolution of the friction stress. This limit dosage is around 0.7%, after which the friction stress remains unchanged regardless the volume of paste. This phenomenon can be explained by the saturation of the boundary layer by the superplasticiser (cement grains are completely deflocculated in the medium) for these three concretes (C6, C7 and C8).

The value of the friction stress cannot be deduced by interpolation of lines at 0 kPa because the behaviour of the fresh concrete is unknown in this pressure range. The apparatus used is unable to measure the friction forces for pressures below 20 kPa.

**CONCLUSION**

The friction tests conducted in this work highlighted the influence of the volume of paste and the superplasticiser dosage on the
friction at the concrete/formwork interface. Firstly, the friction stresses at the interface increase with the paste volume. This augmentation is linked to the enrichment of the boundary layer by the fines. The thickness of this layer becomes greater, which increases the probability of cement grains being trapped in the roughness of the metal plate.

Secondly, a decrease in friction stress with an increase in the superplasticiser dosage indicates improved mobility of grains at the interface. It allows easier deformation of the medium under shear and leads to a decrease in the friction stress. In this configuration, the cement grains and the superplasticiser play a significant role at the interface in reducing friction. Moreover, a superplasticiser dosage of $0.7\%$ cancels the effect of the paste volume on the friction and allows an identical friction stress to be obtained for the concretes $C_6$ and $C_8$. Finally, without superplasticiser, the amount of the fines determines the rheological and tribological behaviour of fresh concrete. Indeed, a high content of paste reduces the rheological parameters and ensures that the concrete has better workability. However, the tribological properties of the concrete/formwork interface increase when the paste volume in concrete increases. Moreover, for concretes with superplasticiser, the yield stress and the friction stress are governed by the lubricating role of this component in volume and in surface contact.

REFERENCES


Figure 15 Evolution of friction stress as a function of superplasticiser dosage at 90 kPa of contact pressure for concretes $C_5$, $C_6$, $C_7$ and $C_8$ ($V_p$: paste volume)
The integration of railway asset management information to ensure maintenance effectiveness

N J van der Westhuizen, P J Grabe

The challenge to sustain railways depends on the ability to optimise the utilisation of the asset base. It demands the establishment of a continuous asset improvement process which requires maintenance personnel continuously to improve their understanding of the infrastructure's performance and the relevance of the configuration to this performance. A successful asset management process incorporates these elements and results in the optimisation of the infrastructure life cycle by extending useful life while minimising the operational interference, thereby increasing capacity. The effect of condition-based maintenance is illustrated, signifying how this maintenance strategy increases maintenance effectiveness (doing the right things at the right places), thus decreasing maintenance costs. It is apparent that with an increase in maintenance effectiveness, less time will be spent on maintaining assets, resulting in an increase of asset availability adding to the business objectives, allowing for an increase in operations. The paper demonstrates how condition-based and condition performance-based maintenance can enable railway organisations to save up to roughly 30% on their maintenance costs, while increasing capacity to approximately the same value.

INTRODUCTION

Since the turn of this century, the Chair in Railway Engineering at the University of Pretoria, Transnet Freight Rail and Amtrak (USA) have collaborated on various infrastructure maintenance projects. During the course of these projects, the need to develop philosophies of railway infrastructure asset management became apparent. Some of the methods and processes that were developed and expanded over the years are presented here.

This paper supplements the work presented by Van der Westhuizen and Van der Westhuizen (2009) at the 28th Southern African Transport Conference. The first paper was an introduction to the literature, effective maintenance management models and the concept of continuous improvement and the current paper presents the results of the research.

The key to asset life cycle optimisation is a decision-support system for analysing maintenance requirements that will enable effective management (Van der Westhuizen 2012). In the case of a railway network, an asset base widely distributed over hundreds to thousands of kilometres has to be managed. It is known that the full potential of the existing railway capacity in South Africa and abroad is not completely utilised. According to Ebersohn and Ruppert (1998), the industry needs to consider other approaches to create additional capacity and ensure better utilisation of available capacity.

This paper illustrates how condition information and its utilisation contribute to effective maintenance management, enabling maintenance managers and engineers to determine when and where to invest resources for optimum maintenance of the railway infrastructure.

RAILWAY: AN ASSET-CENTRIC BUSINESS

Asset-intensive businesses usually rely on distinctly different asset groups to be concurrently available to support operations. Historically, these asset groups were managed in “silos”, concentrating on asset management challenges per responsibility group and neglecting the challenges to coordinate asset group availability to support service. Railway organisations have an even greater challenge considering the large areas over which they are required to deliver a service to customers. The decentralised nature of the railway networks that are traversed and maintained increases the complexities involved in managing the operational performance of the transport service effectively.

Since railways are the only transport business that typically own all the

Keywords: Information integration, asset management, maintenance, continuous improvement, condition-based, condition performance
components to provide a transport service (railway infrastructure, rolling stock, as well as the train control), transparency of operations management information continues to gain momentum as a shared strategic goal for these business units.

Railways should follow an integrated (between business units) and collaborative (strategic, tactical and operational) approach to cause-and-effect-based performance management. The authors favour an asset-centric approach, illustrated in Figure 1, whereby operations and maintenance are integrated elements of a system and need to be managed as such.

From the model it is evident that the traditional, historical approach of “silo” functional activity-centred organisations within the business is not effective and that there should be a paradigm shift in the business’s way of thinking. Top-performing transport companies are distinguished by greater transparency of management information. They make better use of performance data and measurement technology, and have stronger communication channels between senior management and frontline employees. They also create more transparent and timely management reporting and planning processes to facilitate better fact-based decision-making to realise business objectives (Van der Westhuizen 2012).

It is also noted that a paradigm shift from a “silo” functional activity-centred organisation will not be achieved through the transparency of information only, but will also be influenced by coordination between stakeholders, as well as commercial and/or regulatory changes related to the business.

**ASSET AND MAINTENANCE MANAGEMENT**

In general, asset management can be defined as a holistic approach within the entire organisation to improve its objectives. In other words, it ensures that performance-improvement initiatives are a collaborative effort across the strategic, tactical and operational levels in the organisation. Maintenance management is therefore defined as one of many elements within the asset management approach.

**Asset management**

According to Mitchell (2007), Physical Asset Management has a single objective, namely to increase the value and return delivered by the physical assets. It then follows that Physical Asset Optimisation is a programme or business initiative focusing on and determined by opportunities to create value in different areas. These opportunities can include aspects such as the reduction of cost and the improvement of availability or capacity. From this it can be appreciated why this approach requires a close partnership or synergy between production/operations (the asset owner) and maintenance/engineering (the asset caretaker).

Woodhouse (2001) defined asset management as a group of tools, processes, methods and disciplines that are used to optimise the service life cycle of a physical asset. Peterson (2007) described asset management as a process for asset-care decision-making.

Some aspects that were identified to improve the required service objectives for asset management optimisation are listed below (Mitchell 2007; Peterson 2007; Woodhouse 2001):

- Improve production availability
- Increase production
- Reduce operating cost
- Increase asset effectiveness
- Increase reliability and quality
- Develop and use flexible and reliable processes
- Improve efficiency.

These considerations all support the authors’ viewpoint that asset management is a process of continuous improvement. Continuous improvement will ensure that asset-centric businesses apply and implement a business initiative process to optimise the asset with a single tangible result in mind, namely to maximise the value and return delivered by the physical assets.

**Maintenance management**

Maintenance can be defined as the care, correction and servicing of assets and their...
components. The purpose of these interventions is to ensure the satisfactory operation of the assets and their components before physical or functional failure, or before major defects develop that could influence the effectiveness of operations. In short, maintenance can be defined as asset care. It ensures the availability and capacity required by the asset owner from the assets to deliver either a product or a service, depending on the business objective (Mitchell 2007).

According to Mitchell (2007) and Peterson (2007), the benefits of asset care within the asset-optimisation process include, but are not limited to:

- Reducing incidents
- Increasing utilisation
- Improving quality
- Maximising effectiveness
- Reducing and minimising failures
- Reducing operating and maintenance costs
- Reducing the need for spare parts
- Setting operational and maintenance goals.

It is therefore clear that, if maintenance or engineering partners synergistically with operations, value will be added to the operational profit. Maintenance should therefore become a core management responsibility within the asset-centric business, and the continuous participation of maintenance in the business decision-making and production process is not negotiable.

**Maintenance management life cycle**

Ebersöhn and Ruppert (1998) pointed out that infrastructure assets deteriorate under operational loads and various environmental conditions, requiring maintenance to ensure their availability for the required operations at an operational standard.

As part of the decision-making process, the maintenance manager should select among the following maintenance options, based on different maintenance strategies, to provide affordable assets of a required operational standard:

- **Rehabilitation** is required to maintain or restore the facility at or to a serviceable condition or status (e.g. repainting the infrastructure, realigning the track or repainting a steel bridge). These actions are typically reflected as operating expenses.
- **Renewal** is the replacement of the structure or its components with comparable new structures or components when condition and reliability improvements are required (e.g. replacing a worn rail with a new rail or resurfacing a road). These are typically reflected as capital expenses.
- **Upgrading** is the reconstruction of equipment, components or facilities to improve or enhance their physical functionality. It is implemented when enhanced performance and reliability are required (e.g. replacing the track structure with stronger sleepers and heavier rail or improving the substructure to carry heavier loads). These are typically reflected as capital expenses.

To illustrate the maintenance management life cycle, a hypothetical example of infrastructure asset deterioration trends is shown in Figure 2. The graphs illustrate the effects of various maintenance activities on asset performance. These are represented by changes in the asset’s roughness (an indication of condition) over its life cycle, with respect to the traffic volume or million gross tonnes (MGTs) which can be transposed into time.

**Maintenance management process**

Mitchell (2007) argues that leaders within asset-centric businesses consider maintenance as an integral part of the operations and business processes, thereby generating and adding value to the business. The importance of a well-defined and well-documented maintenance process to assist in the delivery of the business objectives cannot be overstressed. Dunn (1997) categorises the maintenance management decision-making process into six phases:

1. **Work identification**
2. **Work planning**
3. **Work scheduling**
4. **Work execution**
5. **Recording work history**
6. **Analysis**

At the end of the sixth phase, a feedback loop follows, whereby the decision-making process restarts at Phase 1 with the identification of new maintenance requirements.

**CONTINUOUS IMPROVEMENT PROCESS: EFFECTIVE MAINTENANCE MANAGEMENT MODEL**

The maintenance management process explained above relates to the familiar management philosophy, Theory of Constraints (Goldratt 2004). It states that at any given point in time, at least one constraint limits the performance of the system. As the process repeats itself over time, the constraint may change, but the same constraint may also reappear over time.

The authors designated this step as “continuous improvement”; it is the platform where opportunity arises to improve the existing and/or current methods and strategies for asset management optimisation (Van der Westhuizen 2012).

Ebersöhn and Ruppert (1998), as well as Woodhouse (2001), include these phases in their maintenance cycle. At first these phases seem somewhat different, but the principles of maintenance management are clarified when the maintenance process models are analysed. The fundamentals of the maintenance management process are therefore confirmed. The continuous improvement maintenance process, as extended from the basic maintenance management process by Dunn (1997), is presented in Figure 3.

**MAINTENANCE STRATEGIES**

Within maintenance, asset management optimisation requires a combination of maintenance strategies to minimise...
Typical and well-known maintenance strategies used for asset optimisation include:
- Corrective maintenance
- Planned/routine maintenance
- Condition-based maintenance.

The advantages of progressing from reactive or corrective maintenance to condition-based maintenance are appreciated by industry and result in increased effectiveness, as well as a decrease in maintenance costs.

**Corrective maintenance**

"Fix it when it breaks!" would be the layman’s term for corrective maintenance. Mitchell (2007) describes corrective maintenance as encompassing problems usually identified by production and/or operations. In the railways industry, examples of corrective maintenance include:
- Derailments
- Rail breaks
- Earthwork slips
- Wash-away of substructure components
- Signalling or electrification failures.

The corrective maintenance tactic is by far the most costly and must be minimised to increase effectiveness and at the same time decrease operational cost. Corrective maintenance has a direct relationship with uncertainty. It is therefore apparent that an organisation utilising a corrective maintenance tactic to deal with a high number of occurrences will have a high uncertainty value, impacting negatively on the reliability of operations, service delivery and operational profits. It addition, corrective maintenance often leads to downtime in production, thereby limiting the capacity available to deliver the required service.

**Preventive maintenance**

Preventive maintenance includes both routine-based and condition-based maintenance, and forms the basis for the maintenance approach followed in this paper. Preventive maintenance can be defined as a tactic that is applied to mitigate failure. In contrast to corrective maintenance, preventive maintenance assists in reducing long downtimes, thereby ensuring the availability of capacity, a decrease in uncertainty and ultimately an increase in reliability and service delivery.

**Routine-based maintenance**

Routine-based maintenance is performed at time-based intervals, which can be either calendar-based or operating time-based. Note that in some cases, routine-based maintenance might include unnecessary maintenance activities, as they will not improve asset reliability. This translates into over-maintained assets, increasing the operational cost and therefore impacting negatively on the operational profit. Mitchell (2007) confirms this notion by stating that only up to 20% of failures are time base-related in the maintenance industry. Therefore, 80% of routine-based maintenance can be regarded as ineffective and unable to prevent failures. Under extreme circumstances, routine-based maintenance can even introduce failures.

This idea is consistent with that of Selig and Waters (1994) with regard to railway ballast maintenance utilising a tamping machine. They explain that the tamping action breaks down the ballast and introduces some additional functional failures of the ballast properties. These include:
- Ballast bed loosening, resulting in further settlement with additional traffic
- Initial reduction in vertical and horizontal resistance
- Increasing the degree of settlement as the ballast deteriorates.

**Condition-based maintenance**

After reviewing routine-based maintenance, it is apparent that some assets are “over-maintained” due to ineffective maintenance activities. Condition-based maintenance was therefore born out of the need to reduce fruitless over-maintenance and can be defined as having the objective to maintain the correct asset/equipment at the right time, thereby ensuring optimal levels of efficiency. As maintenance effectiveness increases, reliability and production increase, resulting in a decrease in the overall maintenance costs.

![Figure 3 The continuous improvement maintenance process (adopted from Dunn (1997))](image)
RAILWAY TERMINOLOGY
This section focuses on railway terminology and presents the fundamentals of permanent way condition parameters which will facilitate a high-level understanding of the components’ functions. It forms the basis of the decision-making analysis (explained later in the paper) to ensure effective maintenance, assisting the business in optimising the management of its assets. This section summarises the basics of rail track components as given by Selig and Waters (1994).

Track geometry terminology
Track geometry refers to the location each rail occupies in space. Track in the longitudinal direction is composed of various track characteristics, such as straight or tangent sections, horizontal curves (transition and circular curves) and vertical curves.

In practice, deviations (i.e. mid-chord offsets) from the design geometry are measured to determine the condition of the track geometry. These deviations contribute to the roughness of the track which determines the riding quality of the right-of-way. Roughness occurs for a large diversity of reasons and root-cause analysis is required to determine the reason(s) for these irregularities. The calculation of track roughness will be discussed later. In general, deviations from the original track design geometry occur due to:

- Variation of the substructure construction and therefore its stiffness
- Overloading of the track structure compared with the design load
- Localised weak spots in the track structure
- Track discontinuities, such as block joints, level crossings and turnouts.

The projection of the track geometry onto various planes (see Figure 4) enables track geometry parameters to be specified. These parameters can be measured and used to determine the condition of a track, highlighting areas with irregularities and requiring maintenance input.

Track geometry parameters are grouped as parameters in the horizontal, longitudinal, vertical, transverse vertical and track plane (located 15 mm below the top of both rails along the track centre line) and are described in the following sections.

Track geometry in the horizontal plane
The horizontal plane (see Figure 4) represents the track geometry parameter “alignment”. Alignment represents a rail with a line along the rail gauge side, 15 mm below the top of the rail. The design alignment of a track is measured in terms of the absolute geometrical location of the track in the horizontal plane. Mid-chord offset measurements of the alignment are also referred to as “versed measurements”.

Longitudinal vertical track geometry
The longitudinal vertical track geometry is called the “vertical profile”. Vertical profile is the projection of each rail onto the longitudinal vertical plane, as indicated in Figure 4. The line along the top of the rail is used for the projection.

The design longitudinal profile is measured in terms of the absolute elevation of each geometrical location of the longitudinal profile. These geometrical locations define the absolute vertical space curve. Mid-chord offset measurements of the vertical profile are also referred to as “top measurements”.

Track geometry in the transverse vertical plane
The transverse vertical plane describes two parameters that need to be managed by maintenance managers. These parameters, “superelevation” and “twist”, are indicated in Figure 5.

Superelevation is the difference in elevation between a point on one rail (S₁) and a point on the other rail (S₂) measured along a
line perpendicular to the track centre line, as indicated in Figure 5.

Twist is the difference in elevation of two points, one on either rail (T_a1 and T_a2 at position A, and T_b1 and T_b2 at position B), a fixed distance apart along the length of the track, as indicated in Figure 5. The distance between the two points (A and B) is referred to as the “twist base”.

Condition data

Track geometry assessment data are usually automated condition assessment data recorded by a track geometry car (e.g. the IM2000 in the case of Transnet Freight Rail). Numerous track geometry parameters can be measured to determine the condition of the track, thus revealing the health of the track structure. The track geometry condition assessment vehicle records these parameters at intervals of between 250 mm and 2 000 mm, depending on the parameter measured. A list of the parameters obtained and relating to the terminology discussed above is as follows:

- Vertical profile left and right
- Horizontal alignment left and right
- Superelevation
- Twist
- Gauge
- Curvature (indication of the radius)
- Radius

Roughness profile left and right
Roughness alignment left and right.

Track roughness

Track roughness is the sum of squares with variable summation lengths. This condition index was developed by Ebersöhn (1995) and is expressed in Equation (1) below:

\[
R^2 = \frac{\sum_{i=1}^{n} d_i^2}{n}
\]

where:

- \( n \) = number of measurements in the summation length
- \( d_i \) = the mid-chord measurement for profile and alignment and the deviation from a target value for cant, twist and gauge.

From Equation 1 it is clear that the Roughness Index is a variation measurement of the specific condition parameter under consideration.

As an example, the mid-chord measurements for a length of track including a good and a poor section are plotted in Figure 6. The corresponding running roughness was calculated using the mid-chord measurement values and a 50 m calculation length. The beginning and end of the good and poor sections can be clearly identified in the roughness plot.

**CONDITION ANALYSIS:**

**CONDITION-BASED AND CONDITION PERFORMANCE-BASED MODELS**

For tamping optimisation, track geometry condition measurement data are used to determine the condition of a track, indicating areas with irregularities requiring maintenance input. According to Gräbe and Maree (1997), the most important maintenance parameters that will influence the condition of track geometry related to tamping are the standard deviations of vertical profile and horizontal alignment. To demonstrate the different condition analysis models, a section of track on the Natal mainline between Pietermaritzburg and Durban in South Africa (number 1 line) was used.

To establish the effectiveness of condition-based maintenance compared with the generally accepted routine-based strategy, two tamping algorithms (condition-based and condition performance-based) were used to assist in the optimisation of the track tamping activity.

**Condition-based model**

The condition-based model assists in identifying maintenance requirements based on condition parameters that can be defined by the user. The condition-based algorithm consists of the following steps:
The selection of up to five track geometry condition assessment parameters to calculate a condition index. Each parameter can be multiplied by a weighted average, and these values are added to produce a single condition index. A variable maintenance intervention limit parameter is used to identify all areas where the condition index calculated exceeds the intervention limit. A variable cluster length is used to cluster maintenance areas into one group if the area not requiring maintenance between identified maintenance areas is less than this length. There is an option to include radius configuration data to ensure that the result is applied to the full extent of curves, thereby preventing undesirable, partial remedial work in curves. By using condition-based maintenance intervention limits and the analysis of condition data, the decision-maker is provided with a list of areas identified as requiring maintenance. From this list, effective maintenance actions can be identified and optimised according to maintenance activity rules, thus ensuring optimisation of the maintenance activity type. The authors focused on optimising track tamping and the results obtained from the condition-based algorithm, applied to a section of track, are presented in Figure 7.

From Figure 7 it is clear that the calculated maintenance intervention limits and the maintenance needs process result in an optimised and continuous maintenance action (tamping), compared with the scattered maintenance exception data. This confirms that the model will provide a result that will ensure that the majority of exceptions are attended to during the execution of maintenance activities. The condition-based algorithm also demonstrates that the analysis model can be used to identify effective maintenance management actions.

**Condition performance-based intervention maintenance needs**

A new method, termed “condition performance-based intervention”, has been developed by the authors. The proposed method will assist in the continuous improvement of track geometry conditions. The model assists in identifying maintenance needs through the identification of areas where the condition is deteriorating at a rate above a specified intervention limit. A condition index is calculated through two condition parameters, \( R_{2n-1}^2 \) and \( R_2^2 \) which can be defined by the user to provide agility to the maintenance identification requirements. These two parameters are subtracted from each other to determine the rate of change over a period; this results in the Condition Performance Index (CPI), also known as Condition Performance Index – \( \Delta R_2 \).
Delta Roughness. In areas where the CPI is greater than zero, the condition has deteriorated from the previous assessment at a rate equivalent to the CPI value and they represent areas requiring maintenance input. If areas have been maintained effectively, the CPI values will be negative, signifying that these areas have improved, resulting in an overall improvement of track quality for the section.

If this process is continuously implemented, the total track quality will improve over time, until it reaches its optimum condition (at which it cannot be further improved). Thereafter it will only be necessary to sustain the optimum condition.

The CPI methodology is described in the following process:

- **Step 1a:** Compare year-on-year roughness data by subtracting the consecutive yearly condition assessment data from each other to determine the areas that have deteriorated over the past year. The result is presented in the Equation (2):
  \[ \Delta R^2 = R^2_{n-1} - R^2_n \]
  where:
  \[ \Delta R^2 = \text{rate of change in the Condition Performance Index (CPI), also known as the Delta Roughness. All values smaller than zero indicate an improvement.} \]
  \[ R^2_{n-1} = \text{roughness condition index for the period } n-1 \]
  \[ R^2_n = \text{roughness condition index for the period } n. \]

- **Step 1b:** Identify maintenance needs using the \( \Delta R^2 \) profile and \( \Delta R^2 \) alignment intervention limits. An intervention limit of 3 mm\(^2\) is used in this case (see Figure 8).

- **Step 2:** Use a cluster length to ensure that where maintenance needs are identified for distances smaller than the defined cluster length, these needs will be optimised to group them together. A cluster length of 200 m is used (see Figure 9).

- **Step 3:** After this result, use the radius configuration information to ensure that if a portion of a curve needs tamping, the result will be extended to include the total curve (see Figure 10).

- **Step 4:** Compare \( \Delta R^2 \) profile and \( \Delta R^2 \) alignment and determine the tamping requirements to ensure that all needs are attended to.

- **Step 5:** Compare the condition-based and condition performance-based analyses and develop a holistic final tamping requirement plan.

**RESULTS OBTAINED FROM ANALYSIS**

The results are based on and evaluated against the generally accepted routine-based maintenance strategy used by the railway industry.
Baseline routine maintenance

In Transnet Freight Rail, most tamping requirements on a network level are routine-based, determined by an equation published by Hall (1985). To determine the time interval requirements for tamping, the equation uses the traffic characteristics of a section, as indicated in Equation (3):

\[
\text{Tamping interval} \approx \frac{48}{\sqrt{\text{MGT}_{\text{section}}} \text{/ year}} \tag{3}
\]

From the above, the MGT (million gross tons) per annum on the section used in the study (Pietermaritzburg–Durban line) amounts to 18 MGT. Using Equation (3), an interval of approximately 11.5 months is calculated. Pragmatically, it is decided to make the routine-based maintenance interval on this section exactly one year.

In addition to routine-based maintenance, Transnet Freight Rail's body of knowledge requires a double tamp action on all curves, as well as tamping of switches (turnout) using a switch-point tamper. One switch point is estimated to be equivalent to 400 m of mainline tamping.

From the analysis it is determined that there are 81 switches on the number one line. This is equivalent to 32.4 km of mainline tamping. Using these figures, the routine tamping requirements can be summarised as shown in Table 1.

It is calculated that at an anticipated tamping rate of 5 km per day, the yearly tamping maintenance activity can be conducted in 38 days.

Condition-based maintenance

A summary of the condition-based (R^2) tamping maintenance model, as determined from the integration of the condition-based analyses, is provided in Table 2. The results presented are an integrated (combined) result obtained from the analysis of the R^2 alignment and R^2 profile condition parameters. The method for determining the maintenance criteria (intervention limits) used in this analysis is discussed in more detail by Van der Westhuizen (2012). Figure 11 presents the integrated R^2 alignment and R^2 profile results in Transnet

---

### Table 1 Routine-based tamping requirements: Natal mainline (Pietermaritzburg–Durban)

<table>
<thead>
<tr>
<th>Equivalent routine-based track distance tamping requirements</th>
<th>Unit</th>
<th>Quantity</th>
<th>Equivalent tamping distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tangent track</td>
<td>km</td>
<td>42.0</td>
<td>42.0</td>
</tr>
<tr>
<td>Curved track</td>
<td>km</td>
<td>55.5</td>
<td>111.0</td>
</tr>
<tr>
<td>Switches</td>
<td>number</td>
<td>81.0</td>
<td>32.4</td>
</tr>
<tr>
<td>Total track distance</td>
<td></td>
<td>97.5</td>
<td>185.4</td>
</tr>
</tbody>
</table>

### Table 2 Tamping maintenance plan derived from integration of the R^2 alignment and R^2 profile analysis results: Natal mainline (Pietermaritzburg–Durban)

<table>
<thead>
<tr>
<th>Maintenance criteria</th>
<th>Distance (km)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roughness (alignment) criteria: 30 mm^2</td>
<td>31.2</td>
<td>32</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roughness (profile) criteria: 17 mm^2</td>
<td>50.2</td>
<td>52</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roughness (profile and alignment)</td>
<td>56.2</td>
<td>58</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total track distance</td>
<td>97.5</td>
<td>100</td>
</tr>
</tbody>
</table>

### Table 3 Tamping maintenance plan by integration of the ΔR^2 alignment and ΔR^2 profile results: Natal mainline (Pietermaritzburg–Durban)

<table>
<thead>
<tr>
<th>Maintenance criteria</th>
<th>Distance (km)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta Roughness (alignment) criteria: 3 mm^2</td>
<td>56.5</td>
<td>58</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta Roughness (profile) criteria: 3 mm^2</td>
<td>62.8</td>
<td>64</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delta Roughness (profile and alignment)</td>
<td>70.0</td>
<td>72</td>
</tr>
<tr>
<td>Cluster length: 200 m and radius integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total track distance</td>
<td>97.5</td>
<td>100</td>
</tr>
</tbody>
</table>

---

**Figure 11** R^2 alignment and R^2 profile analysis integration: Pietermaritzburg–Durban, km 20 to km 26
Freight Rail’s Infrastructure Asset Maintenance Management System for the Pietermaritzburg–Durban mainline section from km 20 to km 26.

As shown in Table 1, it is evident from the condition-based \((R^2)\) maintenance model that only 56.2 km (58%) of the total track distance requires maintenance, subsequently improving maintenance effectiveness by 42% compared with the baseline routine-based maintenance strategy.

**Condition performance-based maintenance**

A summary of the condition performance-based \((\Delta R^2)\) tamping maintenance model, as determined from the integration of the condition performance-based analysis, is provided in Table 3. The results presented are an integrated (combined) result obtained from the analysis of condition parameters \(\Delta R^2_{\text{alignment}}\), \(\Delta R^2_{\text{profile}}\), and \(R^2\). Figure 12 presents the integrated \(\Delta R^2_{\text{alignment}}\) and \(\Delta R^2_{\text{profile}}\) results in Transnet Freight Rail’s Infrastructure Asset Maintenance Management System for the Pietermaritzburg–Durban mainline section from km 20 to km 26.

It is evident from Table 3 that 70 km (72%) of the total track route distance requires maintenance according to the condition performance-based maintenance \((\Delta R^2)\) model, which is an increase of approximately 25% compared with the condition-based model (56.2 km or 58% of the total track distance).

The condition performance-based maintenance approach decreases the tamping distance from 97.5 km to 70.0 km, resulting in a maintenance effectiveness improvement of approximately 28% compared with the baseline routine-based maintenance strategy.

**INTEGRATING THE CONDITION-BASED AND CONDITION PERFORMANCE-BASED MAINTENANCE MODELS**

Table 4 presents a summarised view of the integration of the two maintenance models, condition-based \((R^2)\) and condition performance-based \((\Delta R^2)\) maintenance. The integration of the results provides a combined maintenance tamping model, ensuring a holistic approach to the maintenance needs and thereby optimising the maintenance requirements.

From Table 4 it is evident that when integrating the condition-based \((R^2)\) and condition performance-based \((\Delta R^2)\) maintenance tamping models, the increase is virtually negligible. When analysing the results as presented in Figure 13 it is, however, evident that both models are required. From the area denoted by A it can be seen that at location A1 (solid grey block), there is an improvement in the condition over the period under consideration (condition performance-based \(\Delta R^2\) is smaller than zero), resulting in no maintenance requirements in the condition window denoted by the Condition Performance Index \((\Delta R^2)\) limit areas. However, at the same location, the condition-based maintenance intervention analysis \((R^2)\) indicates (condition window denoted by condition-based \((R^2)\) limit areas) that a deficiency at this location is present and requires maintenance input.

In many instances, the areas identified as requiring maintenance, determined from both analysis methods (condition-based \((R^2)\) and Condition Performance Index \((\Delta R^2)\)), correspond.

The authors suggest using the two condition analysis models in phases. These will depend on the maturity level of the railway’s asset management in utilising condition-based maintenance planning practices. Firstly, it should be ensured that the basics have been implemented successfully, focusing on the condition-based model. As soon as maturity has been reached in preventive condition-based maintenance \((R^2)\) planning, the railway can advance to the next level of predictive/proactive condition performance-based \((\Delta R^2)\) maintenance planning and the integration of the two models.

Table 5 compares the optimised condition performance-based tamping requirements and the baseline routine requirements, and also gives an indication of the effectiveness of the former method in terms of estimated savings potential.
It is apparent from the information in Table 5 that the combination of condition performance-based maintenance strategies and a sound maintenance management process results in maintenance effectiveness and a reduction in maintenance costs. The information presented suggests a reduction in maintenance requirements, allowing an increase in track availability. This could be anticipated as less time is required to maintain the track, resulting in more time available for operation and thus an increase in production capability.

CONCLUSIONS

A maintenance management process model was developed to assist with effective decision-making on maintenance requirements. The maintenance management model is referred to as the Continuous Improvement process. This process (a systematic approach) ensures that constraints and changes are continuously taken into account, assisting in the optimisation of the railway’s asset management process. It enables an organisation to continuously improve the current maintenance strategies, intervention limits and maintenance processes.

The maintenance management process consists of numerous procedures based on condition analysis of the assets. Two tamping analysis models, namely the condition-based and condition performance-based models, were introduced as part of this procedure and resulted in methods that add value in determining preventive and proactive maintenance requirements. The authors suggest using the two condition analysis models in phases, depending on the maturity level of the railway’s asset management in using condition-based maintenance planning practices.

The results of the integration of the condition-based and condition performance-based maintenance methods are increased maintenance effectiveness (doing the right things) and consequently a decrease in maintenance costs. From this follows less maintenance time, increased availability of assets, value adding to the business through increased available capacity and, ultimately, an increase in production capability. This corroborates the asset management optimisation philosophy of Mitchell et al. (2007) – maximising return on investment.

In summary, the analysis indicates that use of the condition analysis models will increase the effectiveness of the maintenance management process. By implementing a condition-based and condition performance-based maintenance strategy, the railway organisation can achieve a total maintenance cost reduction of approximately 20% and an increase in asset availability of approximately 21%. In this study, the cost reduction for the study area analysed was restricted to 20% due to the significant length of curves present in the railway line. The authors are of the opinion that by implementing condition-based and condition performance-based maintenance, railway organisations could save up to 35% on maintenance costs while increasing capacity by approximately the same percentage.

ACKNOWLEDGEMENTS

The following people and organisations are gratefully acknowledged for their contributions in supporting this research:

- Dr Willem Ebersöhn for co-supervising the research and initiating the philosophies behind some of the work presented in this paper.
- Mr Danie Barnard and Mr Pierre Lombard for their encouragement and assistance in completing the study.
- Transnet Freight Rail, e-Logics, Esteq Engineering and the National Research Foundation (NRF) of South Africa for sponsoring this research as part of a THRIP project.

<table>
<thead>
<tr>
<th>Tamping activity type</th>
<th>Routine-based distance requirements</th>
<th>Condition performance-based (ΔR²) distance requirements</th>
<th>Savings (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unit</td>
<td>Quantity</td>
<td>Equivalent tamping distance (km)</td>
</tr>
<tr>
<td>Tangent track</td>
<td>km</td>
<td>42.0</td>
<td>42.0 (km)</td>
</tr>
<tr>
<td>Curved track</td>
<td>km</td>
<td>55.5</td>
<td>111.0 (km)</td>
</tr>
<tr>
<td>Switches</td>
<td>Number</td>
<td>81</td>
<td>32.4 (km)</td>
</tr>
<tr>
<td>Total distance</td>
<td></td>
<td>185.4</td>
<td></td>
</tr>
</tbody>
</table>

Figure 13 Condition-based (R²) and condition performance-based (ΔR²) comparison
REFERENCES

Dunn, S 1997. Implementing a computerized maintenance management system. Why most CMMS implementations fail to provide the promised benefits. Proceedings, Maintenance in Mining Conference, Sydney, Australia, July.


Laser-based assessment of road aggregate particle shape and texture properties with the aim of deriving comparative models

I J Breytenbach, J K Anochie-Boateng, P Paige-Green, J L van Rooy

Research was undertaken using an innovative three-dimensional (3D) laser scanning tool to study the shape and texture characteristics of road aggregate particles. Aggregate materials used for road construction, including G1 crushed rocks of different geological origins, recycled aggregate and alluvial gravel (not used as aggregate) were used for this study. Representative samples were scanned using the laser system to collect 3D aggregate data for analyses and, subsequently, develop comparative models. The objective was to arrange the aggregate particles in a sequence based on their surface texture. Two models were proposed and key aspects evaluated against each other. Ultimately, one model was selected that may be improved and used for further research. The study found that, while it is possible to use the 3D aggregate scan data to produce comparative models, distinguishing between particle shape and texture proved a daunting task. It was also concluded that particle elongation must be considered as a major influencing factor.

INTRODUCTION

The study forms part of a PhD research project (Breytenbach 2013), which is aimed partially at deriving a tabulated range of particles that could graphically or numerically represent a range of aggregate shape and texture properties. The proposed model, if validated, should be able to arrange particles in a certain sequence based on their texture properties. The method would lend scientific substantiation to particle shape characterisation instead of using visual observations, which have previously (to a larger extent) been used to quantify aggregate particle shape and texture.

An innovative study using three-dimensional (3D) laser scanning technology in the application of road materials is under way at the CSIR. The laser scanning method is applied to road construction aggregate using different approaches, attempting to study aggregate particle properties in great detail and for different applications. Emphasis has thus far been placed on particle shape and surface properties (Anochie-Boateng et al 2010) and the Flakiness Index (Anochie-Boateng et al 2011a), among others.

This paper describes the application of the laser scanning method in an attempt to refine the description of particle textures and shapes. The two comparative models proposed endeavour to sort aggregate particle data obtained from coarse/angular to smooth/rounded particles.

EXISTING METHODS

A number of advanced methods have been applied to study the particle shape properties of aggregates. Rao et al (2002) attempted to quantify the angularity of particles and ultimately showed that their approach (using image analysis) could distinguish between rounded gravel and crushed stone. Fletcher et al (2002) discussed an aggregate imaging system (AIMS) using back-lighting. It is clear that this method is able to determine the angularity of coarse and fine aggregate, as well as relative dimensions, thereby enabling the identification of flat or elongated particles. Other back-lit systems have also been used (Fernlund 2005; Ken et al 2009), while Descantes et al (2006) and Bouquet et al (2006) used videographers in combination with a back-lit system.

Kim et al (2003) analysed materials using a laser-based approach. The digital image technology (DIT) calculates the volume of a particle, as well as the minimum square aperture through which a particle could fit. However, the approaches discussed above are based on imaging techniques, which could at best capture two-dimensional (2D) physical attributes of the aggregate particles. In reality,
aggregate particles are three-dimensional; accordingly, any improved techniques or methods should be able to address the physical properties in three dimensions.

The approach used in this research is based on 3D scanning techniques that were recently introduced by the CSIR (Anochie-Boateng et al 2010; Anochie-Boateng et al 2011a; Anochie-Boateng et al 2011b). Details will not be repeated in this paper.

METHODOLOGY

Scanning procedure

Scanning of aggregate particles is done according to a protocol developed specifically for the laser scanner at the CSIR (Anochie-Boateng & Komba 2010). The approach adopted consists of four steps which can be summarised as follows:

1. Scanning: The first step entails physical processing of the particle by means of scanning. During the first phase, four faces of the particle are scanned (the particle is placed on a rotating table in the scanner). The particle is then rotated in order to scan the remaining two faces which were not in line of sight during the first scan phase (i.e. top and bottom).

2. Alignment: During this step the data obtained from the two scanning phases are orientated by the user. In essence, this step simply rotates and moves the two scans (dataset) into the correct orientation relative to each other.

3. Combining: Software is used to combine the correctly orientated data.

4. Merging: The combined data are merged into a single object and saved as a single particle presented in three dimensions. Any refinement (e.g. removal of small imperfections or overlapping areas, etc) is also executed during this step.

The scanning density used by the scanning equipment can be varied according to the user’s requirements. The density used affects the resolution of the final result and also the amount of time required to scan the particle. A higher scan resolution results in a longer amount of time required to scan the particle.

The resolution of the final result and also the amount of time required to scan the particle is also affected by the size of a particle.

Particle selection

The aggregate particles and materials used during this study are part of other ongoing research projects running congruently at the CSIR. The sourced aggregates used in this study included the following geological materials:

- Quartzite (G1 aggregate from stockpile) – Magaliesberg Formation, Pretoria Group, collected in Pretoria, Gauteng
- Granite (G1 aggregate from stockpile) – Johannesburg Dome, collected in Midrand, Gauteng
- Tillite (G1 aggregate from stockpile) – Dwyka Group, Karoo Supergroup, collected in Verulam, KwaZulu-Natal
- Hornfels (G1 aggregate from stockpile) – Tygerberg Formation, Malmsbury Group, collected in Durbanville, Western Cape
- Dolerite (G1 aggregate from stockpile) – Karoo Supergroup, collected in Trichardt, Mpumalanga
- Recycled aggregate – National Asphalt plant in Durban, KwaZulu-Natal
- Alluvial gravel – Quaternary aged surface deposit sampled from the river bed of the Molopo River, some 120 km west of Mafikeng, North West Province; this material is not an aggregate source.

Table 1 Summary of scanned particles

<table>
<thead>
<tr>
<th>Size</th>
<th>Quartzite</th>
<th>Granite</th>
<th>Tillite</th>
<th>Hornfels</th>
<th>Recycled aggregate</th>
<th>Alluvial gravel</th>
<th>Dolerite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.5 mm</td>
<td>–</td>
<td>22</td>
<td>30</td>
<td>18</td>
<td>–</td>
<td>30</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>19.0 mm</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>13.2 mm</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>9.5 mm</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>6.7 mm</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>4.75 mm</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>29</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>172</td>
<td>180</td>
<td>168</td>
<td>150</td>
<td>150</td>
<td>179</td>
<td>1 149</td>
</tr>
</tbody>
</table>

DISCUSSION

Two working concept models were developed using the scan data, one of which was ultimately abandoned as the model was heavily affected by factoring in the so-called “elongation value” (the ratio between the minimum and maximum dimensions of a particle). The second model proved to be superior to the first in that it discerned between differences in particle textures, as opposed to being excessively affected by the particle shape. The second model was further refined to develop a reference system for each particle size to enable comparison of particle textures.

Data manipulation

The data arrangement described above was done specifically with the aim of calculating another parameter to be used in analysis. In spreadsheet form the parameter was labelled “elongation value”. The Flakiness Index was also executed during this step.

In order to make provision for the effects of elongated or flattened particle shapes, data were divided into subgroups to assess whether the elongation/flattened would affect the derived model(s). It was therefore necessary to derive a descriptor for this assessment. Particle data were divided into...
“regular” and “elongated” subsets, based on the ratio between the maximum and minimum dimensions of a particle. An elongation value was calculated as follows:

\[ EV = \frac{\text{Length}}{\text{Depth}} \]  

where 

- \( EV \) = elongation value
- \( \text{Length} \) = the particle’s maximum dimension
- \( \text{Depth} \) = the particle’s minimum dimension

Particles with an \( EV > 2 \) were considered to be “elongated”, while particles with an \( EV < 2 \) were considered to be “regular”. The spreadsheet data were sorted based on the calculated elongation values and subdivided accordingly. Once data sorting had been completed, limited descriptive statistics and histogram analyses were done to verify data properties and identify potential shortages in data ranges.

Data were processed for all particle sizes and subsequently two comparative models (one for elongated particles and one for regular particles) were produced for each size fraction.

**Proposed Model One**

The first model considers three parameters, namely particle volume, particle surface area and the calculated elongation value. The “model value” is calculated as follows:

\[ \text{Model value} = \frac{V}{A} \times EV \]  

where 

- \( V \) = particle volume (\( \text{mm}^3 \))
- \( A \) = particle surface area (\( \text{mm}^2 \))
- \( EV \) = elongation value

Histogram analyses for the regular and elongated datasets are illustrated in Figure 1 and Figure 2 respectively. From the histogram of the elongated particles (Figure 2) it is clear that there are no data for the model value range between 6.0 and 7.0, which produced gaps in the output chart for the model. This underlines the effects of factoring in the elongation value and the limitations it imposes on the derived models.

After revision of the histogram data, model values were sorted in ascending order and model value ranges were calculated. Ten random values within the data range were selected and the closest corresponding particles were entered into a table. Table 2 shows the models selected for elongated particles and Table 3 shows the models selected for regular particles.

The tabled results reflect the findings of the histogram analyses in which gaps were found in the data range proposed for the
Table 3 Regular 6.7 mm particles – Model One

<table>
<thead>
<tr>
<th>Sample/Particle number</th>
<th>Minimum V/A</th>
<th>Maximum V/A</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tillite 10</td>
<td>2.322</td>
<td>2.541</td>
<td>0.219</td>
</tr>
<tr>
<td>Granite 9</td>
<td>1.584</td>
<td>1.965</td>
<td>0.381</td>
</tr>
<tr>
<td>Hornfels 9</td>
<td>1.782</td>
<td>2.154</td>
<td>0.372</td>
</tr>
<tr>
<td>Tillite 1</td>
<td>2.322</td>
<td>2.541</td>
<td>0.219</td>
</tr>
<tr>
<td>Granite 23</td>
<td>1.584</td>
<td>1.965</td>
<td>0.381</td>
</tr>
<tr>
<td>Gravel 23</td>
<td>1.782</td>
<td>2.154</td>
<td>0.372</td>
</tr>
<tr>
<td>Gravel 28</td>
<td>2.322</td>
<td>2.541</td>
<td>0.219</td>
</tr>
<tr>
<td>Quartzite 18</td>
<td>1.584</td>
<td>1.965</td>
<td>0.381</td>
</tr>
</tbody>
</table>

Table 4 Particle selection for 6.7 mm – Model Two

<table>
<thead>
<tr>
<th>Regular particles</th>
<th>Minimum V/A</th>
<th>Maximum V/A</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.910</td>
<td>1.580</td>
<td>0.670</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Increments calculated</th>
<th>Increment value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.074</td>
<td>Nearest match</td>
</tr>
<tr>
<td></td>
<td>Sample/Particle number</td>
</tr>
<tr>
<td>0.910</td>
<td>0.910</td>
</tr>
<tr>
<td>0.985</td>
<td>0.981</td>
</tr>
<tr>
<td>1.059</td>
<td>1.062</td>
</tr>
<tr>
<td>1.134</td>
<td>1.139</td>
</tr>
<tr>
<td>1.208</td>
<td>1.208</td>
</tr>
<tr>
<td>1.282</td>
<td>1.285</td>
</tr>
<tr>
<td>1.357</td>
<td>1.354</td>
</tr>
<tr>
<td>1.431</td>
<td>1.439</td>
</tr>
<tr>
<td>1.506</td>
<td>1.513</td>
</tr>
<tr>
<td>1.580</td>
<td>1.580</td>
</tr>
</tbody>
</table>

ELONGATED PARTICLES

<table>
<thead>
<tr>
<th>Minimum V/A</th>
<th>Maximum V/A</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.675</td>
<td>1.595</td>
<td>0.920</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Increments calculated</th>
<th>Increment value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.102</td>
<td>Nearest match</td>
</tr>
<tr>
<td></td>
<td>Sample/Particle number</td>
</tr>
<tr>
<td>0.675</td>
<td>0.675</td>
</tr>
<tr>
<td>0.778</td>
<td>0.774</td>
</tr>
<tr>
<td>0.880</td>
<td>0.873</td>
</tr>
<tr>
<td>0.982</td>
<td>0.981</td>
</tr>
<tr>
<td>1.084</td>
<td>1.084</td>
</tr>
<tr>
<td>1.186</td>
<td>1.182</td>
</tr>
<tr>
<td>1.289</td>
<td>1.284</td>
</tr>
<tr>
<td>1.391</td>
<td>1.379</td>
</tr>
<tr>
<td>1.493</td>
<td>1.490</td>
</tr>
<tr>
<td>1.595</td>
<td>1.595</td>
</tr>
</tbody>
</table>
elongated particles. Factoring the elongation value into the model also introduced a bias towards elongated particles in both datasets. Subsequently, the results show little refinement in data, particularly for the elongated particles. Considering that the aim of the model was to sort the particle data from coarse/angular to smooth/rounded particles, it is clear that this model was only moderately successful for regular particles and largely unsuccessful for elongated particles.

Proposed Model Two

The results from the first model led to the second model, in which the elongation value is no longer factored into the model value:

\[ V/A \](3)

The subsequent histogram analyses show an improvement in model data presentation. Although the histogram for regular particles (Figure 3) shows little improvement over the same dataset for Model One, the histogram for elongated particles (Figure 4) no longer has the previous gap in data for the second model.

The calculated model values were again sorted in ascending order, but this time 10 equal increments were identified between the minimum and maximum data values. The datasets were used to identify the nearest corresponding value and the applicable particles were again entered into a table. The results are summarised in Table 4, while Table 5 and Table 6 show the results for elongated and regular particles respectively.

In this instance the results showed significant refinement based on the data from the second model. The refinement is perhaps more clearly observed in the elongated particle results than in the regular particle results. Model Two shows good potential for arranging particles from rough/angular to smooth/rounded, although there is still room for improvement.

Model comparison

From a visual comparison of the results of the two proposed models it is clear that the second model refines and arranges data more effectively from rough or angular particles to smooth or rounded particles. The limited calculated model data in Model One may be due to factoring the elongation value into the calculations.

CONCLUSION

The research was undertaken to determine whether a three-dimensional laser-based scanning method could be applied to the modelling of the shape and texture of road aggregate particles in an attempt to arrange particles based on their surface texture. The laser scanning technique applied was relatively successful in producing comparative models that can be used to describe or classify aggregate particle shape and surface texture properties. The scan data obtained from scanning various aggregate materials were used to develop and compare two derived models.

The results indicate that the second proposed model shows better potential for describing the shape and surface texture properties of the aggregates, compared with the first model. Nevertheless, further development is required. It was found that factoring an elongation value (EV) into the proposed model (i.e. Model One) had a negative effect. The data with the elongation value factored into the model were heavily biased towards the inclusion of either elongated or flat particles. Generally, neither of the proposed models could effectively discern between particle roughness and angularity as the two parameters appear to be strongly co-dependent.

RECOMMENDATIONS

It is suggested that the approach used in Model Two be considered for further analyses
and investigations of particle properties. Additional scan data obtained from mixed geological sources will also assist in the further refinement of the models. Further studies may have to investigate how elongation properties can be combined with surface area or volume separately, as alternatives.

ACKNOWLEDGEMENTS

The work presented in this paper received funding from the Council for Scientific and Industrial Research (CSIR) Strategic Research Panel (SRP) project TA 2001-001. The authors would like to acknowledge the CSIR R&D office for funding this SRP project. Mr Julius Komba of the CSIR and the University of Pretoria, and Ms Sibongiseni Lusawana of the CSIR and Tshwane University of Technology in Pretoria assisted in the scanning of the aggregates used in this research.

REFERENCES


The text is a technical paper on incorporating rainfall uncertainty into catchment modelling. It discusses the importance of rainfall uncertainty in hydrological analysis and presents a framework for addressing this uncertainty. The paper is authored by J.G. Nduru and includes an introduction, methodology, results, and conclusions. The abstract, keywords, and introduction are provided, along with references to previous research on rainfall uncertainty. The paper uses Bayesian approaches to incorporate rainfall uncertainties into catchment modelling and demonstrates the effectiveness of this approach through a case study.
(Kavetski et al 2006a, b) that update parameter distributions within a single computation run. The framework is assessed by the split-sample method and a control experiment in which no disturbances (multipliers) are applied. The effect of rainfall uncertainties on the optimisation effort required in calibration is assessed by comparing the validation performance obtained from two levels in rigour of optimisation.

**METHODOLOGY**

Figure 1 presents the framework for incorporating rainfall uncertainty for the common streamflow simulation problem and could be easily adapted to other catchment modelling problems (water quality, groundwater, sediment generation, etc). The areal rainfall obtained by any appropriate method (e.g. Thiessen polygons) is perturbed by multipliers obtained randomly from a probability distribution derived from the rainfall data. An areal rainfall $r_t$ for period $t$ thus becomes $r_t \times m_t$, where $m_t$ is the multiplier for period $t$. The number of perturbed rainfall sequences that need to be generated (ensemble size) is selected and a population of perturbed rainfall sequences is thus obtained. Each of these is used, together with other required inputs, for multiple calibrations of the model. An understanding of the model structure, the catchment characteristics, previous experience and other information is used to establish the starting parameter ranges and the parameter range limits for the calibration. Where the uncertainties regarding the realistic parameter values are large, the starting ranges will be set more widely. The ranges therefore effectively act as quantifiers of parameter uncertainty and define the prior distribution of the parameters. Depending on the purpose of the modelling, an appropriate objective function is also selected for the calibration.

Each calibration run (for each perturbed rainfall sequence) provides an “optimal” parameter set and a population of optimal parameters is finally obtained. An assessment of this population and the calibrated streamflow time series makes it clear how realistic the modelling is and helps to identify any unexpected behaviour. This may then require adjustment of the parameter range limits and could also provide leads to aspects of significant catchment processes that were ignored or not recognised (Ndiritu 2009b). After the practically implementable changes have been made (and the calibration runs repeated if need be), each of the “optimal” parameter sets is used with a perturbed rainfall series (and other required inputs) for a period that was not applied to calibrate the model. The result is a population of validation streamflow time series. A comparison between the observed validation time series and the generated population of validation streamflows shows how suitable the framework is for the specific problem.

**APPLICATION OF FRAMEWORK**

The uncertainty framework was applied to daily streamflow modelling of the Mooi River catchment in South Africa using the Australian Water Balance Model (AWBM) and multiplicative perturbations (multipliers) of rainfall derived from ratios of areal rainfall obtained from various rain-gauge densities. The widely applied SCE-UA optimiser (Duan et al 1992) was selected for calibration and maximising the coefficient of efficiency as the objective function. The ensemble (population) size was subjectively selected as 100.

**The catchment**

The Mooi River headwaters up to river-gauging station V2H002 were included in the analysis and were delineated into three sub-catchments: up to gauging stations V2H005 and V2H007, and the incremental area from these two to V2H002. Figure 2 shows the location of the catchment in South Africa, the three sub-catchments and the four rain-gauging stations used to obtain areal rainfall. Daily evaporation measurements were obtained from station V7E003A located outside the catchment. Flow and evaporation data were obtained from the Department of Water Affairs’ (DWA) website (http://www.dwa.gov.za/hydrology), while rainfall was obtained from a rainfall database and extraction facility (Lynch 2003; Kunz 2009). The period 3 November 1973 to 19 August 1976 was used for calibration and that from 20 August 1976 to 7 June 1979 for validation. The selection was based on the need to
have a continuous dataset with minimal human impacts.

The catchment model
The AWBM model (Boughton 2004) is widely used for daily rainfall-runoff modelling in Australia and for flood hydrograph prediction when applied in hourly time steps. An approach for estimating runoff for ungauged catchments in Australia using the AWBM model has also been developed (Boughton & Chiew 2007). Makungo et al (2010) applied the AWBM to the Nzhelele catchment of Limpopo Province, South Africa. The AWBM was selected on the basis of its robust structure and successful application. The ACRU model (Schulze 1989) is widely applied for daily catchment modelling in South Africa, but is data-intensive and has not been set up for hybrid manual–automatic calibration. ACRU was therefore not an optimal choice for this study, although it is possible to adapt the rainfall uncertainty framework for application with ACRU. The AWBM model (Figure 3) assumes that the catchment consists of three stores of different depths C1, C2 and C3 which respectively occupy different proportions of the catchment, indicated as partial areas A1, A2 and A3 in Figure 3. At each time period, runoff is generated as the sum of the excess (overflow) from each store. The runoff is then divided into surface runoff and baseflow in proportions determined by the baseflow index (BFI). The surface runoff and the baseflow at the catchment outlet are each subjected to linear attenuation and are then summed to give the flow at the catchment outlet. Boughton (2004) provides more details of the AWBM model.

The model applied in this study also included lags for both surface runoff and baseflow, and a coefficient for scaling open-water evaporation to effective catchment evapotranspiration, giving a total of 12 parameters for each sub-catchment. The partial areas A1, A2 and A3 are expressed as proportions of the total area and therefore sum to unity. Only two of the three therefore need to be calibrated and 11 parameters were calibrated for each sub-catchment. These are shown in the first two columns of Table 1. Although the recession constants can be obtained directly from the data, it was decided to calibrate them, as an effective calibrator would have no difficulty obtaining these parameters for a well-structured model. Table 1 shows the starting parameter ranges and the range limits that were used in this study based on the understanding of the model structure, literature sources (Boughton 2004, Boughton & Chiew 2007) and past experience of modelling the Mooi River catchment.

### Probability distribution of multiplicative perturbations
Some studies have assumed that the multiplicative perturbations (multipliers) can be obtained from a log-normal distribution (Kavetski et al 2006b; Thyer et al 2009) and this has been largely supported by an experimental study (McMillan et al 2011), although the log-normal distribution did not capture the upper-end tail of the data.

---

![Image](image_url)
adequately. McMillan et al (2011) therefore proposed trials with other distributions as well. No other field data-based studies on multiplier distributions were found in the literature, and assuming that distributions fitting well in one region will do the same in another may also not be justifiable. For the current study, plausible values of multipliers were determined by computing ratios of daily areal rainfall values obtained from different rain-gauge densities for the study catchment. Thiessen polygons were used to obtain the areal rainfalls and this was confined to the days with non-zero rainfalls at all stations. It is expected that the multiplier values should depend on the magnitude of the areal rainfall as larger rainfall storms cover bigger areas and less variable rainfall would therefore be recorded at the different rain gauges. The observed variation of the multipliers with the areal rainfall (obtained at the highest rain-gauge density) is presented as Figure 4 and it reveals the expected reduction in multiplier variability as areal rainfall increases. Figure 4 also reveals that very large variations of areal rainfall could be obtained by simply omitting one or two rain gauges. It was decided to incorporate the observed reduction in multiplier variability in generating the perturbations by obtaining probability distributions for different ranges of areal rainfall magnitude. After some trial runs, the rainfall ranges selected were: < 3, 3–10, 10–20, 20–40 and > 40 mm. The multipliers within each range were ranked and plotted in order of magnitude, with the rank transformed into a percentage (non-exceedance probability), akin to the plotting of flow-duration curves. This resulted in the cumulative density plots presented in Figure 5. The multiplier to apply for a given areal rainfall was then randomly obtained from the respective probability distribution, based on the rainfall magnitude.

**Experimental set-up**

In order to evaluate the impact of incorporating rainfall perturbations, a control experiment consisting of 100 randomly initialised calibrations of the catchment with the unperturbed rainfall data was included. It was also decided to assess the effect of incorporating uncertainties on the required level of optimisation for calibration because it was considered likely that perturbing data could reduce the effectiveness and therefore the need for high levels of optimisation. The optimiser selected for this study, the SCE-UA (Duan et al 1992), is widely used and has been found to be effective and efficient (Ndiritu 2009a). The SCE-UA generates a population of solutions (parameter values) and divides these into a
number of complexes. Each complex evolves independently, using the downhill simplex method for a set number of evolutions. The complexes are then shuffled to exchange valuable information among them and a new set of independent evolutions (epoch) commences. This process repeats until the set convergence criteria are achieved. The default SCE-UA optimisation parameters as specified by Duan et al. (1994) were applied here and the level of optimisation was varied by setting the two parameters that Duan et al. (1994) did not specify, namely the number of complexes to use and the convergence criterion to apply. The higher optimisation level applied 10 complexes and the convergence criterion was specified as an improvement of less than 10% in the best solution (objective function value) of the current epoch in comparison with the best solution from the epoch two steps before (the one before the previous epoch). For the lower optimisation level, five complexes were applied and convergence was specified as an improvement of less than 10% in the best solution from the current epoch in comparison with the best one from the previous epoch.

A set of 100 calibration runs with and without perturbations was therefore carried out at the higher and the lower levels of optimisation. The analysis reported in the next section thus compares results from the following four experiments: (i) higher optimisation effort with perturbations; (ii) higher optimisation effort with no perturbations; (iii) lower optimisation effort with perturbations; and (iv) lower optimisation effort with no perturbations. The lower level took 110 minutes (on a standard desktop PC), while the higher level of optimisation took 11 hours (six times longer).

RESULTS AND DISCUSSION

Table 2 provides the mean and standard deviations of the 100 values obtained for the three sub-catchments. All the parameter values were found to be realistic. The mean parameter values from the four experiments are very close and mostly within 95–105% of the grand average (average parameter from all four experiments), as seen in Figure 6. Figure 6 also presents plots of the coefficients of variation (mean/standard deviation) of the parameters. It is observed that the coefficients of variation of only the evaporation coefficient (Ke) consistently increase with the inclusion of rainfall uncertainties. This happens for sub-catchment V2H005 and V2H007 but not for V2H002. The mean and coefficient of variation for the surface lag (LagS) is also notably higher for V2H002 than for the other sub-catchments and a

Table 2 Mean and standard deviation of parameters from 100 calibration runs for sub-catchments VH2005, V2H007 and V2H002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Units</th>
<th>Mean</th>
<th>Std dev</th>
<th>Mean</th>
<th>Std dev</th>
<th>Mean</th>
<th>Std dev</th>
<th>Mean</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sub-catchment V2H005</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>mm</td>
<td>9.57</td>
<td>1.58</td>
<td>9.75</td>
<td>1.54</td>
<td>10.0</td>
<td>1.61</td>
<td>9.68</td>
<td>1.79</td>
</tr>
<tr>
<td>C2</td>
<td>mm</td>
<td>76.8</td>
<td>15.1</td>
<td>78.05</td>
<td>13.17</td>
<td>82</td>
<td>17.38</td>
<td>80.6</td>
<td>16.16</td>
</tr>
<tr>
<td>C3</td>
<td>mm</td>
<td>299.5</td>
<td>5.39</td>
<td>299.7</td>
<td>5.37</td>
<td>300.1</td>
<td>6.15</td>
<td>301.4</td>
<td>5.49</td>
</tr>
<tr>
<td>A1</td>
<td>–</td>
<td>0.15</td>
<td>0.014</td>
<td>0.15</td>
<td>0.022</td>
<td>0.15</td>
<td>0.029</td>
<td>0.15</td>
<td>0.036</td>
</tr>
<tr>
<td>A2</td>
<td>–</td>
<td>0.49</td>
<td>0.027</td>
<td>0.5</td>
<td>0.02</td>
<td>0.48</td>
<td>0.026</td>
<td>0.49</td>
<td>0.025</td>
</tr>
<tr>
<td>BFI</td>
<td>–</td>
<td>0.62</td>
<td>0.043</td>
<td>0.69</td>
<td>0.049</td>
<td>0.6</td>
<td>0.052</td>
<td>0.59</td>
<td>0.05</td>
</tr>
<tr>
<td>Ks</td>
<td>–</td>
<td>0.61</td>
<td>0.043</td>
<td>0.61</td>
<td>0.042</td>
<td>0.61</td>
<td>0.041</td>
<td>0.59</td>
<td>0.042</td>
</tr>
<tr>
<td>Kb</td>
<td>–</td>
<td>0.93</td>
<td>0.022</td>
<td>0.92</td>
<td>0.018</td>
<td>0.93</td>
<td>0.025</td>
<td>0.923</td>
<td>0.024</td>
</tr>
<tr>
<td>Ke</td>
<td>–</td>
<td>0.49</td>
<td>0.058</td>
<td>0.45</td>
<td>0.025</td>
<td>0.5</td>
<td>0.065</td>
<td>0.45</td>
<td>0.033</td>
</tr>
<tr>
<td>LagS</td>
<td>days</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LagB</td>
<td>days</td>
<td>2</td>
<td>0</td>
<td>1.99</td>
<td>0.1</td>
<td>2</td>
<td>0</td>
<td>2.01</td>
<td>0.1</td>
</tr>
</tbody>
</table>

| **Sub-catchment V2H007** | | | | | | | | | |
| C1 | mm | 9.67 | 1.59 | 9.72 | 1.39 | 9.94 | 1.63 | 9.77 | 1.87 |
| C2 | mm | 100.6 | 16.1 | 100.5 | 15.48 | 99.9 | 17.55 | 97.8 | 14.96 |
| C3 | mm | 296.0 | 4.86 | 295.5 | 4.68 | 297.3 | 6.078 | 297.6 | 6.26 |
| A1 | – | 0.14 | 0.015 | 0.14 | 0.016 | 0.14 | 0.038 | 0.14 | 0.036 |
| A2 | – | 0.47 | 0.029 | 0.47 | 0.024 | 0.46 | 0.034 | 0.47 | 0.029 |
| BFI | – | 0.54 | 0.042 | 0.5 | 0.036 | 0.53 | 0.05 | 0.5 | 0.043 |
| Ks | – | 0.53 | 0.05 | 0.52 | 0.044 | 0.54 | 0.046 | 0.51 | 0.043 |
| Kb | – | 0.94 | 0.02 | 0.92 | 0.017 | 0.94 | 0.023 | 0.91 | 0.022 |
| Ke | – | 0.38 | 0.034 | 0.35 | 0.017 | 0.38 | 0.044 | 0.35 | 0.018 |
| LagS | days | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| LagB | days | 2 | 0 | 1.99 | 0.1 | 1.99 | 0.1 | 2.01 | 0.1 |

| **Incremental sub-catchment V2H002** | | | | | | | | | |
| C1 | mm | 9.92 | 1.58 | 9.98 | 1.65 | 10.23 | 1.90 | 10.07 | 1.71 |
| C2 | mm | 99.4 | 15.57 | 97.8 | 13.82 | 99.4 | 17.43 | 97.65 | 15.64 |
| C3 | mm | 299.3 | 5.85 | 299.3 | 5.33 | 299.7 | 6.31 | 301.1 | 5.96 |
| A1 | – | 0.13 | 0.017 | 0.13 | 0.016 | 0.13 | 0.016 | 0.13 | 0.016 |
| A2 | – | 0.46 | 0.028 | 0.45 | 0.025 | 0.45 | 0.029 | 0.46 | 0.029 |
| BFI | – | 0.48 | 0.045 | 0.48 | 0.046 | 0.49 | 0.061 | 0.49 | 0.059 |
| Ks | – | 0.54 | 0.05 | 0.53 | 0.048 | 0.52 | 0.061 | 0.52 | 0.054 |
| Kb | – | 0.86 | 0.027 | 0.87 | 0.024 | 0.87 | 0.025 | 0.87 | 0.028 |
| Ke | – | 0.4 | 0.073 | 0.4 | 0.066 | 0.41 | 0.089 | 0.4 | 0.067 |
| LagS | days | 1.17 | 0.378 | 1.22 | 0.416 | 1.19 | 0.394 | 1.11 | 0.316 |
| LagB | days | 2 | 0 | 1.99 | 0.1 | 2 | 0 | 2 | 0 |
Figure 6 Comparison of the averages and coefficients of variation of the 100 parameter values obtained from four experiments. H & U indicate higher optimisation with uncertainty, L & U lower optimisation with uncertainty, H higher optimisation without uncertainty and L lower optimisation without uncertainty.
probable explanation for these differences is offered later in this section. The effect of rainfall uncertainty on parameter Ke could be attributed to the direct impact of perturbations on rainfall on the computed net rainfall (rainfall – Ke × evaporation). The observed dependence of only one parameter on rainfall uncertainty is consistent with the finding by Kuczera et al. (2006) who found that only two out of the seven parameters of the LogSPM model were dependent on rainfall uncertainty.

Figure 7 shows the probability density plots and normal distribution fits for parameters Ke and A2 for sub-catchment V2H005. Although the differences in variability were not substantial for parameter A2, the plot in Figure 7 helps to illustrate the ability of the calibration to search for and obtain optimal parameters beyond the starting range specified in Table 1. This table specifies the starting range as 0.4–0.5 for A2, whereas a substantial proportion of the optimal parameters for A2 in Figure 7 locate beyond 0.5. From Figure 7 it is observed that applying perturbations leads to a notably larger spread in variability for parameter Ke at both optimisation levels, whereas the effect on the variability of A2 was only slight. Incorporating uncertainties shifted the location of the distribution of Ke, but the average Ke values for all four experiments were still reasonably close.

Figure 8 shows the 5–95 percentile range obtained from the 100 ensembles of validation time series for the four experiments for sub-catchment V2H005 and also includes plots of the observed streamflows for the same period (portrayed as circles). It is found that perturbing the rainfall obtains much wider ranges than if this is not done. A more detailed analysis of the effect of rainfall uncertainties is done by obtaining the percentages of the observed flows locating within the 5–95% bounds for different magnitudes of observed flows. The percentages obtained using 10 classes of flow magnitude defined by the 10th percentiles of the respective flow-duration curves are presented in Table 3 and Figure 9. For all three sub-catchments, including rainfall uncertainty obtains a much larger percentage of the flows within the 5–95% bounds for all flow levels, with an overall increase from 25 to 52%. The proportion of observed flows within the percentiles is found to reduce as flow reduces, probably because the applied objective function (maximising the coefficient of variation) favours the replication of higher rather than lower flows. It could also be an indication of an inadequacy of the AWBM model structure in simulating low flows. In addition, Table 3 and Figure 9 reveal that the lower optimisation effort obtains slightly higher percentages of observed flows within

![Figure 7 Normal probability distribution plots for parameters Ke and A2 for subcatchment V2H005](image)

<table>
<thead>
<tr>
<th>Sub-catchment</th>
<th>Experiment</th>
<th>Flow-duration percentile range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0–10</td>
</tr>
<tr>
<td>V2H005</td>
<td>H &amp; U</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>L &amp; U</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>38</td>
</tr>
<tr>
<td>V2H007</td>
<td>H &amp; U</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>L &amp; U</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>36</td>
</tr>
<tr>
<td>V2H002</td>
<td>H &amp; U</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>L &amp; U</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>33</td>
</tr>
</tbody>
</table>
Figure 8 Validation 5–95 percentile plots and observed flows for sub-catchment V2H005
the 5–95% bounds than the higher level for the entire range of flows. Careful selection of the optimisation effort to apply is therefore needed, as an exceedingly high optimisation may over-fit on the calibration dataset, while simultaneously losing the overall fitness of the parameter set.

A probable explanation of the distinct differences in the results obtained for sub-catchment V2H002 in comparison with those for V2H005 and V2H007 is now offered. For V2H002, the variability of parameter Ke is found to be independent of rainfall uncertainty (Figure 6), while the average value and the coefficient of variation of the lag for surface runoff (LagS) is found to be considerably higher than for V2H005 and V2H007 (Table 2). The observed average LagS value ranged from 1.11 to 1.22 days for V2H002, meaning that some calibration runs optimised this to 1 day and some to 2 days since LagS was specified to vary at a daily time step. For V2H005 and V2H007, Table 2 shows that the LagS value optimised to 1 day for all 100 runs. Sub-catchment V2H002 is the most downstream of the three sub-catchments and is expected to generally steep more gently than the other two; consequently it would have slower surface runoff processes. Since V2H002 is also the longest of the three sub-catchments, it is probable that a considerable portion of the surface runoff takes longer than 1 day to reach river gauge V2H002, but would reach it within 2 days, while most surface runoff may be reaching gauges V2H005 and V2H007 within 1 day. Since the calibration constrained LagS to optimise to a daily value, the variability in LagS became artificially larger as it has to take a value of either 1 or 2 days, whereas the more realistic lag time lies in-between. Confining LagS to a daily time could also have caused inaccuracy in the streamflow simulation that perhaps (i) confounded the impact of rainfall uncertainties on Ke, (ii) led to the observed higher variability of the other parameters for V2H002 than for V2H005 and V2H007 (coefficient of variation of 0.089 compared with 0.074), and (iii) led to the lower validation performance for V2H002 as seen in Table 3 and Figure 9. Catchment modelling is mostly carried out at single time steps but the reasoning here, while not proven, gives credence to variable time interval catchment modelling (Hughes & Sami 1994) which seems to have gone dormant in research and practice.

In comparison with the manual rainfall-runoff model calibration approach (the predominant approach in southern Africa) which obtains single parameter values fairly subjectively, the framework applied here obtains a population of realistic parameter sets, while incorporating areal rainfall uncertainty. As revealed in the previous

**Figure 9** Percentage of observed flows that locate within the 5–95 percentile bounds. H & U indicate higher optimisation with uncertainty, L & U lower optimisation with uncertainty, H higher optimisation without uncertainty, and L lower optimisation without uncertainty.
CONCLUSIONS AND RECOMMENDATIONS

A framework for incorporating rainfall uncertainties in catchment modelling has been presented and applied to a daily streamflow simulation problem of the Mooi River catchment in South Africa using the AWBM model. In the absence of any field data-based guideline for quantifying rainfall uncertainties, the ratios of areal daily rainfalls obtained from various rain-gauge densities were used to obtain probable values of multiplicative perturbations. A reasonable probability distribution of perturbations was then conceived from these and it was found that very large variations of areal rainfall can be obtained by omitting one or two rain gauges. This underlines the need to formally incorporate rainfall uncertainty into water resources assessment.

The impact of rainfall uncertainties was assessed by making 100 randomly initialised calibration-validation runs, with and without including rainfall uncertainties, and comparing the resulting distribution of parameter values and the proportion of observed flows falling in the 5–95 percentile bounds of the flows simulated in validation. Applying rainfall uncertainties is found not to impact on the average parameter values and to increase significantly the variability of only the evaporation coefficient Ke of the AWBM model – the only parameter directly associated with rainfall. All the other parameters are for modelling surface and subsurface processes, and the independence of the probability distributions of their calibrated values from rainfall uncertainty is considered to be an indication that the modelling represented the main catchment components and processes realistically. This also indicates that including rainfall uncertainty in calibration did not prevent a realistic quantification of parameter uncertainty, although the framework did not include an explicit procedure to enable this as is done in the more complex and computation-intensive Bayesian approaches (Kavetski et al 2006a, b). The framework applied here could therefore be a credible and practical alternative to these approaches, provided the modelling captures the main catchment processes adequately.

Applying rainfall uncertainties was found to double the proportion of observed flows within the 5–95 percentile bounds from an average of 25 to 52% in validation, indicating that rainfall input uncertainty is indeed highly significant. Two levels of optimisation effort were applied and the lower optimisation level obtained slightly better percentages of the observed flows within the 5–95 percentile bounds, highlighting the need for careful selection of the optimisation effort to apply in model calibration.

Further work needs to consider the following:

- Are multiplicative perturbations the most appropriate for quantifying areal rainfall uncertainties, and does the approach applied here make the best use of the data and other information available? Ongoing analysis indicates that linear perturbations hold much promise.
- How can computational efficiency be maximised/optimised for uncertainty analysis? The SCEM-UA (Vrugt et al 2003), a later development of the SCE-UA calibrator applied here, could be considered.
- How does the choice of the ensemble size and objective function for calibration impact on the uncertainty analysis?
- How would this framework fit into the current water resources planning and management decision-support structures?
- How can the framework be adapted for prediction in ungauged basins and to climate change/variability analysis?

REFERENCES


Kunz, R 2009. Rainfall data extraction, Version Number 1.2, ICFR, PMB, South Africa.


Rocking shear wall foundations in regions of moderate seismicity

J E van der Merwe, J A Wium

This paper presents a study which investigates the feasibility of a concept to reduce the size of shear wall foundations for earthquake forces in regions of moderate seismicity. The approach is to allow rocking of the shear wall foundation and to include the contribution of a shear wall and reinforced concrete frame to assist as a lateral force-resisting system. A simplified multi degree-of-freedom model with non-linear material properties was used to investigate this lateral-force-resisting system subjected to base accelerations from recorded ground motions. An example building was studied with the shear wall foundation designed to resist 0%, 20%, 40%, 60%, 80% and 100% of the design overturning moment from the seismic event. Non-linear time-history analyses were performed with input from seven scaled ground-motion records. It is shown that the concept warrants more detailed studies and that a significantly reduced shear wall foundation size is possible without failure of the lateral force-resisting system.

INTRODUCTION
Reinforced concrete structures consisting of flat slabs, columns and shear walls are common structural systems in many parts of the world. These structures are usually designed so that the shear walls resist all lateral forces, which can be either wind or seismic loads. The flat slabs and columns are designed to resist gravity loads only.

In regions of moderate seismicity it has been shown that a suitable structural system is created when designing the shear wall with a plastic hinge zone at the lower part of the wall, with the shear walls resisting lateral loads, and all other structural elements designed to resist gravity loads. It is common to verify the behaviour of the columns and flat slabs against lateral drift criteria.

For the assumption of plastic zones at the bottom of shear walls to hold true, a sufficiently stiff foundation is required. This foundation should have limited rotation and should remain linear elastic when lateral loads are applied to the structure. Buildings with at least one basement level may provide a shear wall with a sufficiently stiff foundation.

However, if a building has no basement level, the stiff support of the shear wall will have to be provided by the foundation. Shear walls that are designed to resist seismic loads require significantly larger foundations than in the case of wind loading as the dominant lateral load condition, depending on the height of the building and the number of shear walls in the building. Traditionally the shear wall foundation is designed to have a larger bending moment capacity than the shear wall to ensure that plastic deformation occurs in the wall and not the foundation. The result is that excessively large shear wall foundations are required even in regions of moderate seismicity.

This paper presents a study into the feasibility of reducing the size of shear wall foundations in regions of moderate seismicity in buildings with no basement level. A simplified approach was taken to determine the merit of a more sophisticated approach in a subsequent study. This investigative study was aimed at allowing shear wall foundation rocking, taking into account the contribution of structural frames consisting of flat slabs and columns to the lateral stiffness of the structure.

The shear wall, rocking shear wall foundation and the structural frame will therefore work together to resist seismic loading on the building, the main mechanism being the rocking motion of the shear wall foundation. Such a reduction in the shear wall foundation could result in a significant reduction in cost.

Analysis methods that are not normally used for building structures in regions of moderate seismicity were implemented in this study. It is the view of the authors that current simplified analysis methods are not capable of investigating the feasibility of the concept of this study.

LITERATURE REVIEW
The concept of rocking foundations has received much attention to date.
Gazetas (2006) states that the deformability of soil increases the natural vibration period of the structure, which in turn leads to smaller accelerations and stresses in the superstructure and foundation. The use of overstrength factors in the capacity design method may prevent structural yielding of the footing as well as bearing capacity failures. A limited amount of foundation uplift can, however, still occur. When foundation rocking takes place, other structural elements should be designed for the associated shedding of load from the shear wall to the structural frame.

Kawashima and Hosoiri (2003) have shown that foundation rocking has a beneficial effect on the dynamic performance of bridge piers. They found that the plastic deformation of the bridge pier decreases if uplifting of the foundation occurs as a result of softening of the moment-rotation hysteresis loops of the foundation.

The similarities and differences between the oscillatory response of a single-degree-of-freedom (SDOF) oscillator (regular pendulum) and the rocking response of a slender rigid block (inverted pendulum) were investigated by Makris and Konstantinidis (2001). They found that there are fundamental differences in the mechanical structure of these two dynamic systems, and consequently that the rocking structure cannot be replaced by an equivalent SDOF oscillator. Based on the findings by Makris and Konstantinidis (2001), it was decided that the simplified model created for the purpose of this investigation should contain all the horizontal degrees of freedom of the investigated structure. Therefore an equivalent multiple-degree-of-freedom (MDOF) model was created rather than an SDOF model. The MDOF model is discussed later in this paper.

Anderson (2003) investigated the effect of a rocking shear wall foundation to determine how this approach can be used to reduce shear wall foundation sizes – the main response investigated was the drift ratio of the structure. The shear wall and foundation, not including any other structural elements, were modelled using a number of soil springs with zero tension gap elements to allow for rocking of the shear wall foundation. The investigation by Anderson (2003) confirmed that the concept of rocking foundations can reduce the foundations considerably to sizes smaller than the size required to resist the moment capacity of the shear wall without the building falling over.

This concept was extended to include the contribution of other structural elements to the investigative study described in this paper.

Anderson (2003) investigated the effect of a rocking shear wall foundation to determine how this approach can be used to reduce shear wall foundation sizes – the main response investigated was the drift ratio of the structure. The shear wall and foundation, not including any other structural elements, were modelled using a number of soil springs with zero tension gap elements to allow for rocking of the shear wall foundation. The investigation by Anderson (2003) confirmed that the concept of rocking foundations can reduce the foundations considerably to sizes smaller than the size required to resist the moment capacity of the shear wall without the building falling over.

This concept was extended to include the contribution of other structural elements to the investigative study described in this paper.

### DESCRIPTION OF STRUCTURE

In this study an example building was used to perform various comparative analyses. The structural elements in the building were detailed according to the assumption that only the shear walls resist lateral forces, and that columns and flat slabs are designed to resist gravity loads only. Seismic excitation was only considered in the north-south direction and therefore shear walls were included to provide lateral stiffness in this direction.

A plan layout of the chosen building is shown in Figure 1.

The properties used in this investigation are shown in Table 1.

Calculation of the above parameters is discussed in more detail by Van der Merwe (2009).

The shear wall located on grid A of the plan layout, the foundation of this shear wall, and an internal structural frame located on grid B of the plan layout were investigated.

---

**Figure 1 Plan layout of chosen building**

**Table 1 Properties used in the investigation**

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slab thickness</td>
<td>250 mm</td>
</tr>
<tr>
<td>Storeys</td>
<td>Eight at 3.5 m floor-to-floor height</td>
</tr>
<tr>
<td>Column grid spacing</td>
<td>6 m</td>
</tr>
<tr>
<td>Column dimensions</td>
<td>600 mm × 600 mm</td>
</tr>
<tr>
<td>Wall dimensions</td>
<td>6 000 mm × 300 mm</td>
</tr>
<tr>
<td>Soil-bearing capacity</td>
<td>750 kPa</td>
</tr>
<tr>
<td>Flexural slab reinforcement</td>
<td>Bottom: $A_s = 447 \text{ mm}^2/m$ Top: Central column strip $A_s = 1 340 \text{ mm}^2/m$ Remainder column strip $A_s = 670 \text{ mm}^2/m$</td>
</tr>
<tr>
<td>Column reinforcement</td>
<td>Ground to 2nd floor: 4 474 mm$^2$ (K = 1.049) 2nd floor to roof: 2 767 mm$^2$ (K = 1.058)</td>
</tr>
<tr>
<td>Wall reinforcement</td>
<td>Ground to 4th floor: End zones = 2 513 mm$^2$ (K = 1.1) Remainder = 1 608 mm$^2/m$ 4th floor to roof: 908 mm$^2/m$</td>
</tr>
<tr>
<td>Structural wall floor loads for seismic force calculation</td>
<td>31 050 kg for each floor</td>
</tr>
</tbody>
</table>

K = Confinement factor (Paulay & Priestley 1992)
(see Figure 1). The stiff shear wall governs the dynamic response of the entire edge frame, and for this reason the contribution of the frame elements in the plane of the wall was deemed to be insignificant. The reinforcement layout for the shear wall resulted from flexural resistance to lateral loading, leading to an axial load of only 10% of the axial load capacity.

Foundation sizes were obtained by using different percentages of the overstrength bending moment \( (M_{O}d^+) \) as the applied bending moment for the design of the foundation. Foundation sizes obtained by applying 0%, 20%, 40%, 60%, 80% and 100% of the overstrength bending moment were investigated. Table 2 shows the investigated shear wall foundation sizes.

Table 2 Shear wall foundation dimensions

<table>
<thead>
<tr>
<th>Percentage of ( M_{O}d^+ ) applied</th>
<th>Length (m)</th>
<th>Width (m)</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>14.0</td>
<td>4.0</td>
<td>1.3</td>
</tr>
<tr>
<td>80</td>
<td>12.0</td>
<td>4.0</td>
<td>1.3</td>
</tr>
<tr>
<td>60</td>
<td>11.5</td>
<td>3.0</td>
<td>1.3</td>
</tr>
<tr>
<td>40</td>
<td>9.5</td>
<td>3.0</td>
<td>1.3</td>
</tr>
<tr>
<td>20</td>
<td>8.0</td>
<td>2.5</td>
<td>1.3</td>
</tr>
<tr>
<td>0</td>
<td>6.5</td>
<td>2.0</td>
<td>0.8</td>
</tr>
</tbody>
</table>

For the purpose of the numerical analyses, criteria were determined that identified different modes of failure. The mean material properties were used to determine the failure criteria rather than the design material properties.

Material strain criteria

The strain limits of reinforcement steel and concrete material were calculated for different modes of failure and are presented in Table 3 and Table 4. Refer also to Figure 2 for definitions of the symbols. Calculation of these parameters is discussed in more detail by Van der Merwe (2009).

The mean reinforcement steel material properties were determined from the South African concrete design code SABS 0100-1 (2000) and Mirza and MacGregor (1979). The fracture strain of reinforcement steel was determined from FEMA 273 (1997). Expressions presented by Paulay and Priestley (1992) were used to determine the strain limits of concrete material. The structural analysis software package SeismoStruct (SEISMOSOFT 2007) was used in this investigation. Structural sections were created and the above-mentioned material strain limits were assigned to the different materials as performance criteria.

This enabled the identification of different types of failure as shown in Figure 4 and Figure 6. The properties of the reinforcement material used in this investigation are shown in Table 3.

For unconfined concrete the following values were used:

- \( K = 1.001 \)
- \( \varepsilon_c = 0.002 \)

Element rotation criteria

FEMA 273 (1997) prescribes limits to the chord rotation of plastic hinges that may form in different structural elements depending on the reinforcement detail and desired performance level. The assumption of a Life Safe performance level for the chosen building structure leads to the following plastic hinge rotation limits:

- Shear wall: 0.01 radians
- Columns: 0.005 radians

CAPACITY CURVES

Non-linear capacity against lateral loading effects was determined for the three systems assumed to contribute to the lateral stiffness of the building:

- Shear wall
- Structural frame
- Shear wall foundation

The capacity curves of these systems are discussed in the following paragraphs.

Table 3 Properties of reinforcement steel

<table>
<thead>
<tr>
<th>Material property</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modulus of elasticity (GPa)</td>
<td>( E_s )</td>
<td>200.0</td>
</tr>
<tr>
<td>Yield strength (MPa)</td>
<td>( f_y )</td>
<td>569</td>
</tr>
<tr>
<td>Strain-hardening parameter (-)</td>
<td>( M )</td>
<td>0.005</td>
</tr>
<tr>
<td>Specific weight (kN/m²)</td>
<td>( \gamma_s )</td>
<td>78.0</td>
</tr>
<tr>
<td>Fracture strain (m/m)</td>
<td>( \varepsilon_{su} )</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 4 Properties of confined concrete (refer to Figure 2)

<table>
<thead>
<tr>
<th>Material property</th>
<th>Symbol</th>
<th>Ground to 2nd floor columns</th>
<th>2nd floor to roof columns</th>
<th>Shear wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compressive strength (MPa)</td>
<td>( f'_{cc} )</td>
<td>33.0</td>
<td>33.0</td>
<td>33.0</td>
</tr>
<tr>
<td>Strain at peak stress (m/m)</td>
<td>( \varepsilon_{cc} )</td>
<td>0.0025</td>
<td>0.0026</td>
<td>0.0030</td>
</tr>
<tr>
<td>Confinement factor (-)</td>
<td>( K )</td>
<td>1.049</td>
<td>1.058</td>
<td>1.100</td>
</tr>
<tr>
<td>Specific weight (kN/m³)</td>
<td>( \gamma_c )</td>
<td>24.0</td>
<td>24.0</td>
<td>24.0</td>
</tr>
<tr>
<td>Collapse strain (m/m)</td>
<td>( \varepsilon_{cu} )</td>
<td>0.0041</td>
<td>0.0041</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

Figure 2 Stress-strain behaviour of confined and unconfined concrete
Shear wall capacity
Pushover analyses were performed on the shear wall to obtain the lateral force-displacement characteristics. The structural analysis software package SeismoStruct (SEISMOSOFT 2007) was used for this purpose. Lumped masses were included on each floor level to account for the effect of gravity of the contributing slab area on the shear wall. The designers' guide to EN 1998-1 and EN 1998-5 (2005) states that the base shear force should be distributed in a uniform and a triangular pattern over the height of the building and that the most unfavourable resulting curve should be used.

The result of the pushover analysis of the shear wall is a non-linear relationship of base shear force and lateral roof displacement. From Figure 3 it can be seen that the triangular force distribution results in the largest lateral roof displacement for a given base shear force and is therefore the most unfavourable result if lateral roof displacement is considered to be the determining parameter.

A bilinear approximation of the resulting curve was obtained by following the procedure described in FEMA 440 (2005). The prescribed material strain limits for steel yielding and cover spalling were highlighted during the pushover analysis as can be seen from Figure 4.

Structural frame capacity
Following the clauses for laterally loaded frames in the South African concrete design code SABS 0100-1 (2000), the slab width was taken as half the distance between the centres of the panels, resulting in an effective slab width of 3 m.

Both a uniform and triangular lateral load distribution was investigated in the pushover analysis using SeismoStruct (SEISMOSOFT 2007), and bilinear approximation curves were obtained from the FEMA 440 (2005) Procedure. Comparing the capacity curves (Figure 5), it is clear that a triangular distribution of the base shear force leads to the most unfavourable result if roof displacement is the parameter that is considered.

Prescribed material strain limits for steel yielding and concrete cover spalling were highlighted during the pushover analyses, as well as the slab element rotation limit (Figure 6).

Foundation capacity
Analytical expressions for the moment-rotation response of a rigid foundation on a Winkler soil model were presented by Allotey and Naggar (2003). Equations were derived for four main states. These conditions and the resulting supporting soil pressure on the foundation are shown in Figure 7.

An allowable soil-bearing pressure of 750 kPa and a soil density of 1 800 kg/m³ were chosen for the comparative analyses. For the building investigated using the chosen parameters, it was found that foundation uplift would occur prior to yielding of the
supporting soil for all investigated foundation sizes.

The non-linear moment-rotation capacity curves of the investigated foundation sizes are shown in Figure 8, along with the positions of the foundation uplift and yielding of the supporting soil.

**TIME-HISTORY DATA**

Non-linear time-history analyses were performed using recorded ground motions. It is important that the chosen ground motion records are representative of the geological and seismological conditions at the location of the investigated structure. Two approaches can be followed. Either three spectrum-compatible records are used with the design response taken to be the maximum of the three, or seven ground motions can be used with the design response taken to be the average of the seven (Priestley, Calvi and Kowalski 2007). In this study the latter approach was used.

**Selection of appropriate ground motion records**

An initial set of 20 ground motion records were chosen from the PEER Strong Motion Database (University of California 2001). These ground motions are representative of the geological conditions of the chosen structure. These geological conditions included firm soil to soft rock soil profiles and a 15.1 km to 31.7 km closest distance to rupture.

Seven ground motion records, for which the 5% damped response spectra within the same range of vibration periods best fit the shape of the elastic response spectrum used to analyse the investigated structure, were then selected from the initial set of 20 records.

The seven selected ground motion records used, together with the resulting scaling factors, are shown in Table 5.

**Scaling of ground motion records**

The response spectra of the above ground motion records were scaled to fit the elastic response spectrum in the expected range of vibration periods. For the purpose of scaling the chosen ground motion records, it was assumed that the shear walls have a dominant effect on the vibrating response of the building. This was justified from a comparison of the first mode of vibration of the entire building with that of the shear wall, which showed only a 2.15% difference. Natural vibration periods of the shear wall were therefore used to determine the period range in which to scale the earthquake response spectra. In order to determine the...
vibration periods an eigenvalue analysis was performed on the shear wall using SeismoStruct (SEISMOSOFT 2007) together with the appropriate contributing mass.

To account for the higher vibration mode response, a vibration period value that will result in a cumulative effective modal mass percentage of 90% was chosen as the lower limit for the range of vibration period values in which to scale the earthquake response spectra. The fundamental period of the wall was increased by a factor of 1.5 to account for the increase in vibration period that could result from non-linear material behaviour. This factor of 1.5 was also used by Naeim et al (2004) in presenting a procedure for scaling of ground motion time-histories. From the output of the eigenvalue analysis it was observed that the first three natural vibration modes lead to a cumulative modal mass percentage of 91.8%, with $T_1 = 1.264$ s and $T_3 = 0.071$ s. The second fundamental period ($T_2$) lies between these two values and is therefore of no importance with regard to the scaling of the selected ground motion. The earthquake response spectra were therefore scaled to fit the elastic response spectrum within a period range of 0.071 s to 1.896 s. A peak ground acceleration of 0.15 g applies to the Cape Town region and was used to compute the elastic response spectrum.

It was attempted to obtain a scaling factor that ensures an equal area between the elastic response spectrum and the earthquake response spectra above and below the elastic response spectrum. An attempt was also made to obtain a good fit between the curves at the fundamental period of vibration. Figure 9 shows this graphic procedure for the ground motion recorded at Plaster City during the Imperial Valley earthquake of 1979.

**SIMPLIFIED MODEL**

Simplified geometries were chosen to represent the various structural systems assumed to contribute to the lateral stiffness of the building to reduce computational effort when performing non-linear time-history analyses. The finite element analysis software system Strand7 (2005) was used to perform the non-linear time-history analyses. This software package allows the modelling of non-linear spring elements, various types of link elements, beam elements with non-linear material behaviour and lumped mass elements, all of which were used for the simplified model.

**Foundation**

A spring element with non-linear rotational stiffness was used to model the shear wall foundation. This enabled the direct use of the moment-rotation response obtained for each of the investigated foundation sizes to define the rotational stiffness of the spring element. It must be noted that the simplified model implies that the centre of rotation would always be about the centre line of the wall. Wall rotation due to true foundation
rocking response would, however, not be about a fixed position. Rather, the position of rotation would vary as foundation uplift occurs. This limitation was accepted due to the investigative nature of the study.

An elastic-plastic non-linear material response with an isotropic hardening rule was used for the foundation. This way, elastic deformations are recovered, but not plastic deformations.

The simplified model used in this investigative study is not able to capture all the possible dynamic foundation behaviour, such as radiation damping and differential settlement. The hysteresis response of the rotational spring was not considered in the analyses owing to the investigative nature of the study. This is an aspect which needs attention in a more detailed follow-up study.

Shear wall

The shear wall was modelled with beam elements and lumped mass added at each floor level to account for the mass of the floor slab. The bilinear pushover curve of the shear wall was used to determine the required non-linear moment-curvature response for the shear wall material. The non-linear material response was therefore determined to ensure that an accurate displacement response is obtained. Elastic-plastic hysteresis behaviour with isotropic hardening was assigned to the non-linear material behaviour. Considering that this model was the best available from the software package, this approach is justified by the investigative nature of this pilot study.

A Rayleigh damping coefficient was assigned to the shear wall material to ensure behaviour similar to that of an equivalent single-degree-of-freedom system with a 5% viscous damping ratio.

Structural frame

A two-column model was used to represent the internal structural frame to ensure that the lateral displaced shape would represent frame action. The two columns were connected with rigid link elements to ensure that the columns undergo the same lateral displacement and rotation over the full height. Lumped mass was added at every floor level to account for the weight of the floor slabs which were not modelled. Note that the study did not investigate the degree to which the columns in the frame were stressed.

The non-linear moment-curvature behaviour was determined for the column material in much the same manner as for the shear wall to enable the column material to lead to the required lateral force-displacement behaviour as defined by the bilinear approximation of the pushover curve of the frame. Here also elastic-plastic hysteresis behaviour with isotropic hardening was assigned to the column material response. A damping coefficient was assigned to the column material to ensure a 5% viscous damping ratio.

Combined simplified model

It is assumed that the building will have a uniform lateral floor displacement and therefore the lateral displacement of the shear wall model and structural frame model is linked. Link elements that enforce equal lateral displacement were included at every floor level (Figure 10).

ANALYSES, RESULTS AND ASSESSMENT

Non-linear time-history analyses were performed by applying the ground acceleration to the base nodes of the model. Seven ground motion records were used for each of the six models with different foundation sizes. The maximum and average response were evaluated for the combined system (global assessment), as well as for each of the investigated systems individually (local assessment).

Local assessment of structural frame

Lateral roof displacement response was used to evaluate the performance of the structural frame. These response output results are shown in Figure 11.

From the pushover analysis of the structural frame (Figure 6), the first performance criterion that was exceeded was the yielding strain limit of reinforcing steel. Since this performance criterion was only exceeded at a lateral roof displacement of 390 mm, and the lateral displacement of the combined system is much less, no failures are expected in the structural frame.

The increasing trend in lateral roof displacement response with decreasing foundation size can be expected as the bending moment resistance of the foundation decreases with decreasing size.

Local assessment of foundations

Relative rotation response of the non-linear spring element was used to assess the performance of the various foundation sizes. Rocking of the foundation is allowed and
therefore yielding of the supporting soil is the only performance criterion investigated in this study (identified in Figure 8). The maximum and average response output are shown in Figure 12.

The general trend of increasing footing rotation with decreasing foundation size can be expected because of the resulting decrease in bending moment resistance.

From Figure 12 it is observed that the rotation limit resulting in yielding of the supporting soil is only exceeded for foundation sizes designed to resist 20% or less of the overstrength bending moment. Yielding of the supporting soil to the shear wall foundation can therefore be expected for only rather small foundation sizes.

**Local assessment of shear wall**

Lateral roof displacement relative to the footing rotation was used to evaluate the performance of the shear wall, as the shear wall capacity curve (Figure 4) was obtained by assuming a fixed support to the wall (i.e. relative wall displacement = total displacement – displacement due to footing rotation). The yielding strain limit of steel material was the first performance criterion to be exceeded that was identified from the pushover curve of the shear wall at a relative lateral roof displacement of 148.7 mm. Maximum and average relative lateral roof displacement response are shown in Figure 13. The large contribution of the wall stiffness to the total system stiffness, as explained earlier, allows the assumption that wall behaviour in the system can be compared to the behaviour of the wall on its own.

From Figure 13 a general decreasing trend is observed. The rocking effect of the foundation therefore generally has the effect of reducing strains in the shear wall. Very little rocking action can be expected from the foundation size designed to resist the full overstrength bending moment, and therefore a plastic hinge can be expected to form in the lower part of the shear wall to dissipate energy. From Figure 13 it is observed that the maximum response exceeds the steel material yielding strain limit for this foundation size. It follows that a plastic hinge mechanism can be expected to form in the lower part of the wall when the foundation is designed to resist the full overstrength bending moment.

Foundation rocking can be expected to contribute to energy dissipation. Rocking of the foundation (and therefore the contribution to energy dissipation) increases as the foundation size decreases. The required contribution from the plastic hinge mechanism to energy dissipation can therefore be expected to decrease with decreasing foundation size. This is evident from the decreasing trend in the relative lateral roof displacement of the shear wall.

**Global assessment**

Assessment of the lateral roof displacement of the combined investigated models...
was performed using the performance levels described in the Vision 2000 report (SEAOC 1995). This global evaluation requires a lateral capacity curve combining that of the shear wall, foundation and structural frame.

Assuming the same lateral displacement for all frames of the building due to stiff floor diaphragms, it follows that the lateral wall displacement due to footing rotation and wall flexibility should be equal to that of the internal frame. With the rotational stiffness of the shear wall footing converted to a corresponding translational stiffness, the spring analogy depicted in Figure 14 can be used to determine the contribution of the different systems to global (combined) stiffness.

The global lateral stiffness can therefore be calculated from Equation (1), where \( k_i \) represents the lateral stiffness of system \( i \).

\[
k_{\text{global}} = k_{\text{frame}} + \frac{k_{\text{wall}} - k_{\text{footing}}}{k_{\text{wall}} + k_{\text{footing}}}
\]

This combination of lateral stiffness resulted in global capacity curves for the different combined models. Yield and ultimate lateral roof displacement values of bilinear approximations of each of the capacity curves compared well with that of the shear wall. It can be seen from Figure 15 that all the investigated models had the same values for yield and ultimate lateral roof displacements.

The displacement values required for global assessment of the response output, as taken from Figure 15, are shown in Table 6.

These values were then used to obtain the limits shown in Table 7.

### Table 6: Global capacity displacement parameters

<table>
<thead>
<tr>
<th>Displacement</th>
<th>Symbol</th>
<th>Value [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yield</td>
<td>( \Delta_y )</td>
<td>76</td>
</tr>
<tr>
<td>Ultimate</td>
<td>( \Delta_u )</td>
<td>350</td>
</tr>
<tr>
<td>Plastic</td>
<td>( \Delta_p = \Delta_u - \Delta_y )</td>
<td>274</td>
</tr>
</tbody>
</table>

Performance levels defined in the Vision 2000 report (SEAOC 1995) were used for the global assessment, with lateral roof displacement limits as shown in Table 7.

### Table 7: Vision 2000 performance limits (SEAOC 1995)

<table>
<thead>
<tr>
<th>Performance level</th>
<th>Lower limit</th>
<th>Value [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Fully operational</td>
<td>( \Delta_y )</td>
<td>76.0</td>
</tr>
<tr>
<td>Operational</td>
<td>( \Delta_y + 0.3\Delta_p )</td>
<td>158.2</td>
</tr>
<tr>
<td>Life safe</td>
<td>( \Delta_y + 0.6\Delta_p )</td>
<td>240.4</td>
</tr>
<tr>
<td>Near collapse</td>
<td>( \Delta_y + 0.8\Delta_p )</td>
<td>295.2</td>
</tr>
<tr>
<td>Collapsed</td>
<td>( \Delta_u )</td>
<td>350.0</td>
</tr>
</tbody>
</table>

A graphic representation of the lateral roof displacement response is presented in Figure 16.

The investigated office building for general use is classified as a Basic Facility and the Vision 2000 report (SEAOC 1995) prescribes a minimum performance level of Life Safe for this type of structure. From Figure 16 it is observed that only the maximum lateral roof displacement response of the model with a foundation designed not to resist any bending moment exceeds the Life Safe displacement limit.

### CONCLUSIONS

Conclusions regarding the performance of the investigated structure can be made from the assessment of the non-linear time-history analysis response results.

### Structural frame

Neither concrete crushing failures nor flexural failures were identified for the structural frame. The increase of shear forces in slab elements at slab-column connections due to the transfer of unbalanced bending moments were not explicitly investigated in this study, and it is recommended that a future study should be focused on investigating this mode of failure. However, rotation limits used in this study provide a suitable approximate...
The aim of this study was to investigate the feasibility of reducing the size of shear wall foundations of buildings with no basement level in regions of moderate seismicity. From the assessment of displacement responses (of an example structure) resulting from non-linear time-history analyses, it follows that allowing the shear wall foundation to rock, may result in significantly smaller shear wall foundations within acceptable deformation limits.

With this feasible concept of using smaller shear wall foundations it is proposed that the investigation be extended to include a range of structures with different sizes and heights. It is proposed that full three-dimensional analyses and experimental tests be performed to further investigate the feasibility of implementing this concept. Subsequent research can also proceed to investigate the feasibility of the concept in terms of project cost and risk.

ACKNOWLEDGEMENTS
The comments by Dr Alessandro Dazio of the Rose School on this study and on the paper are much appreciated.

LIST OF SYMBOLS
- $k_i$: stiffness of system $i$
- $k$: confinement factor
- $M_{OBD}$: overstrength bending moment at base of shear wall
- $T_i$: period of the $i$th natural mode of vibration
- $\epsilon_c$: cover concrete spalling strain
- $\epsilon_{cc}$: confined concrete crushing strain
- $\epsilon_{s}$: steel material yield strain
- $\epsilon_{f}$: steel material fracture strain
- $\Delta_{pl}$: plastic lateral roof displacement
- $\Delta_{u}$: ultimate lateral roof displacement
- $\Delta_{y}$: yield lateral roof displacement

REFERENCES
A field study of the in situ moisture regime during active hydraulic tailings deposition

C J MacRobert, G E Blight

A common method of managing tailings in semi-arid environments is to self-impound the waste as it dries. This paper presents the results of research carried out on the in situ drying behaviour of platinum tailings. Monitoring of gravimetric water contents following deposition and the quantity of water released during sedimentation indicated that a tailings beach acts as a gravitational thickener. After sedimentation, water contents were observed to decrease at a rate correlated with reference evapotranspiration, reaching a steady-state condition. Field capacity values determined from laboratory experimentation and predictive modelling correlated closely with this steady state. Further near-surface moisture reduction is inhibited due to the relative abundance of moisture rising from below to replenish deficits. Liquidity indices demonstrated that, as a result, only the head of the beach dries to a steady-state condition which provides strength to impound the tailings.

INTRODUCTION
The world is faced with a growing problem of how to utilise its mineral resources in a sustainable manner. With the depletion of high-grade ore bodies and advances in extractive metallurgy, the volume of tailings requiring storage is ever-increasing. Typically, mill-circuits extract small percentages of the sought mineral from ore via a wet process, resulting in slurried tailings. The need to dispose of the tailings in an environmentally friendly, structurally stable and economically viable manner has resulted in a number of available solutions. Tailings disposal involves both process engineering for the thickening of tailings and geotechnical engineering for the deposition of tailings. Tailings may be thickened1 to a low-density slurry or a paste, or even filtered to a cake, increasing the surface area over which low-density slurries are deposited, water recoveries achieved from the dam can be similar to those from paste thickening. Tailings then require impounding or can self-impound if the geotechnical behaviour results in sufficient strength within realistic time-frames. No single solution is universally applicable as each has its merits. This paper builds on research previously presented by the authors (Blight et al 2012) which highlighted the gravitational thickening of low-density slurries during subaerial deposition, and extensively investigates the drying phase leading to strength gain in platinum tailings.

Moisture loss and strength gain
The strength gain of deposited “conventional” slurries is dependent on increasing inter-particle forces during sedimentation, consolidation and, finally, drying. Pane & Schiffman (1985) linked the Kynch (1952) theory of sedimentation and the Gibson et al (1967) theory of non-linear, finite strain consolidation of saturated clays, to model the pelagic (i.e. in the open ocean) sedimentation and simultaneous consolidation of sediments. Although applied to tailings deposition, conditions are often not pelagic, with deposition characterised by unstable, meandering, turbulent rill flow (Blight & Bentel 1983). Wells & Robertson (2003) concluded that only water liberated during initial sedimentation and bleed water during consolidation are available for recovery. Lyell et al (2008) proposed that by reducing the surface area over which low-density slurries are deposited, water recoveries achieved from the dam can be similar to those from paste thickening.

As solids settle out of the slurry a viscous mass results, with the quantity of pore water above the liquid limit, \( w_c \). Donaldson (1960) showed that in semi-arid environments evaporative drying results in the tailings losing pore water and gaining strength at a rate faster than can be estimated by the self-weight consolidation theory. The movement of pore water is governed by gravitational forces and suction-induced potentials. The reduction in pore water, quantified by the gravimetric water content, \( w_c \), through seepage and evaporation, is proportional to the reduction in volume while the soil skeleton...
remains saturated. When the solids skeleton reaches a maximum density, air enters the voids and the resulting pore water menisci create an increase in suction. The relationship between volume change and water content or suction at constant (zero) total stress is termed the Volumetric Shrinkage Curve (VSC), while the corresponding relationship between water content and suction is termed the Suction-Water Content Curve (SWCC) (Fredlund & Rahardjo 1993).

As suctions develop, gravitational seepage reduces and eventually water may be drawn up from within the deposit via the suction potentials (i.e. capillary action). When gravitational seepage becomes negligible, the soil has reached its field capacity, with classical soil science defining the corresponding suction to be 33 kPa, although this is often closer to 10 to 20 kPa (Miller & Donahue 1990). Koorevaar et al (1999) suggested that when the water table is at shallow depths, field capacity corresponds to hydrostatic equilibrium. Due to the reduction in permeability as suctions develop, Meyer & Gee (1999) proposed that field capacity occurred when the hydraulic conductivity decreases to between $10^{-9}$ and $10^{-11}$ m·s$^{-1}$, depending on soil type and texture. Ratliff et al (1983) investigated 401 soils and found the average standard deviation of laboratory and in situ field capacity values of similarly textured soils to be 3.6% and 4.9% respectively. Depending on the availability of water to replenish deficits, soils can remain at field capacity for considerable periods.

Evaporation from bare soil is often characterised as occurring in two distinct stages: the “energy limited” stage and the “falling rate” or “soil limited” stage. During the energy-limited stage, the rate is defined by the evaporative energy available to vaporise water within the soil. During the soil-limited stage, the hydraulic properties of the soil are unable to supply moisture at the potential evaporation rate. Allen et al (2005) analysed a number of soils and determined that the change from stage one to stage two evaporation occurs at a $w_c$ between the field capacity and wilting point. The wilting point is defined as the $w_c$ at a suction of 1 500 kPa (Miller & Donahue 1990). Stage two continues until the $w_c$ is half the wilting point value, which is presumed to approximate air-dry conditions.

**Reference evapotranspiration**

Evapotranspiration is the combined process whereby liquid water contained within soils and plant tissue is vaporised. The rate of evapotranspiration is influenced by macroclimatic conditions, microclimatic conditions and the development and management of plant type. A means to determine a reference evapotranspiration, $ET_0$, to reflect purely macroclimatic conditions, was developed by the Food and Agriculture Organization (FAO) (Allen et al 1998). Two methods are outlined: the FAO Penman-Monteith Method, Equation 1, when sufficient climatic data are available, and the Hargreaves Method, Equation 2, for limited climatic data. To further improve the performance of the Hargreaves method, linear regression between $ET_0$ values from the two methods is proposed to develop a regional calibration.

$$ET_0 = \frac{0.408\Delta(R_u - G) + \gamma \frac{900}{T + 273} u_2 (e_a - e_s)}{\Delta + \gamma (1 + 0.34u_2)}$$

(1)

where

- $ET_0 =$ reference evapotranspiration (mm·day$^{-1}$)
- $R_u =$ net radiation at the crop surface (MJ·m$^{-2}$·day$^{-1}$)
- $G =$ soil heat flux density (MJ·m$^{-2}$·day$^{-1}$)
- $T =$ daily mean temperature (°C)
- $u_2 =$ wind speed at a height of 2 m (m·s$^{-1}$)
- $e_a =$ actual vapour pressure (kPa)
- $e_s =$ saturation vapour pressure (kPa)
- $\Delta =$ slope of the vapour pressure curve (kPa°C$^{-1}$)
- $\gamma =$ psychrometric constant (kPa°C$^{-1}$)

$$ET_0 Hargreaves = 0.0023 \times \left( \frac{T_{max} + 17.8}{T_{max} - T_{min}} \right) \times R_a$$

(2)

where

- $ET_0 =$ reference evapotranspiration (mm·day$^{-1}$)
- $T_{mean} =$ daily mean air temperature (°C)
- $T_{max} =$ daily maximum air temperature (°C)
- $T_{min} =$ daily minimum air temperature (°C)
- $R_a =$ extraterrestrial radiation (mm·day$^{-1}$)

A review of the literature and analysis of carefully collected lysimeter data over a three-year period by Benli et al (2010) in arid to semi-arid climates indicated the FAO Penman-Monteith method to be generally superior, with variable results from other methods. Hargreaves and Allen (2003) outlined the historical development of the Hargreaves method and assessed it against 59 000 global sites finding it universally applicable, although it tended to underpredict in climates with high wind and overpredict in climates with high relative humidity.

**AIMS AND OBJECTIVES**

As outlined above, it has long been accepted that strength gain in tailings is driven by the drying effect of climate aridity. Although
extensive laboratory research has been conducted, field experiments have often been limited. The aim of this study, therefore, was to investigate the in situ moisture regime of platinum tailings over 11 months on two active, back-to-back tailings dams with the objectives of:

1. Determining the quantity of water released by gravitational thickening during deposition.
2. Developing an empirical correlation between ET₀ and the rate of moisture loss.
3. Demonstrating the steady state that limits the degree of moisture loss during active deposition.

**TEST WORK**

Test work was carried out at Anglo American Platinum Mogalakwena mine, 30 km north of Mokopane, South Africa (23°59’S, 28°56’E). During testing the mine operated two back-to-back facilities: Dam 1, a conventional upstream spigoted facility (raised at 2.3 m·year⁻¹) and Dam 2, a waste rock impoundment, filled via a series of spigots (raised at 4.6 m·year⁻¹). Two separate processing plants supplied similar tailings: the South Plant to Dam 1 and the North Plant to Dam 2.

Figure 1 shows the site plan, indicating the test strips, locations of the rain gauge, A-pan and site office which housed a small laboratory for determining gravimetric water content.

<table>
<thead>
<tr>
<th>Test strip</th>
<th>Deposition Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailings Dam 1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>12 January – 17 January</td>
</tr>
<tr>
<td>2</td>
<td>2 March – 10 March</td>
</tr>
<tr>
<td>3</td>
<td>12 April – 23 April</td>
</tr>
<tr>
<td>4</td>
<td>8 June – 18 June</td>
</tr>
<tr>
<td>Tailings Dam 2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21 July – 26 July</td>
</tr>
<tr>
<td>2</td>
<td>10 September – 19 September</td>
</tr>
<tr>
<td>3</td>
<td>19 October to 26 October</td>
</tr>
</tbody>
</table>

† On the 4th of May the test strip was flooded beyond a distance of 100 m from the head of the beach due to operational constraints; however, measurements were still carried out.

### Table 1 Details of raw data

<table>
<thead>
<tr>
<th>Data source</th>
<th>Raw data obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beach sampling via bulk samples, grab samples and auger samples</td>
<td>Particle size distributions, particle specific gravities, gravimetric water content, calibrated gypsum block suction tests and triaxial permeability tests</td>
</tr>
<tr>
<td>Site climatic data</td>
<td>A-Pan evaporation, rainfall and daily minimum and maximum temperatures</td>
</tr>
<tr>
<td>South African Weather Service, Mokopane Station</td>
<td>Daily temperature, wind speed and relative humidity</td>
</tr>
<tr>
<td>Historical monitoring and design data</td>
<td>Atterberg limits, evaporative drying tests, filter paper suction tests and triaxial permeability tests</td>
</tr>
</tbody>
</table>

![Figure 2 Beach surveys](image)

![Figure 3 Particle size distribution on Dam 1](image)

intention of depositing 400 mm of material each time on each test strip. However, the depth of material deposited was not uniform due to the inherent beaching behaviour. The rate of rise along each test strip was determined by survey data, shown in Figure 2, and the time between depositions, given in Table 1. For the outer 100 m of both test strips, the rate of rise was roughly 2.5 m·year⁻¹. On Dam 1 the rate for the remainder of the 400 m beach was roughly 1.2 m·year⁻¹, whereas on Dam 2 the rate for the remainder of the 200 m beach was roughly 4.0 m·year⁻¹. Table 2 details the raw data collected.

### ANALYSIS OF RESULTS

#### Basic geotechnical parameters

To fully liberate Platinum Group metals, IsaMill® technology (trademark of Xstrata Technology) was being introduced at the two plants to produce a finer grind of material. During testing the technology had been partially introduced at the South Plant and...
was fully operational at the North Plant. Consequently, the grading envelope along the Dam 1 beach was wider than that along the Dam 2 beach, as illustrated by Figures 3 and 4. Grading curves at the point of deposition (0 m) contained slightly more coarse material with little segregation along the rest of the beach.

The mean particle specific gravity, $G_s$, was 3.10, as determined by the vacuum method from 31 samples, with a standard deviation of 0.03. Atterberg limits, determined from five data sets, indicated that the liquid limit, $w_L$, varied between 25% and 20% with a mean of 23%, and that the plastic limit, $w_P$, varied between 21% and 18% with a mean of 19%. Based on the average plasticity index and average clay fraction, the activity ($w_L – w_P / %$ clay fraction) was 0.3, indicating that the clay fraction contains only a small portion of clay minerals.

The permeability was determined on two remoulded samples of representative material during the consolidation stages of triaxial testing during the study. Permeabilities were also determined on remoulded samples, during the consolidation stage of triaxial testing, as part of mine operational monitoring. The results from both sets of test work are summarised in Table 3.

### Table 3: Triaxial permeability test results

<table>
<thead>
<tr>
<th>Study</th>
<th>Cell pressure (kPa)</th>
<th>Permeability (m·s⁻¹)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Current</td>
<td>100</td>
<td>9.0 ×10⁻⁸</td>
<td>7.5 ×10⁻⁸</td>
<td></td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>6.0 ×10⁻⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operational monitoring</td>
<td>350</td>
<td>6.5 ×10⁻⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>5.9 ×10⁻⁸</td>
<td>5.2 ×10⁻⁸</td>
<td></td>
</tr>
<tr>
<td></td>
<td>700</td>
<td>3.2 ×10⁻³</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Water recovery during gravitational thickening**

The water released from the slurry during sedimentation was investigated by taking successive grab samples after cessation of deposition. Relevant average slurry properties during deposition, and the corresponding values after sedimentation, are given in Table 4. These values were calculated from $w_c$, using the mean $G_s$, due to the small scatter in values and assuming full saturation, as samples were taken from the beach surface soon after deposition. Based on the successive grab samples, sedimentation was observed to be complete within 65 hours (7 hours standard deviation) with the time increasing down the beach towards the pool.

### Table 4 Water release during gravitational thickening

<table>
<thead>
<tr>
<th></th>
<th>Relative density (kg·m⁻³)</th>
<th>Solids content, $c_w$ (%)</th>
<th>Water content, $w_c$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dam 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average slurry properties</td>
<td>1 500</td>
<td>49</td>
<td>104</td>
</tr>
<tr>
<td>Average settled slurry properties</td>
<td>1 800 – 2 000</td>
<td>66 – 74</td>
<td>52 – 35</td>
</tr>
<tr>
<td>Dam 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average slurry properties</td>
<td>1 400</td>
<td>42</td>
<td>138</td>
</tr>
<tr>
<td>Average settled slurry properties</td>
<td>1 925</td>
<td>71</td>
<td>41</td>
</tr>
</tbody>
</table>

**Table 5 Water balance**

<table>
<thead>
<tr>
<th>Component</th>
<th>Methodology of determining value</th>
<th>Dam 1 (April to November)</th>
<th>Dam 2 (April to June)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed water (Mm³)</td>
<td>Calculated from half-hourly slurry density and flow rates</td>
<td>2 941</td>
<td>1 796</td>
</tr>
<tr>
<td>Dry tons (Mt)</td>
<td></td>
<td>2 964</td>
<td>1 431</td>
</tr>
<tr>
<td>Decant flow (Mm³)</td>
<td>Calculated from daily flow over calibrated weirs</td>
<td>1 681</td>
<td>705</td>
</tr>
<tr>
<td>Drain flow (Mm³)</td>
<td>Calculated from daily drain flow measurements</td>
<td>22</td>
<td>47</td>
</tr>
<tr>
<td>Evaporation (Mm³)</td>
<td>Calculated by applying 60% of measured A-Pan evaporation over the entire dam area†</td>
<td>803</td>
<td>196</td>
</tr>
<tr>
<td>Precipitation (Mm³)</td>
<td>Calculated by applying measured rain gauge rainfall over entire dam area</td>
<td>250</td>
<td>180</td>
</tr>
<tr>
<td>Water recovered from tailings (Mm³)</td>
<td>Decant flow + Drain flow</td>
<td>1 703</td>
<td>752</td>
</tr>
<tr>
<td>Percentage water recovered (%)</td>
<td>Ratio of water recovered to feed water</td>
<td>58</td>
<td>42</td>
</tr>
<tr>
<td>Apparent settled relative density (kg·m⁻³)</td>
<td>Calculated from unrecovered water and dry tons</td>
<td>1 900</td>
<td>1 600</td>
</tr>
<tr>
<td>Apparent settled solids concentration, $c_w$ (%)</td>
<td>Calculated from unrecovered water and dry tons</td>
<td>71</td>
<td>58</td>
</tr>
<tr>
<td>Apparent settled water content, $w_c$ (%)</td>
<td>Calculated from unrecovered water and dry tons</td>
<td>41</td>
<td>72</td>
</tr>
</tbody>
</table>

† The large difference in evaporation quantities on the two dams is due to the duration of investigation and seasonal differences during which the respective water balances were determined.
One anomalous result, not included in Table 4, was the average settled conditions following the first deposition on Dam 1. Owing to operational constraints, the test section had remained dormant for six months, resulting in a dry beach, and sampling only took place five days after deposition had ceased. Consequently, a large volume of water was drawn into the underlying dry layers, resulting in average settled $w_c$ values along the beach of 20% ($c_w = 83\%$).

An investigation by the mine (Vietti et al. 2010) with a view to implementing paste technology illustrated that a slurry density of 1 900 kg·m$^{-3}$ ($c_w = 70\%, w_c = 43\%$) could be achieved for similar material. The experimental evidence discussed in this paper suggests that, provided deposition is not onto a dry beach, and all decant water is recovered, the beach acts as a gravitational thickener at no additional cost.

To investigate how much of the supernatant water released during sedimentation is recovered, water balances were determined from April to November on Dam 1 and from April to June on Dam 2. Table 5 shows the water balances components, the quantity of water recovered and the apparent settled conditions determined from the unrecovered water volume and dry tons deposited. On Dam 1 the apparent and measured settled conditions were comparable, confirming that the water released was recovered. On Dam 2 the apparent settled conditions and measured values differed, suggesting that some of the water released during sedimentation was not recovered. A possible reason for this is that 75% of the outer wall on Dam 2 and the pool-training wall were constructed from highly pervious waste rock. During field work, ponded water was observed between the toe of the rock wall and the solution trench, even though it was the dry season. It is plausible to say that the unaccounted water seeped through the rock wall interface.

### Correlating reference evaporation and moisture loss

#### Moisture loss regression model

A multiple straight-line regression model was fitted to the observed moisture loss based on gravimetric $w_c$. The initial $w_c$ was taken from plant slurry densities. Surface
samples were recovered and tested during sedimentation, drying and apparent steady-state. On Dam 2 a total of three depositions were sampled at five sampling points 50 m apart, resulting in 15 data sets. An iterative statistical method was used to intersect three separate linear regression curves through the sedimentation stage, drying stage and steady-state condition. The methodology followed is outlined in Table 6.

Figures 5 to 9 illustrate the fitted regression models, along with the raw data from the 15 test depositions. An iterative statistical method was used to intersect three separate linear regression curves through the sedimentation stage, drying stage and steady-state condition. The methodology followed is outlined in Table 6.

Figures 5 to 9 illustrate the fitted regression models, along with the raw data from the 15 test depositions. An iterative statistical method was used to intersect three separate linear regression curves through the sedimentation stage, drying stage and steady-state condition. The methodology followed is outlined in Table 6.

Figures 8 and 9 illustrate the fitted regression models, along with the raw data from the 15 test depositions. An iterative statistical method was used to intersect three separate linear regression curves through the sedimentation stage, drying stage and steady-state condition. The methodology followed is outlined in Table 6.

Figures 10 illustrates the performance of all the multiple straight-line regression models against the field data. Individual scatter plots showed that the performance of the model was poorer for the sampling points at 150 m and 200 m along the beach.

Reference evapotranspiration
A regional calibration of the Hargreaves method was developed from climatic data obtained from the South African Weather Bureau station at Mokopane:

$$ET_0^{\text{FAO}} 56 – \text{Penman Monteith} = 1.1 \times ET_0^{\text{Hargreaves}} - 0.7 \quad (R^2 = 0.91) \quad (3)$$

$ET_0$ values calculated for the test site were found to compare adequately with historical data from Polokwane and Mokopane. The mean ratio between the A-Pan data and $ET_0$ values was 0.65, which is consistent with expected behaviour reported by Allen et al. (1998). However, a great deal of scatter was observed. This is considered a consequence of variable microclimatic conditions and operational constraints with the A-Pan.

Empirical correlation
An empirical correlation between the change in $w_c$ per day during the drying stage for the outer 100 m and $ET_0$ was developed, as shown in Table 7. In order for the coefficient to be compared with other sites, it was necessary to correlate the rate of drying with a macroclimatic measure of evaporative energy. This is independent of the dam surface microclimate that resulted in scattered A-Pan data. It is clear that the rate of drying is slower during the winter periods when evaporative energy is low and high during the summer period. The steady state reached, as shown in the next subsection, is field capacity, when seepage losses become negligible. Up to this point water loss is due to both evaporation and seepage.

Evaporative energy varies throughout the year. However, the rate of seepage is controlled by the materials’ permeability, which remains essentially constant (assuming the material remains saturated and density does not change significantly). During winter the evaporative energy is lower and more moisture may be lost through seepage, with the opposite being the case during summer. It is also likely that this water bleeds up to the surface and is recovered as the material consolidates. This is illustrated by the k-values being slightly higher during winter than in summer, but a longer study would be required to quantify this variation. The relationship would need to be developed taking into account evaporation, seepage and bleed water. Figure 11 illustrates the time required...
for the sedimentation step and then, using the average k-value, the number of days to reach steady state based on daily ET₀ values. Such relationships can be used to optimise the safe development of tailings dams.

**Steady state**
The wₑ regime was investigated by analysing auger samples taken at 200 mm intervals to a depth of 1 m, increasing to 2.5 m as the study progressed. The large quantity of data generated enabled the apparent steady state to be investigated relative to sampling position, sampling time and material parameters.

**Laboratory geotechnical model**
Five laboratory evaporative drying tests and a filter paper test were used to determine the Volumetric Shrinkage Curves (VSC) illustrated in Figure 12. All the samples dried along the zero air voids curve to a wₑ of roughly 27%, after which air entry occurred. After air entry the dry density remained essentially constant, at roughly 1,700 kg·m⁻³, consistent with accepted in situ dry densities for the material. Box plots (indicating the minimum, maximum and interquartile values) of all the wₑ values determined for the respective dams are included in the figure for comparison. This illustrates that roughly 25% of the samples were taken before peak density, whereas 75% of the samples were taken at varying degrees of desaturation.

Three separate Suction-Water Content Curves (SWCCs) were determined. Two drying and subsequent wetting curves were determined with calibrated gypsum blocks on samples before and after the introduction of the finer grind (Tests 1 and 2 in Figure 13). The third drying curve was determined by the filter paper method on a prepared sample of finer grind material (Test 3 in Figure 13). The computer program RETC (Van Genuchten et al. 1991) was used to estimate the SWCC from the raw data based on the Van Genuchten model. Figure 13 illustrates the curves obtained: the range of air entry values, the range of field capacity values at 33 kPa and asymptotic suction increase below a wₑ of 14%. Air entry occurred at an average wₑ of 25% and a standard deviation of 4.9%. The average field capacity was 21% with a standard deviation of 4.3%, similar to the range reported by Ratliff et al. (1983). The range in values reflects the slight differences in grading of the samples tested. Due to the fine nature of the material, the SWCCs have a steep gradient which results in air entry and field capacity occurring at close wₑ values.

The computer program SEEP/W (GeoStudio 2007) was used to predict the relationship between wₑ and hydraulic conductivity. The volumetric water content function was determined using the modified Kovács method proposed by Aubertin et al. (2003). This method assumes that both capillary and adhesive forces act simultaneously to induce suction and is applicable to tailings, silts and various other soils. The input variables are the particle size corresponding to 10% and 60% passing on the particle size distribution curve, porosity, and liquid limit.
To account for the range in particle sizes along each beach, the minimum, maximum and quartile grading parameters were used. The porosity was calculated to be 0.48 from the dry density of 1 700 kg·m⁻³ and particle specific gravity of 3.10. The method uses the liquid limit in an empirical expression to calculate the specific surface area for fine-grained plastic-cohesive materials. This expression, however, is required only when the \( w_L \) is above 30%, which was not the case for this material as the average \( w_L \) was 23%.

The hydraulic conductivity function was estimated from the volumetric water content function, saturated hydraulic conductivity and residual water content using the Van Genuchten method. The average saturated hydraulic conductivity of 7.5×10⁻⁸ m·s⁻¹ was used. The residual volumetric water content, 4.6%, was determined from the \( w_c \) at a suction of 1 500 kPa (Van Genuchten 1980) using the filter paper SWCC. This is roughly 10% of the saturated volumetric water content. Volumetric water contents were converted to \( w_c \) using phase relationships. Meyer & Gee (1999) suggested that field capacity occurs when the hydraulic conductivity decreases to between \( 10^{-9} \) and \( 10^{-11} \) m·s⁻¹, depending on soil type and texture. Figure 14 illustrates the hydraulic conductivity curves and the range in field capacity values at a hydraulic conductivity of \( 10^{-9} \) m·s⁻¹. The values at \( 10^{-11} \) m·s⁻¹ were lower than the laboratory-determined values and therefore not realistic. The average and standard deviation of predicted field capacity values were 22% and 6% for Dam 1, and 23% and 4% for Dam 2. This was comparable to the laboratory-determined values (average 21% and standard deviation of 4.3%). The larger variability of the predicted values on Dam 1 compared with Dam 2 is a consequence of the wider grading envelope for the former, highlighting the sensitivity of soil-suction characteristics to particle size distribution.

**Field behaviour**

Figures 15 and 16 illustrate that there was no correspondence between \( w_c \) values at incremental depths over time. Rather, \( w_c \) values varied within a narrow band that did not change within the time-frame of a depositional cycle. The focus of the study therefore was to explore what this distribution was, whether it changed with depth or over longer time steps, what impact the weather had and what material properties controlled it. Finally, the impact of this limitation on drying, to strength gain, was assessed.

The entire data set was analysed against depth, and Figure 17 shows the mean and variance of \( w_c \) at each depth increment. The mean is essentially constant, although the variance is roughly 50 for the top 500 mm, 35 from 500 to 1 000 mm and 20 below 1 500 mm. The large dispersion of the data for the first 500 mm is a consequence of...
samples being at varying stages of drying, from sedimentation to air-dry. As the depth increases, this dispersion decreases as the influence of evaporation becomes less.

Figures 18 and 19 indicate the deposition sampling periods for the respective dams, with black dots indicating the sampling dates. The ranges of $w_c$ determined for each sampling period are represented by separate histograms, with row lengths representing the relative frequency. The mean $w_c$ for the entire data set was 27% with a standard deviation of 6%. Therefore, assuming a normal distribution, 98% of the data lie between 41% and 14%. These values correspond to the mean and the degree of skew about the mean of the entire data set.

Initial sampling on Dam 1 was to a depth of 1 m and on a beach that had been dormant for six months due to catamaran commissioning constraints. Figure 18 illustrates that, as a result, the $w_c$ values were prominently skewed to low values indicative of large suctions. The first deposition “hung up” on the dry beach with virtually no material deposited past 150 m. The large skew in the baseline sampling was not observed during active deposition, although the distributions for the first three depositions during summer and autumn were skewed to lower values. As the beach did not dry out as much during active deposition, subsequent depositions were more uniform. Substantial rainfall during the second deposition highlighted the limited effect of these events on drying. The final deposition during winter on Dam 1 was slightly skewed to values higher than the mean of the entire data set. During winter the $ET_0$ values were markedly lower, reflecting the lower evaporative conditions. Although the sampling period was similar to previous depositions, less drying took place.

Before the test work, Dam 2 was raised at roughly twice the rate of Dam 1 and consequently the initial sampling resulted in a distribution skewed to high values. Three test depositions on Dam 2, in spring and summer, were at similar rates of rise to that of the outer portion of Dam 1. Despite the high evaporative conditions, the distributions remained slightly skewed to values above the mean. Dam 2 was therefore significantly moister than Dam 1, although the sampling window and evaporative conditions were similar. The wet baseline conditions appeared to have persisted throughout. It is also postulated that the finer material and shorter beach length affected these results.

To explore the controlling effect of field capacity on the degree of moisture loss, the following null and alternative hypotheses were tested:

$H_0$: Water contents at each sampling point have a different population mean to the field capacity mean.

$H_1$: Water contents at each sampling point have the same population mean to the field capacity mean.

The two-tailed $t$-test with unequal variances was used to test the hypotheses. The variances were assumed to be different as the field samples were taken at various stages of drying, whereas field capacity has a narrower variance. The predicted field capacity values for each dam were used as these were assumed to better represent the grind differences.

Figure 20 illustrates the distribution of $w_c$ at each position along the beach against the results of the hypothesis testing for Dam 1. The alternative hypothesis is accepted for the outer 250 m, although the degree of confidence decreases from being high at the head of the beach. Past 250 m the hypothesis is rejected as the probability is less than 0.05. Following air entry, determined to occur
from a \( w_c \) of 27%, suctions start to develop. As the material reaches field capacity, seepage becomes negligible with moisture loss being dominated by evaporation. Initially, capillary rise will provide moisture at a rate equal to the evaporation rate. If suctions can continue to develop, the hydraulic properties will restrict this moisture loss and stage two evaporation starts, continuing until the material is air-dry. During active deposition only a few samples indicated that suctions larger than field capacity had developed. It is assumed that seepage into base layers during deposition rises during drying to replenish the developing deficits. As illustrated earlier, it was only during the six months of dormancy on Dam 1 that this moisture was depleted and stage two evaporation was observed to have occurred. Hence, during active deposition the degree of moisture loss is controlled by the material’s field capacity. Further along the beach moisture can also be obtained from the closer phreatic surface springing from the pool. This is illustrated by the larger portion of saturated \( w_c \) values past 250 m due to equilibrium with the phreatic surface.

Figure 21 illustrates the relationship for Dam 2, with the alternative hypothesis accepted for the outer 100 m, albeit with less confidence, and rejected for the remainder of the beach. From the baseline sampling, this dam contained significantly more moisture from previous depositions to replenish any developing deficits. Equilibrium with the phreatic regime appears to occur higher up the beach and this is postulated to be the result of the phreatic surface not becoming as depressed along the shorter beach.

To illustrate the effect of the limited drying that occurs on strength gain, liquidity indices \( I_L \) were calculated from the average Atterberg limits.

\[
I_L = \frac{w_c - w_p}{w_L - w_p} (4)
\]

The \( I_L \) values have been divided into three categories: less than 0 (i.e. \( w_c < w_p \)), between 0 and 1 (i.e. \( w_c \) between the \( w_p \) and \( w_L \)) and greater than 1 (i.e. \( w_c > w_L \)). Negative \( I_L \) values are indicative of heavily overconsolidated soil deposits (Das 2008) and high shear strengths (Bovis 2003). Values between 0 and 1 are indicative of soils that behave plastically, resisting shear failure, whereas values greater than 1 are indicative of soils that are extremely sensitive to breakdown if sheared (Holz & Kovács 1981). These categories are shown graphically in Figures 22 and 23 respectively, in relation to sampling positions relative to the final elevation at the end of the test work.

On Dam 1 it is apparent that only the outer section reached a state of high shear strength, with 48% of the samples having an \( I_L \) less than 1 and 34% being negative-based on the average Atterberg limits. However, 75% of the \( w_c \) values at this position were lower than the maximum \( w_L \) of 25%. After 50 m the proportion of samples with an \( I_L \) less than 1 was on average 15% for all sampling points, being slightly higher at 50 m and decreasing towards the pool. Thus, the majority of the interior is prone to fail under shear, although it is able to support a man and prevent the auger hole from collapsing. At the head of the beach the distribution with depth is fairly constant for the baseline beach and the first two depositions. As
mentioned, the baseline conditions were dry due to the extensive dormant period. The first two depositions during months with high evaporative conditions resulted in a similar strength gain. The third deposition was done under evaporative and rainfall conditions similar to those of the second deposition, but appeared to result in less strength gain. On closer inspection, 65% of the samples were below the maximum $w_c$ so it is likely that this material was slightly finer. Due to the low evaporative energy during the final deposition, only 30% of the samples reached a high strength state during the sampling period.

Baseline sampling on Dam 2 indicated that it had not gained significant strength, with only 30% of the samples at the head of the beach having an IL below 1 and the average for the remainder of the beach being 5%. This is attributed to the high rate of rise prior to the test deposition. Analysis at the head of the beach within the first two test depositions indicated that 51% of the samples had an IL below 1, with 75% of the samples being below the maximum $w_c$, as was the case for Dam 1 at this position. This drying front also appeared to extend roughly 1 m below the first deposition, yet insufficient strength gain appeared to have occurred during the shorter sampling window of the final deposition.

**DISCUSSION AND CONCLUSIONS**

This paper has reported on extensive field and laboratory work to guide mining houses in disposing of tailings. Test work was carried out on two back-to-back facilities receiving similar tailings at Anglo American Platinum Mogalakwena Mine. Dam 1 is an upstream spigot facility and Dam 2 is a waste rock impoundment facility. On-beach sampling was undertaken with laboratory test work carried out on site and at the University of the Witwatersrand, and additional data were sourced from monitoring records. The following conclusions are drawn from the study in order of the stated objectives.

1. On-beach sampling illustrated that within 65 hours (7 hour standard deviation) the settled $w_c$ values equalled those anticipated from mechanical paste thickening. Deposition onto a dry beach indicated that cycle times should maintain wet beaches to release water. A rate of rise of 2.5 m/year$^1$ over the 100 ha Dam 1 was adequate to achieve this. Water balances indicated that the water released on Dam 1 was indeed recovered, providing a significantly cheaper means of recovering water. Similar water recoveries were not observed on Dam 2; it is assumed that this was due to water losses through the highly pervious waste rock impoundment.

2. Following sedimentation, the rate of change of $w_c$ values during the drying stage was correlated with ET$_0$, a macroclimatic measure of evaporative energy. The average ratio or k-value was calculated to be 0.21, being slightly lower during the winter months due to lower evaporative energy. The effects of seepage and consolidation bleed water on moisture loss over the course of the year need to be investigated further to fully define the k-value. Such relationships can be used to optimise the safe development of tailings dams.

3. After the drying stage a steady state developed, with $w_c$ values varying within a narrow band. The mean $w_c$ for the entire 11-month data set was 27% (standard deviation of 6.0%) which was also determined to be the air entry value. Assuming a normal distribution, 98% of the data were between 41% and 14%. These values corresponded to the average settled $w_c$ and the $w_c$ below which asymptotic suctions developed in suction tests.

With depth the mean $w_c$ remained constant, although the dispersivity decreased substantially. The freshly deposited layer exhibited the greatest dispersivity, with samples at varying stages from settled to air dry. Below this layer the dispersion was lower, decreasing steadily to 1 500 mm before remaining constant, suggesting that this is the limit of evaporative influence.

The $w_c$ distribution was observed to vary slightly from one deposition to the next, reflecting changes in evaporative conditions and baseline conditions. The Dam 1 baseline was significantly skewed to low values, indicative of high suctions due to six months’ dormancy over spring and summer. Test depositions in summer and autumn resulted in more values below the mean $w_c$, with the opposite being the case for the one winter deposition.

Baseline sampling on Dam 2 indicated skew to high $w_c$ values due to the high rate of rise prior to test depositions. Test depositions were similarly negatively skewed, albeit to a much smaller extent, although the sampling period and evaporative conditions were similar to those of Dam 1. The wet baseline conditions appeared to have restricted drying.

During active deposition $w_c$ distributions were controlled by the material reaching field capacity. Distributions for the outer section of the beach matched those of classically defined material field capacity values as partial suctions developed. During deposition, seepage into the beach rises to replenish deficits, preventing the development of suctions greater than field capacity. The baseline dormancy on Dam 1 illustrates that over long periods this source of moisture can be diminished and air-dry conditions may be approached. Closer to the pool the distributions were observed to be in equilibrium with the phreatic surface due to the greater portion of saturated values. This observation was more pronounced on Dam 1 than on Dam 2, presumably due to the fact that phreatic surfaces become more depressed along longer beaches.

Particle size distributions at the head of the beach contained slightly coarser material which contributed to the drier conditions at this point, despite the minimal beach segregation. This was a result of the IsaMill™ technology utilised by the mine, leading to a predominantly silty material being deposited.

As a consequence of these factors controlling moisture loss, liquidity indices showed that substantial shear strength developed only within the first 25 m to 50 m of the beach. Liquidity indices for the remainder of the beach indicated that the material was prone to fail under shear. This observation was made for all test depositions where the rate of rise for the outer sections was 2.5 m/year$^1$. The baseline conditions on Dam 2 did not exhibit this strength gain at the beach head due to the 4.6 m/year$^1$ rate of rise — hence the requirement of a waste rock impoundment.

**ACKNOWLEDGEMENTS**

The authors wish to acknowledge Anglo American Platinum, Fraser Alexander Tailings, SRK Consulting and the South African Weather Bureau for making this study possible.

**LIST OF SYMBOLS**

- $\Delta$: slope of the vapour pressure curve (kPa·°C$^{-1}$)
- $c_w$: solids concentration by mass
- $e_a$: actual vapour pressure (kPa)
- $e_s$: saturation vapour pressure (kPa)
- ET$_0$: reference evapotranspiration (mm·day$^{-1}$)
- $G$: soil heat flux density (MJ·m$^{-2}$·day$^{-1}$)
- $G_s$: particle specific gravity
- IL: liquidity index
- $R_e$: extra-terrestrial radiation (mm·day$^{-1}$)
- $R_n$: net radiation at the crop surface (MJ·m$^{-2}$·day$^{-1}$)
SWCC Suction-Water Content Curve

T daily mean temperature (°C)
T\text{\textsubscript{max}} daily maximum air temperature (°C)
T\text{\textsubscript{mean}} daily mean air temperature (°C)
T\text{\textsubscript{min}} daily minimum air temperature (°C)
u\text{\textsubscript{2}} wind speed at a height of 2 m (m·s\textsuperscript{-1})
VSC Volumetric Shrinkage Curve
w\textsubscript{c} gravimetric water content
w\textsubscript{L} liquid limit
w\textsubscript{P} plastic limit
γ psychrometric constant (kPa·°C\textsuperscript{-1})

NOTE
1 Here thickening is a process step whereby the solids concentration by mass, c\text{\textsubscript{w}}, is increased by a mechanical process. Low-density slurry has c\text{\textsubscript{w}} values between 30% and 50%, paste has c\text{\textsubscript{w}} values between 50% and 65% and cake has c\text{\textsubscript{w}} values above 80%.

REFERENCES
The use of ground-penetrating radar to develop a track substructure characterisation model

D J Vorster, P J Gräbe

The characterisation of track substructure is an integral component of track maintenance and rehabilitation. Traditionally this is done by determining the geotechnical properties of the discrete track formation layers by excavating test pits and sampling the substructure layers. The development of geophysical investigation techniques such as ground-penetrating radar (GPR) allows continuous assessment of the condition of the track substructure. The research described in this paper uses GPR to develop a track substructure characterisation model and provides classifications for both the ballast and formation layers. The ballast and formation were classified into four classes, namely very good, good, moderate and poor. When applying the model to two sections of track (with generally good and poor quality) 82% and 100% of classes had only one class difference compared to the traditional characterisation tests for the formation and ballast layers respectively. The GPR track substructure characterisation model that was developed therefore compares well with traditional characterisation techniques and will result in significant cost and time reduction. The characterisation of the track using GPR provides a continuous classification and enhances the accuracy of the data on which a maintenance engineer can base decisions.

INTRODUCTION

The aim of a railway substructure investigation is to characterise the track substructure based on its geotechnical properties. Care should be taken when conducting a substructure investigation to establish the correct remedial action (Brough et al. 2003). Traditionally, for a typical rail line investigation, a linear investigation is conducted with in situ testing at 200 m intervals (SAICE 2010). To classify the soil and ballast, test pits are excavated for profiling and sampling. Dynamic cone penetrometer (DCP) tests are also conducted to further classify the soil and the layer depths (Clayton et al. 1995).

However, with the development of geophysical techniques during the last two decades, methods such as ground-penetrating radar (GPR) have increasingly been used for track substructure investigations (Saarenketo 2006). These techniques have several benefits. Firstly, GPR is a non-destructive geophysical technique that uses electromagnetic waves to investigate nonconductive materials (Daniels 1996). It is used in road, railway and geotechnical environments to detect concealed objects or to determine the internal structure of materials (Saarenketo 2006).

Secondly, the use of GPR to develop a track substructure characterisation model would allow a fast and cost-effective way to classify the track substructure. The capabilities of GPR allow the classification of the ballast and the formation conditions separately. This study uses typical GPR deliverables to develop a GPR track substructure characterisation model and to verify the model with traditional track investigation techniques. The GPR deliverables include the subballast and subgrade layer profiles, the GPR ballast fouling index and the GPR moisture condition of the track. The traditional testing techniques used for the verification of the model are in situ ballast and soil sampling and profiling.

The GPR survey data and characterisation was further compared with the results of specialist geotechnical techniques. These tests included the light-weight deflectometer (LWDD), remote video monitoring (RVM) and multi-depth deflectometers (MDD). These tests are not part of the scope of the development of the GPR substructure characterisation model and will be discussed in a future publication.

GROUND-PENETRATING RADAR

Over the years, the uses of GPR have been explored in various ways. The first developments in GPR railway classification were the determination of ballast layer thickness and locating mud holes and ballast pockets through field and laboratory investigations. He also lectures undergraduate and postgraduate courses in civil engineering.

Contact details:
Department of Civil Engineering
University of Pretoria
Lynnwood Road
Pretoria
0002
T: +27 12 402 2718
E: hanns.grabe@up.ac.za

JACO VORSTER is a lecturer and researcher at the University of Pretoria’s Civil Engineering Department. He completed his MEng degree in Transportation Engineering at the same institution in 2012. He enjoys research into railway engineering and has done research on track geotechnology, railway maintenance methods and instrumentation development through field and laboratory investigations. He also lectures undergraduate and postgraduate courses in civil engineering.

Contact details:
Department of Civil Engineering
University of Pretoria
Lynnwood Road
Pretoria
0002
T: +27 12 402 4723
E: jaco.vorster@up.ac.za

PROF HANNES GRÄBE, who is a professional engineer and a Fellow of the South African Institution of Civil Engineering, is passionate about railways, with experience in the fields of track technology, geotechnology, advanced laboratory testing, field investigations, maintenance models and numerical analysis of track structures. He is currently employed by the University of Pretoria as Associate Professor: Transnet Chair in Railway Engineering, where he lectures undergraduate and postgraduate courses in railway engineering. He is also responsible for railway research, as well as continued professional education in the form of short courses presented to industry.

Contact details:
Department of Civil Engineering
University of Pretoria
Lynnwood Road
Pretoria
0002
T: +27 12 402 4723
E: hanns.grabe@up.ac.za

Keywords: track substructure, characterisation, ground-penetrating radar, ballast fouling, earthworks

**Principles**

The basic principles of and factors influencing the functioning of GPR are discussed in this section. GPR originated from electromagnetic (EM) theory and uses the transmission and reflection of EM pulses into different mediums, as shown in Figure 1 (Daniels 1996). The reflected energy is displayed in waveform where the difference in amplitude shows the interfaces between wave pulses (Daniels 1996; Saarenketo 2006). A ground profile can be generated by repeating the measurements while moving the antenna across a target area with a continuous series of radar pulses (Saarenketo 2006; Sussman et al. 2003; Hyslip et al. 2003).
The material properties that influence the propagation and reflection of the radar pulses are the dielectric permittivity, the electrical conductivity and the magnetic permeability, as described by constitutive relationships (Daniels 2004; Jol 2009). Dielectric permittivity and its effect on GPR signal velocity in materials is the most important property that affects GPR survey results. It is important to know how this value relates to the medium being tested in order to calculate the depth of the target. The pulse travels slower through material with a high dielectric permittivity and faster through material with a low dielectric permittivity (Saarenketo 2006). The formulae used to calculate the dielectric permittivity are shown in Equation 1 and Equation 2.

\[ \nu = \frac{c}{\sqrt{E_r}} \]  
\[ s = \frac{vt}{2} \]

where:
\( \nu \) = wave propagation speed (m/ns)  
\( c \) = speed of light in a vacuum (0.3 m/ns)  
\( s \) = interface depth (m) from the surface of the medium  
\( t \) = two-way travel time from the soil surface to the interface depth (ns = 10^-9)  
\( E_r \) = relative dielectric permittivity of the soil

The dielectric permittivity used in GPR surveys ranges from 1 for air to 81 for free polar water (Saarenketo 2006). The dielectric permittivity of ballast and subgrade material varies between 3.0 and 38.5 for clean dry ballast, repeatable results can be obtained (Clark et al 2001). However, the effect of moisture on dielectric permittivity is greater than that of the material type itself (Clark et al 2001; Leng & Al-Qadi 2010). Information on the material quality can be determined by using the time domain GPR data and converting it to the frequency domain with a Fourier transform (Silvast et al 2006). This is used to determine the ballast fouling and the moisture condition and is calibrated with field measurements (Silvast et al 2010).

**Equipment and data processing**

The above process requires specific equipment and tools. A typical railway track GPR survey is carried out with a 400 MHz air-coupled antennae system that can penetrate up to 1 m and can be performed at rail operating speeds. With the suspension of the antennae 300 mm above the surface of the ballast, repeatable results can be obtained regardless of the changes in surface height (Saarenketo 2006). By using a multiple antennae configuration, a cross-section of track can be obtained (Morey 1998; Clark et al 2001). A typical sampling density (10 scans/m) is used for railways (Saarenketo 2006). Figure 2 shows a typical layout of this type of instrumentation.

A range of accessories can be used in conjunction with GPR systems. However, digital video and global positioning systems (GPS) are most commonly used. A sample drilling rig (Saarenketo 2006), as well as infrared thermography (Clark et al 2003, 2004), can also be integrated with GPR data. Smekal et al (2003) used a track loading vehicle in conjunction with GPR results. Digital video recordings allow the interpreter to evaluate the surroundings of the GPR survey after the initial survey (Clark et al 2004). By using these accessories together with the GPR survey data, a comprehensive understanding of the site can be achieved (Saarenketo 2006).

GPR processing software is used to detect layer interfaces and individual objects within the ground from GPR data and to transform the data from the time domain into depth scale (Saarenketo 2006). Accurate estimates of layer dielectric values are important for GPR data processing. Traditionally, dielectric values are back-calculated from reference sampling. Another method in use is the surface reflection method (Maser & Scullion 1991). If the dielectric values are not available for a GPR survey, general dielectric values can be used or calculated from laboratory tests. These values are required for the successful interpretation of the railway structure, defining the substructure layers, ballast fouling and determining the moisture content.

GPR survey data contain reflections from various components within the entire surveyed structure. Therefore, to interpret GPR survey data, a thorough understanding of the surveyed structure is needed. In some cases reflection from components next to the track may influence the data. The interpreter should therefore interpret the main components first. It should also be noted that GPR

---

**Figure 4 Flow diagram to highlight the processes used to develop the GPR substructure characterisation model**

\[ \text{Equation 1} \]
\[ \text{Equation 2} \]
survey data could be linked to other survey methods in order to get a better understanding of the GPR data itself (Clark et al 2003, 2004; Saarenketo 2006).

Results
The use of GPR for railway track investigations has advanced over the years to deliver the following results:
- Substructure layer thicknesses up to 1.5 m deep
- A ballast fouling index
- The relative moisture condition of the substructure.

Typical results from a GPR survey are shown in Figure 3. Video, GPS and other interpretation results can also be added and used for post-processing and validation of the GPR results.

GPR RAILWAY LINE SURVEY
To develop a track substructure investigation tool or model from GPR data, a comparative study was conducted between GPR and other railway track substructure investigation tools. Laboratory tests as well as a field survey were conducted in the present study. Two test sections with different foundation properties were selected for the study. A flow diagram of the processes followed to characterise the track substructure using GPR is shown in Figure 4. The tasks involved in each step of the survey and characterisation process are detailed below.

In situ test selection
Standard geotechnical field test results were compared to the GPR survey results to complete the comparative study. The tests were done on both test sections of track at the sites that were identified from the GPR results. Six test sites were investigated at each of the two test sections. The sampling and profiling of the layerworks provided the following deliverables:
- Ballast fouling index
- California Bearing Ratio (CBR)
- Foundation indicators (FI)
- In situ moisture content.

Site selection
The two test sections for this study, each with unique substructure properties, were selected from the South African rail network. This included a section on the heavy-haul coal export line near Vryheid between Komvoorhoogte and Nhlazatshe (km 50 to km 70). This was considered as a good section of track because the substructure was reconstructed in 2005 according to Transnet’s S410 earthworks specification. A second site was selected on the railway line near Rustenburg between Northam and Thabazimbi (km 203 to km 223). This was considered as a track with a poor foundation since the formation consisted predominantly of clay and was constructed in 1934. The test sections are shown in Figure 5. The abbreviations KN (Komvoorhoogte–Nhlazatshe) and NT (Northam–Thabazimbe) will be used in the remainder of this report. Having two sections with different quality track allowed a range of track substructure conditions to be identified for the characterisation model.

Laboratory testing
Laboratory tests at the Civil Engineering Laboratory, University of Pretoria, using GPR equipment were conducted on ballast...
The ballast boxes used for the GPR laboratory tests are shown in Figure 6.

Three tests were conducted on each box with a 400 MHz antenna (Figure 7): a static test 300 mm above the surface of the material, a lifting test from 300 mm and a moving test across the width of the box. The first two tests were conducted to remove background noise. The third was used to determine the dielectric permittivity of the ballast material with the addition of moisture (water) to the boxes. The moisture content of the ballast in the boxes with no fouling material could only be increased to 10% due to drainage of the boxes. The moisture content in the boxes with fouled material was increased to 27.0% and 33.5% for the quartzite and dolerite respectively due to moisture retention by the clayey material. The moisture levels were verified by piezometers.

The main deliverable from the laboratory study was the dielectric permittivity of the ballast materials. This was done by analysis of the GPR signal obtained with the moving test. A typical cross-section of a test is shown in Figure 8. The colour scheme used for the interpretation shows the largest reflections in red or white and the smallest reflections in purple. The green lines are areas between large and small reflections. The processed data from each of the boxes was visually inspected to determine the travel time of the signal between the antenna and the layer interfaces. This process was repeated three times to obtain an average. A clear distinction was observed between the ballast rock and the concrete floor on which the ballast boxes were placed.

From the ballast GPR laboratory test results, it was found that the dielectric permittivity of the dolerite was higher than that of the quartzite. It was also determined that once the ballast material was moist, the dielectric permittivity would increase by about 0.8 to 2.2. Furthermore, the dielectric permittivity of the quartzite ballast was influenced more significantly by the fouling than the influence of the fouling on the dielectric permittivity of the dolerite.

**Field testing**

A full GPR line survey was conducted on both test sections in South Africa. The tests were conducted using a road-rail vehicle as shown in Figure 2. The GPR equipment was attached to the vehicle before testing. The equipment used for the GPR survey consisted of the following:

- GSSI SIR-20 amplifier
- Two GSSI 400 MHz antennae (model 5103A)
- Three industrial cameras (Firewire camera, resolution 1 024 x 768)
- Railway Doctor (RD) Camlink software
- GPS system (used with RD Camlink)
- Distance measurement device.

The GPR antennae were attached 1 m behind the vehicle with one antenna 300 mm above the edge of the sleeper and the other 300 mm above the centre of the sleeper. The three video cameras and the GPS antenna were placed on the roof of the vehicle. One camera faced directly in front of the vehicle and the other two covered the two adjacent sides inclined to the front. The placement of the cameras therefore created a panoramic view of the track and its surroundings. In
addition, the distance measurement device was attached to one of the rear wheels and was calibrated according to a known distance.

The GPR survey was carried out at a speed of 40 km/h. While the vehicle was travelling, datum points were taken at each kilometre interval. This was done to establish the corrections required for long and short chainages (i.e. marked kilometre posts covering distances of not exactly 1 000 m). At the end of the section, if possible, the vehicle was turned around facing the opposite direction. If this was not possible, the antenna on the left-hand side was moved to the right-hand side, 300 mm above the edge of the sleeper. The position of the antennae is shown in Figure 2. The vehicle then travelled back to the starting location at 40 km/h. After completion, the equipment was removed from the road-rail vehicle, packed away and the vehicle was manoeuvred off the tracks.

The GPR data, GPS data, the video of the track, as well as the distance measurements, were then processed by the GPR contractor. The processed data included the following:

- GPR data interpretation
- GPR layer thickness interpretation
- GPR ballast fouling index
- Relative moisture condition with depth
- Survey inventory
- Video feed
- GPS placement.

**Ballast fouling and soil classification**

Ballast fouling is determined by the amount of fouling material and is usually expressed as the percentage fines contained in the ballast (Selig & Waters 1994). The ballast fouling index can be determined by doing a grading analysis or by using specific methods recommended by the various railroads in the world (Arangie 1997). South Africa uses a method proposed by Pretorius (1993), based on the percentage material passing the 19.0 mm, 6.7 mm, 1.18 mm and 0.15 mm sieves.

Substructure layerworks is classified by excavating a test pit where samples are taken from each discernable layer that was profiled. The profiling is done according to the method proposed by Jennings et al (1973), where the layers are described in terms of moisture condition, colour, consistency, structure, soil type and origin. The gradation, soil type, California Bearing Ratio (CBR) and foundation indicators (FI) are obtained from laboratory tests on disturbed samples (Clayton et al 1995). The gradation and other deliverables can then be used to classify the track subballast and subgrade layers according to an earthworks specification. The specification proposed by Transnet (2006) is shown in Table 1.

**GPR SUBSTRUCTURE CHARACTERISATION MODEL**

The standard results from the GPR line survey were used to develop a GPR substructure characterisation model. These included the subballast and subgrade layer profiles, the GPR ballast fouling (GBF) index and the GPR moisture condition. The results were first analysed and interpreted to determine the potential of each deliverable to be used as a characterisation parameter. The data were then divided into ranges that best suited the conditions of the two test sections. The exact classification ranges were then obtained from these ranges. The individual steps followed through each of these procedures are discussed in this section.

<table>
<thead>
<tr>
<th>Layer</th>
<th>SAR index</th>
<th>Min. grading modulus</th>
<th>% By mass passing sieve (mm)</th>
<th>PI</th>
<th>Max. CBR swell %</th>
<th>Min. compaction % of modified AASHTO density</th>
<th>Min. strength after compaction CBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSB</td>
<td>&lt;50</td>
<td>2.0</td>
<td>100</td>
<td>60–85</td>
<td>20–50</td>
<td>5–15</td>
<td>3–10 98 (1.5-3 MPa)</td>
</tr>
<tr>
<td>SB</td>
<td>&lt;80</td>
<td>1.8</td>
<td>100</td>
<td>70–100</td>
<td>20–60</td>
<td>5–20</td>
<td>3–10 95 (1.5-3 MPa)</td>
</tr>
<tr>
<td>A</td>
<td>&lt;110</td>
<td>1</td>
<td>&lt;40</td>
<td>&lt;12</td>
<td>95 100*</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>&lt;155</td>
<td>0.5</td>
<td>&lt;70</td>
<td>&lt;17</td>
<td>93 98*</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Bulk earthworks</td>
<td></td>
<td></td>
<td>&lt;25</td>
<td>2</td>
<td>90 95*</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

* These densities apply to non-cohesive soils

(o) Strengths in brackets apply in place of CBR values where sub-ballast is stabilised

+ Increase to 45 in the absence of layer SSB unless otherwise specified (increase not normally required in dry areas)

SSB = Special Subballast
SB = Subballast

**Figure 9** Typical subballast and subgrade profile roughness

**Table 1** Earthworks specifications for subballast and subgrade (Transnet 2006)
Firstly, the interpretation of the GPR signal from the field survey, as shown in Figure 3, was exported for further analysis. This process was repeated for the subballast profile, subgrade profile, GBF index and GPR moisture condition. The reason for this is that the GPR software only allows basic GPR signal interpretation, such as depth determination, from the dielectric permittivity of the material. The GBF index and the GPR moisture condition are determined by FFT algorithms applied by the GPR survey contractor.

Next, the subballast and subgrade profile roughness values were determined using the root mean square method (RMS) as described by Equation 3. The roughness of the track will allow classification of the subballast and subgrade due to differential settlement of the track.

\[
R^2 = \left( \sum \frac{d_i^2}{n} \right) \frac{1}{n}
\]

where:
- \( R^2 \) = roughness, route mean square calculation or variance
- \( n \) = number of measurements in the length of track under consideration
- \( d_i \) = difference between the elevation of the point being measured and mean filtered elevation

The roughness calculation exponentially increases or decreases the magnitude of the results, and the difference between small and larger values then becomes more pronounced. Different lengths for the determination of the profile roughness were considered (100 m, 200 m, 300 m and 400 m), and it was found that the 200 m length provided acceptable results for the classification as shown in Figure 9. The roughness values had magnitudes of thousands and were therefore divided by a factor of a thousand for the classification.

The average of the GBF index results of the GPR survey across the width of the track was calculated, allowing easier interpretation, as shown in Figure 10. The GPR moisture condition was determined by first averaging moisture condition with depth, as the total moisture condition of the track influences the GPR signal. Thereafter, the RMS method was used over a 200 m length to calculate the GPR moisture index. This was done to create more discernable values for the substructure classification. These values were also factored by a thousand as shown in Figure 11.

**Characterisation model**

To develop the characterisation model, the results from the GPR field survey were analysed and grouped into four classes. The classes were, from 1 to 4, very good, good, moderate and poor. The different GPR deliverables were each classified separately and then combined in the characterisation model.

The same process was followed for the determination of the classification ranges of the GPR deliverables. The process required the average value of each section to be determined, assuming that the KN test section was good and the NT section was poor. The different GPR deliverables were each classified separately and then combined in the characterisation model.

The average of the total data set was then determined, which provided an indication of the magnitude of the moderate value. These three values were then adjusted to obtain clear ranges for the four classes. The classification ranges for the subballast surface roughness, subgrade surface roughness, GBF index average and GPR moisture index are given in Table 2.

<table>
<thead>
<tr>
<th>Class 1</th>
<th>Class 2</th>
<th>Class 3</th>
<th>Class 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>Good</td>
<td>Moderate</td>
<td>Poor</td>
</tr>
<tr>
<td>Subballast surface roughness (mm$^2 \times 1000$)</td>
<td>0 to 0.5</td>
<td>0.5 to 1</td>
<td>1 to 1.5</td>
</tr>
<tr>
<td>Subgrade surface roughness (mm$^2 \times 1000$)</td>
<td>0 to 2</td>
<td>2 to 4</td>
<td>4 to 6</td>
</tr>
<tr>
<td>GBF index average</td>
<td>0 to 35</td>
<td>35 to 70</td>
<td>70 to 105</td>
</tr>
<tr>
<td>GPR moisture index $R^2$</td>
<td>0 to 40</td>
<td>40 to 80</td>
<td>80 to 120</td>
</tr>
</tbody>
</table>

**GPR data interpretation**

Finally, the interpretation of the GPR signal from the field survey, as shown in Figure 3, was exported for further analysis. This process was repeated for the subballast profile, subgrade profile, GBF index and GPR moisture condition. The reason for this is that the GPR software only allows basic GPR signal interpretation, such as depth determination, from the dielectric permittivity of the material. The GBF index and the GPR moisture condition are determined by FFT algorithms applied by the GPR survey contractor.
classification for the track formation. The GBF index and the GPR moisture index were also combined for the classification of the ballast condition, as the GPR moisture readings were influenced mostly by the ballast layer. When combining two classifications, the most critical class was used for the final classification. Combining the classifications then provided two main classification criteria, one for the classification of the ballast material and the other for the classification of the track foundation. In doing so, rehabilitation or maintenance of the different components can be planned separately.

It is realised that the characterisation model incorporates a significant amount of averaging to arrive at the final classifications. This is required for a broad and practical classification. However, the un-averaged data plots are ideal for identifying critical and problem areas for future or more detailed investigations.

**APPLICATION OF GPR CHARACTERISATION MODEL**

The classification of the two test sections was done in accordance with the GPR characterisation model. The classification results were further simplified for future maintenance planning. This was done by limiting the classified section length to 2 km unless there was a bridge, tunnel or any other discernable structure that would influence the GPR signal. The classifications of the KN test section and the NT test section are shown in Figure 12 and Figure 13 respectively.

By applying the simplified classification, which takes the length of the classified sections into account, the final classification percentages of the two test sections were determined and are shown in Figure 14.

The percentages of the different classes for the two test sections highlight the difference in quality between them. Before the simplification of the classification, the KN test section had 75.4% of the ballast and 70.8% of the track formation in the good and very good classification ranges. On the NT test section, 15.6% of the ballast and 43.0% of the track formation was good to very good. The change in classification percentages after the simplification showed a decrease in good to very good sections of the track. For the KN test section it decreased to 62.0% for the ballast classification and 59.2% for the formation classification. For the NT test section the good to very good track sections decreased to 0% and 25.0% for the ballast and formation respectively.

**VERIFICATION OF THE MODEL**

The classification of the track with the GPR substructure characterisation model and the results of the traditional in situ testing were compared for verification of the model. The in situ soil tests were classified in accordance with the Transnet (2006) S410 earthworks specification and the ballast fouling index according to the method proposed by Arangie (1997). The classification is shown in Table 3.

The traditional earthworks classification could only be used at the 10 test sites where
EVALUATION OF THE MODEL

To fully investigate the track substructure condition for broad classification purposes by using in situ tests, a test pit has to be excavated at least every 40 km and therefore 200 test pits would have to be excavated. The tests carried out at each test pit would be as follows:

- Profiling of the test pit
- Two foundation indicator samples
- Two CBR samples that include Mod AASHTO testing
- Ballast fouling sample.

In situ tests would also be required for the GPR substructure characterisation to calibrate the GPR results. Each classification section would need at least one test pit per section and one every 2 km in the section. From this it was determined that 22 test pits would be required for the GPR survey.

For the most effective substructure characterisation, it is recommended that GPR substructure characterisation be used in conjunction with traditional in situ classification. The aim would then be to obtain a continuous characterisation of the substructure from the GPR characterisation model and a more in-depth classification from the in situ testing at problematic sections. This will provide engineers making decisions on track substructure rehabilitation with a complete and comprehensive overview and understanding of the condition of the track. A complete track condition maintenance plan can be developed by using the substructure characterisation data from GPR results in conjunction with the continuous geometric classification of the track roughness.

CONCLUSION

This study proves that it is possible to develop an effective and accurate GPR substructure characterisation model. This was done by using typical GPR survey deliverables, namely layer thickness, GPR ballast fouling and GPR moisture content. These deliverables were further analysed for ease of classification in the model. The subballast and subgrade profile roughness values were used for foundation classification and the GPR index and the GPR moisture index roughness values were used for the ballast classification. The GPR substructure classification model was evaluated by comparing its classification with the classification obtained by using in situ investigation techniques.

The comparison of the GPR substructure classification model with typical in situ classification techniques showed good correlation. Only 18% of the formation and earthworks classifications differed by a maximum of two classes, while 100% of the ballast classification differed by one class or less.

In conclusion, the use of GPR in a substructure classification model is not only possible but also provides a continuous characterisation compared to the fragmented nature of a traditional in situ railway track investigation. The most effective use of the GPR substructure characterisation model is in conjunction with in situ investigation techniques and track surface geometry. The GPR model provides continuous characterisation of the substructure, whereas in situ tests will provide a more in-depth classification at problematic areas.

ACKNOWLEDGEMENTS

The following organisations are gratefully acknowledged for their invaluable contributions to this research:

- Roadscanners for carrying out the GPR survey and initial data interpretation.
- Aveng Manufacturing Lennings Rail Services for providing the road-rail vehicle and field testing, and Aurecon assisting with the arrangements.
- Geostrada for the soil and laboratory testing.
- Transnet Freight Rail for sponsoring the research carried out by the Chair in Railway Engineering and for allowing access to their network for the field tests.
- The University of Pretoria laboratory and staff for assistance with laboratory and field testing.

REFERENCES

Berggren, E G, Smekal, A & Silvast, M 2006. Monitoring and substructure condition assessment of existing railway lines for upgrading to higher axle loads and speeds. Paper presented at the 7th World Congress on Railway Research, 4–8 June, Montréal, Canada.


Saarenketo, T 2006. *Electrical properties of road materials and subgrade soils and the use of ground penetrating radar in traffic infrastructure surveys*. Department of Geosciences, University of Oulu: Oulu University Press.


INTRODUCTION

Since the advent of multi-party democracy in Malawi in 1994, the construction industry has experienced a major growth in the number of participants, especially contractors. The construction industry in Malawi, as in most countries, has made a significant contribution to the growth of the economy through infrastructure development and job creation, apart from the multiplier effects on other sectors of the economy.

The contractors and consultants have varying experience, capabilities and management skills, all of which have a major impact on the completion times of construction projects. The growth in the number of these players in the industry has not seen a corresponding improvement in the timely delivery of projects, although with more contractors and consultants, there is increased competition among themselves and the clients have a greater variety of service providers from which to select. The construction industry in Malawi is now at a stage where most contractors, both emerging as well as long established, can hardly deliver their projects on schedule, not to mention failing to perform all together. This failure to deliver road projects on time annoys both clients and road users who expect to benefit from the completed roads. This state of affairs is undesirable to both the contractors and clients, as it is costly for both parties and has the potential to trigger disputes whose resolution is time-consuming and expensive.

Construction delays are often responsible for turning profitable projects into loss-making ventures (Sweis et al 2008). While delays are endemic in the construction industry, this need not be so. The consequences of these delays, which include cost overruns, loss of profits, increased overheads, stress, acrimony between parties, litigation and loss of opportunities because resources are tied up in delayed projects, warrant a study of this nature. The first step in correcting this anomaly is to identify the root causes of the delays so that corrective measures can be devised. Project managers will then be in a better position to monitor and control their plans. Projects that are on track give implementers satisfaction and stress-free hours of work, as they know that they are in control of their projects. All stakeholder (contractors, consultants, clients and others) should benefit from the findings of this study.

It appears that it has become the norm rather than the exception for road construction projects in Malawi to experience delay. Ellis and Thomas (2003) argue that a significant annoyance to the public is when important projects are not completed in a timely manner and when the actual construction work takes longer than necessary, thereby prolonging the inconvenience. Apart from inconveniencing road users, various studies (Al-Khalil & Al-Ghaffly 1999; Ahmed et al 2002; Aibinu & Jagboro 2002; Assaf & Al-Hejjji 2006) have shown that a delay usually leads to cost overruns and disputes, and negatively impacts the economic feasibility of such projects. Projects that are delayed are not just costly for the contractor and client,
but also for other stakeholders. The cost of deprived benefits to the users, which by definition is higher than the cost of the project, is a major result of construction delays (Malotaux 2009).

The objectives of this paper are to document:
- The range of identified causes of delay in completing road construction projects in Malawi
- The most important causes of delay in road construction projects in Malawi
- Identified differences in perception of contractors, consultants and clients regarding causes of delay in delivering projects by the intended completion date.

This paper thus answers the following three questions:
- What are the general causes of delay in road construction projects in Malawi?
- What are the most important causes of delay in road construction projects in Malawi?
- What are the perceptions of contractors, consultants and clients regarding the causes of delay in delivering these projects by the intended completion date?

THEORETICAL FRAMEWORK

Theories on the causes of delay in construction projects and methods used to develop the theories have been presented by various researchers (Mansfield et al 1994; Assaf & Al-Hejji 2006; Odeh & Battaineh 2002; Aibinu & Jagboro 2002; Tumi et al 2009).

Most of these researchers based their research on reviews of publicly available literature. Others researchers (Mansfield et al 1994; Othman et al 2006; Aibinu & Jagboro 2002) based their theories on a combination of literature reviews (publicly available literature) and project files of completed projects or projects in progress to determine construction time performance. The third procedure (Al-Tahtabai 2002; Assaf & Al-Hejji 2006) was based on field visits to construction sites where discussions were held with parties in the construction industry.

After the literature review and, in some cases discussions with parties on project sites, the identified causes of construction delay were tabulated into a questionnaire which was completed by consultants, contractors and clients. Seventy-two causes of delay were identified, and they were divided into six categories related to consultants, clients, contractors, projects, resources and external causes. Analysis of the completed questionnaires indicated that the top three causes of delay are improper planning or lack of planning (Chan & Kumaraswamy 1997; Odeh & Battaineh 2002; Al-Tahtabai 2002; Tumi et al 2009; Mansfield et al 1994; Assaf & Al-Hejji 2006; Ogunlana et al 1996), variations (Chan & Kumaraswamy 1997; Mansfield et al 1994; Assaf & Al-Hejji 2006; Sullivan & Harris 1985; Ellis & Thomas 2003; Ahmed et al 2002) and changed site (ground) conditions (Chan & Kumaraswamy 1997; Acharya et al 2006; Mansfield et al 1994; Sullivan & Harris 1985; Ellis & Thomas 2003; Vidalis & Najafi 2002). The information obtained from these studies form an integral part of the questionnaire used for the present study.

The Relative Importance Index (RII) (Equation 1) is a favourite method for ranking causes of delay (Chan & Kumaraswamy 1997; Aibinu & Jagboro 2002; Odeh & Battaineh 2002; Al-Tahtabai 2002). It is used to rank the different causes of delays from the perspective of clients, consultants, contractors and other stakeholders.

$$RII = \frac{\sum_{i=1}^{5} WiXi}{\sum_{i=1}^{5} Xi}$$  \hspace{1cm} (1)

Where:
- $i =$ response category index for 1 (never), 2 (rare), 3 (occasional), 4 (frequent) and 5 (continual)
- $W =$ the weight assigned to the $i$th response
- = 0, 1, 2, 3, 4 and 5 respectively
- $Xi =$ frequency of the $i$th response given as a percentage of the total responses for each cause.

The indices for the causes are ranked for each group. The cause with the highest index is the most important, while that with smallest number is the least important.

Spearman's rank correlation is a relationship measure among different parties or factors and the strength and direction of the relationship (Assaf & Al-Hejji 2006). This study uses Spearman's rank correlation to show the level of agreement between any two parties (Equation 2).

$$r_s = 1 - \frac{6\sum d^2}{n(n^2 - n)}$$  \hspace{1cm} (2)

Where:
- $r_s =$ Spearman's rank correlation coefficient
- $d =$ the difference in ranking between any two parties
- $n =$ the number of factors

The correlation coefficient varies between +1 and –1, where +1 implies a perfect positive relationship (agreement), while –1 results from a perfect negative relationship (disagreement). Sample estimates of correlation close to unity in magnitude imply good correlation, while values near zero indicate little or no correlation (Assaf & Al-Hejji 2006).

RESEARCH METHODOLOGY

Self-administered surveys were used and questionnaires were delivered to participants by post, e-mail and in person. Participants filled in the questionnaires in their own time without any assistance from the researcher. This approach removes any undue pressure from the respondents and gives them the freedom to fill in the questionnaire as truthfully as possible, unlike one-on-one interviews, where interviewees may be influenced by the interviewer’s attitude. The study started with a literature review, followed by identification of the survey participants. A questionnaire was developed for data collection, focused on the defined research questions. The study participants (population) comprised engineers working at the Malawi Roads Authority (client organisation), consultants (highway engineers and team leaders) who have been supervising road works and contractors (contract managers, site agents and managing directors) who have been involved in the actual construction of roads.

The client for this study is the Malawi Roads Authority, which is headquartered in Lilongwe, Malawi’s capital city. The Roads Authority is a natural choice as the client, considering that almost all road construction projects in Malawi are administered by them. They also have regional offices in the three regions of the country. Questionnaires were sent to senior managers/engineers at the headquarters as well as at regional offices.

Other questionnaires were sent to team leaders or highway engineers of consultants who have been involved in the design and supervision of contracts administered by the Roads Authority. The third set of questionnaires was sent to managing directors, contractors managers and site agents of contractors who have been involved in roads projects under the Roads Authority.

Of the 29 questionnaires distributed to the Roads Authority, 13 were returned (response rate of 44.8%). Since the Malawi construction industry is quite small, there are also a few consulting firms that are involved in the supervision of projects administered by the Roads Authority. Of the 27 questionnaires that were distributed to them, 12 were returned (44% response rate). For contractors, those with more than two years’ experience working with the Roads Authority were randomly sampled since there are a large number of these stakeholders – 43 questionnaires were sent out and 20 were returned (response rate of 46.5%).

A combination of three methods was used to analyse the data:
- Relative Importance Index (RII)
- Spearman’s rank correlation
- Probability values (p-values).
Table 1: Study results

<table>
<thead>
<tr>
<th>Delay factors</th>
<th>Client RII</th>
<th></th>
<th>Rank</th>
<th>Consultants RII</th>
<th></th>
<th>Rank</th>
<th>Contractors RII</th>
<th></th>
<th>Rank</th>
<th>Average RII</th>
<th></th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete drawings/specifications</td>
<td>0.615</td>
<td>26</td>
<td>0.568</td>
<td>43</td>
<td>0.461</td>
<td>53</td>
<td>0.548</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design errors and omissions</td>
<td>0.615</td>
<td>27</td>
<td>0.458</td>
<td>62</td>
<td>0.474</td>
<td>52</td>
<td>0.516</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excessive extra works</td>
<td>0.673</td>
<td>16</td>
<td>0.604</td>
<td>30</td>
<td>0.450</td>
<td>54</td>
<td>0.576</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inadequate design team experience</td>
<td>0.596</td>
<td>31</td>
<td>0.500</td>
<td>57</td>
<td>0.500</td>
<td>47</td>
<td>0.532</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delays in producing design documents</td>
<td>0.673</td>
<td>16</td>
<td>0.542</td>
<td>49</td>
<td>0.538</td>
<td>39</td>
<td>0.584</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excessive variations in quantities</td>
<td>0.654</td>
<td>20</td>
<td>0.625</td>
<td>27</td>
<td>0.500</td>
<td>47</td>
<td>0.593</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rework due to wrong drawings</td>
<td>0.346</td>
<td>65</td>
<td>0.417</td>
<td>66</td>
<td>0.250</td>
<td>71</td>
<td>0.338</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient data collection and survey before design</td>
<td>0.538</td>
<td>42</td>
<td>0.604</td>
<td>30</td>
<td>0.550</td>
<td>33</td>
<td>0.564</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow response</td>
<td>0.558</td>
<td>38</td>
<td>0.568</td>
<td>43</td>
<td>0.605</td>
<td>20</td>
<td>0.577</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow decision-making</td>
<td>0.577</td>
<td>34</td>
<td>0.500</td>
<td>57</td>
<td>0.625</td>
<td>13</td>
<td>0.567</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long period for approval of tests and inspections</td>
<td>0.442</td>
<td>58</td>
<td>0.417</td>
<td>66</td>
<td>0.513</td>
<td>46</td>
<td>0.457</td>
<td>63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unfamiliarity with or lack of knowledge by the consultant’s supervision staff regarding new construction methods, materials and techniques</td>
<td>0.481</td>
<td>52</td>
<td>0.438</td>
<td>65</td>
<td>0.525</td>
<td>42</td>
<td>0.481</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of application of construction management tools and techniques by consultant’s project and site staff</td>
<td>0.500</td>
<td>48</td>
<td>0.563</td>
<td>45</td>
<td>0.375</td>
<td>64</td>
<td>0.479</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflicts between drawings and specifications</td>
<td>0.365</td>
<td>63</td>
<td>0.500</td>
<td>57</td>
<td>0.250</td>
<td>71</td>
<td>0.372</td>
<td>67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequent design changes requested by client during construction</td>
<td>0.417</td>
<td>61</td>
<td>0.455</td>
<td>64</td>
<td>0.450</td>
<td>54</td>
<td>0.441</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inaccurate initial project scope estimate</td>
<td>0.538</td>
<td>42</td>
<td>0.625</td>
<td>27</td>
<td>0.553</td>
<td>32</td>
<td>0.572</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow payment procedures adopted by client in making progress payments</td>
<td>0.654</td>
<td>20</td>
<td>0.792</td>
<td>3</td>
<td>0.813</td>
<td>3</td>
<td>0.753</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrealistic time estimation</td>
<td>0.596</td>
<td>31</td>
<td>0.667</td>
<td>22</td>
<td>0.638</td>
<td>10</td>
<td>0.634</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive bureaucracy at client’s offices</td>
<td>0.500</td>
<td>48</td>
<td>0.771</td>
<td>6</td>
<td>0.605</td>
<td>21</td>
<td>0.625</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slow decision-making process by client’s departments</td>
<td>0.481</td>
<td>52</td>
<td>0.688</td>
<td>16</td>
<td>0.625</td>
<td>13</td>
<td>0.398</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inefficient flow of information from client’s departments</td>
<td>0.423</td>
<td>60</td>
<td>0.583</td>
<td>38</td>
<td>0.525</td>
<td>42</td>
<td>0.510</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No or small time extensions associated with change orders initiated by client</td>
<td>0.365</td>
<td>63</td>
<td>0.614</td>
<td>29</td>
<td>0.625</td>
<td>13</td>
<td>0.535</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inefficient pre-qualification procedures by client, which result in the selection of incompetent contractors</td>
<td>0.346</td>
<td>65</td>
<td>0.604</td>
<td>30</td>
<td>0.638</td>
<td>10</td>
<td>0.529</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understaffed client’s project and site personnel</td>
<td>0.429</td>
<td>59</td>
<td>0.563</td>
<td>45</td>
<td>0.550</td>
<td>33</td>
<td>0.514</td>
<td>51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor communication and coordination by client and other parties</td>
<td>0.308</td>
<td>71</td>
<td>0.583</td>
<td>38</td>
<td>0.525</td>
<td>42</td>
<td>0.472</td>
<td>61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delays in work approval</td>
<td>0.346</td>
<td>67</td>
<td>0.688</td>
<td>16</td>
<td>0.575</td>
<td>25</td>
<td>0.536</td>
<td>44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client-initiated variations</td>
<td>0.481</td>
<td>52</td>
<td>0.591</td>
<td>37</td>
<td>0.487</td>
<td>50</td>
<td>0.520</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insufficient contractor cash flow/difficulties in financing projects</td>
<td>0.827</td>
<td>3</td>
<td>0.813</td>
<td>1</td>
<td>0.825</td>
<td>2</td>
<td>0.822</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor qualifications and inadequate experience of contractor’s supervisors</td>
<td>0.731</td>
<td>9</td>
<td>0.729</td>
<td>12</td>
<td>0.563</td>
<td>29</td>
<td>0.674</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ineffective planning and scheduling of project</td>
<td>0.731</td>
<td>9</td>
<td>0.750</td>
<td>7</td>
<td>0.563</td>
<td>29</td>
<td>0.681</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment allocation problems</td>
<td>0.654</td>
<td>20</td>
<td>0.708</td>
<td>14</td>
<td>0.600</td>
<td>22</td>
<td>0.654</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Materials management problems</td>
<td>0.615</td>
<td>28</td>
<td>0.750</td>
<td>7</td>
<td>0.563</td>
<td>29</td>
<td>0.643</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Misinterpretation of drawings and specifications</td>
<td>0.538</td>
<td>42</td>
<td>0.458</td>
<td>62</td>
<td>0.363</td>
<td>65</td>
<td>0.453</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rework due to errors during construction</td>
<td>0.558</td>
<td>38</td>
<td>0.521</td>
<td>54</td>
<td>0.350</td>
<td>66</td>
<td>0.476</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor communication and coordination with other parties</td>
<td>0.500</td>
<td>48</td>
<td>0.521</td>
<td>54</td>
<td>0.434</td>
<td>42</td>
<td>0.485</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor contractor’s site management and supervision</td>
<td>0.615</td>
<td>28</td>
<td>0.688</td>
<td>16</td>
<td>0.538</td>
<td>39</td>
<td>0.614</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delay in site mobilisation</td>
<td>0.692</td>
<td>12</td>
<td>0.729</td>
<td>13</td>
<td>0.650</td>
<td>9</td>
<td>0.690</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conflict between/with contractor and other parties (consultant and client)</td>
<td>0.538</td>
<td>42</td>
<td>0.479</td>
<td>61</td>
<td>0.425</td>
<td>61</td>
<td>0.481</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improper construction methods implemented by contractor</td>
<td>0.538</td>
<td>42</td>
<td>0.500</td>
<td>57</td>
<td>0.400</td>
<td>62</td>
<td>0.479</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Late delivery of materials and equipment</td>
<td>0.673</td>
<td>16</td>
<td>0.750</td>
<td>7</td>
<td>0.613</td>
<td>18</td>
<td>0.679</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The association between the ranking of parties is verified by a hypothesis testing at 95% significance (thus $\alpha = 5\%$) (Odeh & Battaineh 2002). The p-value is the probability of observing a sample value as extreme as, or more extreme than, the value actually observed, given that the null hypothesis is true. The p-value is then compared to the significance level ($\alpha$), and on this basis the null hypothesis is either rejected or not rejected. If the p-value is less than the significance level, the null hypothesis is rejected ($p\text{-value} < \alpha$, reject null). If the p-value is greater than or equal to the significance level, the null hypothesis is not rejected ($p\text{-value} \geq \alpha$, do not reject null) (Blumberg et al. 2008).

### RESULTS

A complete set of the survey results is shown in Table 1. Analysing the data from the viewpoints of the three major stakeholder types, the following are observed:

#### Clients’ viewpoints

The top five causes of delay identified by clients are:
- **Shortage of fuel** ($RII = 0.904$)
- **Insufficient equipment** ($RII = 0.865$)
- **Insufficient contractor cash-flow/difficulties in financing projects** ($RII = 0.827$)
- **Delays in paying compensation to land owners** ($RII = 0.827$)
- **Shortage of technical personnel** ($RII = 0.813$)

#### Clients’ viewpoints

The association between the ranking of parties is verified by a hypothesis testing at 95% significance (thus $\alpha = 5\%$) (Odeh & Battaineh 2002). The p-value is the probability of observing a sample value as extreme as, or more extreme than, the value actually observed, given that the null hypothesis is true. The p-value is then compared to the significance level ($\alpha$), and on this basis the null hypothesis is either rejected or not rejected. If the p-value is less than the significance level, the null hypothesis is rejected ($p\text{-value} < \alpha$, reject null). If the p-value is greater than or equal to the significance level, the null hypothesis is not rejected ($p\text{-value} \geq \alpha$, do not reject null) (Blumberg et al. 2008).

#### RESULTS

A complete set of the survey results is shown in Table 1. Analysing the data from the viewpoints of the three major stakeholder types, the following are observed:

<table>
<thead>
<tr>
<th>Delay factors</th>
<th>Client</th>
<th>Consultants</th>
<th>Contractors</th>
<th>All parties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor procurement programming of materials</td>
<td>0.615</td>
<td>28</td>
<td>0.708</td>
<td>14</td>
</tr>
<tr>
<td>Type of project bidding and award (lowest bidder)</td>
<td>0.500</td>
<td>48</td>
<td>0.750</td>
<td>7</td>
</tr>
<tr>
<td>Ineffective delay penalties</td>
<td>0.404</td>
<td>62</td>
<td>0.542</td>
<td>49</td>
</tr>
<tr>
<td>Inadequate definition of substantial completion</td>
<td>0.327</td>
<td>69</td>
<td>0.354</td>
<td>69</td>
</tr>
<tr>
<td>Legal disputes between/with various parties</td>
<td>0.346</td>
<td>67</td>
<td>0.229</td>
<td>71</td>
</tr>
<tr>
<td>Unrealistic project construction duration as specified in the contract</td>
<td>0.558</td>
<td>38</td>
<td>0.563</td>
<td>45</td>
</tr>
<tr>
<td>No financial incentives for contractors to finish ahead of schedule</td>
<td>0.462</td>
<td>56</td>
<td>0.646</td>
<td>24</td>
</tr>
<tr>
<td>No application of construction management procedures on the part of client contributes to late detection of construction problems</td>
<td>0.462</td>
<td>56</td>
<td>0.523</td>
<td>53</td>
</tr>
<tr>
<td>Unrealistic schedule programme submitted by contractor</td>
<td>0.731</td>
<td>9</td>
<td>0.604</td>
<td>30</td>
</tr>
<tr>
<td>Contractor’s staff are not properly trained in professional construction management techniques</td>
<td>0.692</td>
<td>12</td>
<td>0.646</td>
<td>24</td>
</tr>
<tr>
<td>Poor judgement and inexperience in estimating procedures by contractor</td>
<td>0.635</td>
<td>23</td>
<td>0.583</td>
<td>38</td>
</tr>
<tr>
<td>Shortage of construction materials (bitumen, cement and steel)</td>
<td>0.750</td>
<td>8</td>
<td>0.667</td>
<td>22</td>
</tr>
<tr>
<td>Shortage of technical personnel</td>
<td>0.813</td>
<td>5</td>
<td>0.688</td>
<td>16</td>
</tr>
<tr>
<td>Insufficient equipment</td>
<td>0.865</td>
<td>2</td>
<td>0.75</td>
<td>11</td>
</tr>
<tr>
<td>Shortage of fuel</td>
<td>0.904</td>
<td>1</td>
<td>0.792</td>
<td>2</td>
</tr>
<tr>
<td>Shortage of labour</td>
<td>0.481</td>
<td>52</td>
<td>0.417</td>
<td>66</td>
</tr>
<tr>
<td>Price escalation</td>
<td>0.577</td>
<td>34</td>
<td>0.646</td>
<td>24</td>
</tr>
<tr>
<td>Low level of equipment operators’ skills</td>
<td>0.673</td>
<td>16</td>
<td>0.542</td>
<td>49</td>
</tr>
<tr>
<td>Low productivity and efficiency of equipment</td>
<td>0.596</td>
<td>31</td>
<td>0.604</td>
<td>30</td>
</tr>
<tr>
<td>Lack of high-technology mechanical equipment</td>
<td>0.692</td>
<td>12</td>
<td>0.688</td>
<td>16</td>
</tr>
<tr>
<td>Unqualified workforce</td>
<td>0.635</td>
<td>23</td>
<td>0.542</td>
<td>49</td>
</tr>
<tr>
<td>Low productivity of labour</td>
<td>0.635</td>
<td>23</td>
<td>0.604</td>
<td>30</td>
</tr>
<tr>
<td>Shortage of foreign currency (importation of materials and equipment)</td>
<td>0.808</td>
<td>6</td>
<td>0.792</td>
<td>2</td>
</tr>
<tr>
<td>Delays attributed to third-party testing of materials</td>
<td>0.558</td>
<td>38</td>
<td>0.521</td>
<td>54</td>
</tr>
<tr>
<td>Differing or unexpected geotechnical conditions during construction</td>
<td>0.577</td>
<td>34</td>
<td>0.604</td>
<td>30</td>
</tr>
<tr>
<td>Effect of rain on construction activities</td>
<td>0.577</td>
<td>34</td>
<td>0.583</td>
<td>38</td>
</tr>
<tr>
<td>Effect of hot weather on construction activities</td>
<td>0.327</td>
<td>69</td>
<td>0.292</td>
<td>70</td>
</tr>
<tr>
<td>Theft of contractor’s resources</td>
<td>0.692</td>
<td>12</td>
<td>0.583</td>
<td>38</td>
</tr>
<tr>
<td>Vandalism of works (in progress or finished)</td>
<td>0.538</td>
<td>42</td>
<td>0.563</td>
<td>45</td>
</tr>
<tr>
<td>Delay in paying compensations (land-owners)</td>
<td>0.827</td>
<td>3</td>
<td>0.688</td>
<td>16</td>
</tr>
<tr>
<td>Delay in relocating utilities</td>
<td>0.808</td>
<td>6</td>
<td>0.792</td>
<td>2</td>
</tr>
<tr>
<td>Industrial action (strike/sit-in)</td>
<td>0.250</td>
<td>72</td>
<td>0.229</td>
<td>71</td>
</tr>
</tbody>
</table>
Table 2 Correlation test of all factors among respondents

<table>
<thead>
<tr>
<th>Group</th>
<th>Client &amp; consultants</th>
<th>Client &amp; contractors</th>
<th>Consultant &amp; contractor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Correlation coefficient</td>
<td>P-value</td>
<td>Correlation coefficient</td>
</tr>
<tr>
<td>All factors</td>
<td>0.601</td>
<td>0.09</td>
<td>0.503</td>
</tr>
</tbody>
</table>

Table 3 Top ten factors that cause delay

<table>
<thead>
<tr>
<th>Cause of delay</th>
<th>All parties</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortage of fuel</td>
<td>0.853</td>
<td>Resource-related</td>
</tr>
<tr>
<td>Insufficient contractor cash-flow/difficulties in financing projects</td>
<td>0.821</td>
<td>Contractor-related</td>
</tr>
<tr>
<td>Shortage of foreign currency (importation of materials and equipment)</td>
<td>0.800</td>
<td>Resource-related</td>
</tr>
<tr>
<td>Slow payment procedures adopted by client in making progress payments</td>
<td>0.753</td>
<td>Client-related</td>
</tr>
<tr>
<td>Insufficient equipment</td>
<td>0.747</td>
<td>Resource-related</td>
</tr>
<tr>
<td>Delay in relocating utilities</td>
<td>0.716</td>
<td>Externally related</td>
</tr>
<tr>
<td>Shortage of construction materials (bitumen, cement and steel)</td>
<td>0.714</td>
<td>Resource-related</td>
</tr>
<tr>
<td>Delay in paying compensations (land owners)</td>
<td>0.705</td>
<td>Externally related</td>
</tr>
<tr>
<td>Shortage of technical personnel</td>
<td>0.704</td>
<td>Resource-related</td>
</tr>
<tr>
<td>Delay in site mobilisation</td>
<td>0.690</td>
<td>Contractor-related</td>
</tr>
</tbody>
</table>

**Consultants’ viewpoints**

The top five causes of delay identified by consultants are:
- Insufficient contractor cash-flow/difficulties in financing projects (RII = 0.813)
- Slow payment procedures adopted by the client in making progress payments (RII = 0.792)
- Shortage of fuel (RII = 0.792)
- Shortage of foreign currency (importation of materials and equipment) (RII = 0.792)
- Delay in relocating utilities (RII = 0.792).

**Contractors’ viewpoints**

The top five causes of delay identified by the contractors are:
- Shortage of fuel (RII = 0.863)
- Insufficient contractor cash-flow/difficulties in financing projects (RII = 0.825)
- Slow payment procedures adopted by the client in making progress payments (RII = 0.813)
- Shortage of foreign currency (importation of materials and equipment) (RII = 0.800)
- Price escalation (RII = 0.763).

**Spearman’s rank correlation coefficient**

Table 2 shows the values of correlation coefficients among the parties and their corresponding p-values. These values show that there is a positive correlation between client and consultant, with a correlation of 0.601 and a corresponding p-value of 0.09 (greater than the level of significance, α = 0.05), therefore there is no significant relationship between causes of delay ranked by client and consultants. Similarly, there is a positive correlation of 0.503 between client and contractors and a p-value of 0.13, an indication of insignificant relationships between causes of delay ranked by these two respondent groups. The correlation coefficient between consultants and contractors is also positive, but their corresponding p-value is greater than 0.05, denoting an insignificant relationship between causes of delay ranked by consultants and contractors.

Since all three p-values in Table 2 are greater than the significance level (α = 0.05), the null hypothesis cannot be rejected.

**All respondents’ viewpoints**

The combined views of all three parties to the survey are shown in Table 3. They show that five out of ten causes of delay are linked to shortage of resources. These results also show that the causes of delay ranked by the consultants did not contribute to the top ten causes of delay.

Some of the results in Table 3 are very similar to those obtained by researchers in South Africa, Botswana and Swaziland. An exploratory study of problems facing small-scale contractors in the North West province of South Africa conducted by Thwala and Phaladi (2009) revealed that government was not paying on time. Lack of capital and difficulty in arranging guarantees, as well as lack of technical skills, were cited as other problems facing small-scale contractors. Four listed construction companies downed tools on 23 road contracts in the Free State because of non-payment of hundreds of millions of rand (Carte 2012).

A study of the current challenges and problems facing small and medium-sized contractors in Swaziland (Thwala & Mvubu 2008) also showed that slow payment and non-payment by government after a government project has been completed is common, leading to many construction firms suffering financial ruin and bankruptcy. Just like in Malawi, contractors in Swaziland also experience inadequacy in technical and managerial skills required for project implementation. Lack of resources also hampers effective delivery of large or complex projects in Swaziland.

Delays in payment for completed works has been cited as being among the top five key factors responsible for time delays of large construction projects in Botswana by Mathumo (2012). The other factors include poor project management skills, poor planning and lack of skilled subcontractors.

Another study on delays in completion of building construction projects in the Botswana public sector by medium and large contractors reveals that poor management is the biggest cause of delays (Joseph 2004).

**CONCLUSIONS**

This study was aimed at finding the causes of delay in road construction projects in Malawi. Seventy-two causes of delay were extracted from the literature on the subject. The seventy-two causes of delay were divided into six categories related to consultants, client, contractors, projects, resources and external. A questionnaire based on these causes of delay was sent to the client, consultant and contractor representatives. The collected data were analysed using the Relative Importance Index (RII) and Spearman’s rank correlation coefficients. From this study a collective analysis of all three groups show that among the top ten causes of delay, five are related to resource shortages, two are contractor related, two are related to external factors, while one is client related. It was observed that there is no project-related or consultant-related delay factors among the top ten causes of delay, which are:

1. Shortage of fuel.
2. Insufficient contractor cash-flow/difficulties in financing projects.
3. Shortage of foreign currency for importation of materials and equipment.
4. Slow payment procedures adopted by the client in making progress payments.
5. Insufficient equipment.
6. Delay in relocating utilities.
7. Shortage of construction materials such as bitumen, cement and steel.
8. Delay in paying compensations to land owners.
9. Shortage of technical personnel.
10. Delay in site mobilisation.

RECOMMENDATIONS

On the basis of these findings, the following recommendations are made (each bullet refers to one of the top ten causes of delay listed):

- Shortage of fuel hinges on the shortage of foreign currency used for the importation of fuel and other goods not available in the country. The country must increase its exports, and at the same time reduce its imports to enable its balance of trade to be in favour of exports.

- Insufficient contractor cash-flow/difficulties in financing projects develop either as a lack of liquidity on the part of the contractor and/or client delays in making progress payments. Assaf and Al-Hejji (2006) recommend that contractors should manage their financial resources and plan cash-flow by utilising progress payment. Contractors should ensure that the advance payment is used to finance project activities. The contractor can be paid on time if a clause is introduced in the contract whereby the client is required to pay the contractor the amount certified in an interim payment certificate within seven calendar days of the date of issue of the payment certificate, as is the case with the Joint Building Contracts Committee (JBCC) Series 2000 Principal Building Agreement. The client should not award a contract to the contractor when the client does not have adequate finances to execute the project. Mansfield et al (1994) recommend that clients should ensure that adequate funds are available before projects are started, so that contractors can be paid in accordance with the contract agreement. The FIDIC Multilateral Development Bank Harmonised Edition (2010) requires the employer to give the contractor evidence that it has access to or has the funds necessary to pay the contract price. Clause 2.4 reads in part “The Employer shall submit, before the Commencement Date and thereafter within 28 days after receiving any request from the contractor, reasonable evidence that financial arrangements have been made and are being maintained which will enable the Employer to pay the Contract Price punctually (as estimated at that time) in accordance with Clause 14 [Contract Price and Payment]. Before the Employer makes any material change to his financial arrangements, the Employer shall give notice to the Contractor with detailed particulars.” Mansfield et al (1994) further recommend that comprehensive economic analysis and workable financial plans should be prepared before contracts are awarded.

- While it is common practice for contracts to include a performance guarantee clause, there should also be a payment guarantee clause so that if a duly issued payment certificate is not paid within the stipulated period, the contractor may demand his payment from the guarantor.

- In most contracts funded by development partners there is a currency split provision for paying the contractor in more than one currency. The contractor should, when tendering, assess his foreign currency requirement for importation of materials and equipment and factor it into the currency split. As for contracts funded by local resources, government must put in place policies that encourage export growth that will generate foreign currency for the country. The government should also create an economic climate that will see the country importing only those goods and services that are not locally available, thereby reducing the outflow of foreign currency.

- The inclusion of a clause in the contract, as is the case with the Joint Building Contracts Committee (JBCC) Series 2000 Principal Building Agreement (PBA) (2007), requiring that the employer (client) shall pay to the contractor the amount certified in an interim payment certificate within seven calendar days of the date of issue of the payment certificate, should considerably quicken payment procedures. Another clause should be introduced in the contract requiring the client to pay interest to the contractor for delaying his payment.

- Contractors should consider buying their own equipment from the proceeds of their contracts. There is an opportunity for investors to set up private plant and equipment hire organisations. Local investors can also invite international investors to invest in plant and equipment hire organisations, since there is a shortage of equipment in the country.

- Utility organisations should be involved at the planning stage so that there is coordination and cooperation in locating and relocating these services before construction works start. Goodrum et al (2009) recommend the establishment of utility corridors and systematic location of facilities. They further recommend avoiding the need to relocate many utility lines by collecting and mapping underground utility data that was primarily unknown, using subsurface utility engineering early in the design phase. Utility organisations should produce accurate and clear as-built drawings to provide utility location information.

- Concerted efforts should be directed towards research and development in the use of local materials. In order to encourage international contractors and investors to undertake greater foreign direct investment in such areas as materials development and production, it is necessary for the host government to relax ‘interventionist monetary policies’ and other strict economic measures (Mansfield et al 1994). This incentive will enable the local currency to find its true value in the free market, and thereby curtail excessive price fluctuations associated with imported construction materials, equipment and other plant items.

- At the planning stage, people whose properties would be affected by the construction works should be identified and compensated agreed with the property owners. After compensations have been agreed, they should be paid so that the affected people can relocate well in advance of the commencement of the construction works.

- All three parties (client, consultant and contractor), should put in place policies that will help them retain their valuable human resources, thereby reducing their high staff turnover. The parties should conduct continuous training programmes to improve the competency of personnel carrying out designs, supervision and construction at all levels, not just at the top, but all the way down to craftsmen and casual workers. There is an urgent need for offering training courses in scheduling, time and cost control, information systems, and management of human resources (Odeh & Bataine 2002). Project personnel should also be trained in critical chain project scheduling.

- Most contracts stipulate the time frame within which a contractor should mobilise. The consequences of failing to mobilise must also be detailed, and include cancellation of contract.

ACKNOWLEDGEMENTS

The invaluable contribution to this work of the Chief Executive Officer of the Roads Authority (Malawi), Engineer Paul Kulemeka, and his entire team of engineers,
as well as the contractors and consultants who participated in the survey is gratefully acknowledged.

REFERENCES


Critical normal traffic loading for flexure of skew bridges according to TMH7

A D Malan, G C van Rooyen

Different types of loading due to traffic may act on bridges. The focus of this paper is on normal traffic loading according to the South African specification TMH7 (Committee of State Road Authorities 1988). TMH7 represents the code of practice for the design of highway bridges and culverts in South Africa. The aim of the paper is to provide additional insight into the effect of normal traffic loading patterns on the flexural analysis of skew bridges. This is necessary since TMH7 does not explicitly specify application patterns for normal traffic loading. Only the intensity of normal traffic loading is specified and it should be applied to yield the most adverse effects. In this paper a set of so-called “standard application patterns” are investigated and compared with the application patterns that yield the most adverse effects. The patterns are compared for flexural effects in selected design regions of the bridge deck. Numerical experiments were performed and are presented for a typical single-span bridge deck. The results of the numerical experiments are compared as the angle of skew of the bridge deck increases.

INTRODUCTION
This paper is concerned with flexure of skew bridges under the effects of normal traffic loading according to TMH7. Normal traffic represents a formula loading consisting of the most severe arrangement of legal vehicles that is probable (Committee of State Road Authorities 1988). The incorporation of heavy vehicles, requiring abnormal load permits, is excluded from normal traffic loading. The problem statement and particular aim of the investigation are formulated in the next section.

Problem statement and aim of investigation
Bridge structures are subject to dead loads and live loads. Dead loads refer to permanent loads, e.g. the own weight of the structure. Live loads, such as live loads due to traffic, refer to loads that may act at various possible locations on the bridge deck. Without the aid of specialised software or influence surfaces, the determination of the positions where traffic loads should be applied to produce the most adverse effects is not obvious. The problem is exacerbated by the following:

- The most adverse effects are not necessarily obtained when the bridge deck is fully loaded.
- The South African specification TMH7 does not explicitly specify application patterns for live loading due to traffic.

The problem arises as to how the traffic loads should be applied to produce the most adverse effects.

Influence surfaces and design charts
For complex structures with complex loading arrangements, influence surfaces and design charts are usually employed to determine the effect in a specific region resulting from the application of a load in another region. However, influence surfaces and design charts have the following disadvantages (Hambly 1976):

- They are not always easy to use.
- They do not give the user a complete picture of the force system in the bridge deck under the effect of a particular load case.
- They cannot be used for orthotropic, cellular or beam-and-slab decks due to their different distortional and torsional characteristics.

Furthermore, published influence surfaces and design charts are not available for all types of structural geometries and support conditions in each of the regions of the bridge deck.

Load application patterns
Given the problems associated with the application of normal traffic loading described above, the aim of this paper is to answer the following questions:

- Is it possible to use a set of standard application patterns to approximate the most adverse effects of normal traffic loading?
If standard patterns are used, by how much will their results differ from the most adverse results?

How well do the standard application patterns perform when the angle of skew of the bridge deck increases in plan view?

The set of standard application patterns mentioned above will be developed and presented after TMH7’s specifications concerning normal traffic loading have been described. Thereafter, the generation of application patterns that produce the most adverse effects will be presented. Numerical experiments, in the form of qualitative, comparative finite element analyses, will be performed for a typical single-span bridge deck. The results of the standard application patterns will be compared with those of the most adverse application patterns as the angle of skew of the bridge deck increases in plan view.

NORMAL TRAFFIC LOADING

The geometries of the components of the bridge deck and normal (NA) traffic loading are presented here.

Bridge deck geometry

This investigation of critical normal traffic loading is restricted to the modelling and analysis of the bridge deck, i.e. the whole bridge structure is not considered. The carriageway, supports and notional lanes are defined as follows:

- **Carriageway**: The carriageway includes all traffic lanes and represents that part of the bridge deck where traffic loads must be applied.
- **Supports**: Supports are represented as line-type supports beneath the carriageway. Vertical displacement is restricted by the supports, i.e. the displacement perpendicular to the plane of the bridge deck.
- **Notional lane**: Notional lanes represent longitudinal strips along the carriageway, used for the application of normal traffic loading.

The bridge deck geometry, as defined above, is presented in Figure 1 for a skew bridge deck in plan view.

The notional lanes do not represent the actual traffic lanes. The width and number of notional lanes are specified by TMH7 and they are dependent only on the effective width of the carriageway.

**Definition of NA loading**

Type NA loading, as specified by TMH7 (1981, revised 1988), consists of a distributed part and a concentrated part acting in conjunction with each other. The distributed part can either be in the form of two equal and parallel line loads spaced 1.9 m apart, or in the form of a distributed load over the full width of the notional lane. The concentrated part of type NA loading can either be in the form of two equal point loads spaced 1.9 m apart, or in the form of a knife-edge load over the full width of the notional lane. The two possible application methods of type NA loading are illustrated in Figure 2.

The intensity of the distributed part of type NA loading is dependent on the effective loaded length of the part where it is applied. The concentrated axle part is dependent on the loading sequence of the relevant notional

---

**Figure 1 Bridge deck geometry**

**Figure 2 Normal traffic loading**
The intensities of type NA loading are described further in the sections below.

**Distributed load**

The distributed part of type NA loading represents a nominal distributed lane load, acting on the whole or parts of the length of any notional lane or combination of such lanes. In the longitudinal direction the loading is uniformly distributed for any continuous part of a notional lane, but the intensity may be different for separate parts:

for \( L_c \leq 36 \text{ m} \)  \( Q_a = 36 \text{ kN/m} \)  (1)

for \( L_c > 36 \text{ m} \)  \( Q_a = \frac{180}{L_c} + 6 \text{ kN/m} \)  (2)

where

- \( L_c \) = effective loaded length (m)
- \( Q_a \) = intensity of the loading (kN/m).

The specification of TMH7 provides a procedure whereby the loading intensities are based on the assumption that the total loading is dependent on the aggregate loaded length of all the notional lanes being considered (see Section 2.A.2.1 of Part 2 of TMH7). The intensities of the loading in these separate parts are not necessarily equal. The loading procedure for the distributed part of type NA loading is presented according to method A(1) of TMH7 (Committee of State Road Authorities 1988):

That part of any notional lane, which has the maximum average influence value (positive or negative as the case may be), is loaded at an intensity determined by the NA uniformly distributed loading formula for that loaded length. Thereafter, that part of the same or any other notional lane with the next highest average influence value of similar sign is loaded at an intensity such that the total load on the two loaded parts corresponds to the formula loading for a loaded length equal to the sum of the two loaded lengths.

If \( \sum_{i=1}^{p} L_i \) is the sum of all the loaded lengths up to and including the \( p^{th} \) part, the intensity of loading \( Q_{ap} \) on the \( p^{th} \) part of length \( L_p \) is defined as follows:

\[
Q_{ap} = \frac{(Q_{a} \sum_{i=1}^{p-1} L_i - \sum_{i=1}^{p-1} Q_i L_i) L_p}{\sum_{i=1}^{p} L_i}
\]

where

- \( Q_a = \) intensity of the loading for a length of \( \sum_{i=1}^{p} L_i \)
- \( Q_i = \) intensity of loading applied to previously calculated base length portion \( i \)
- \( L_i = \) dimension of any previously calculated base length portion \( i \).

In this procedure \( Q_{ap} \) reduces as \( p \) increases with no limiting value.

**Concentrated load**

The concentrated part of type NA loading is in the form of one nominal axle load per notional lane with the following intensity:

\[
\frac{144}{\sqrt{n}} \text{kN}
\]

where

- \( n = \) loading sequence number of the relevant notional lane.

For example, \( n = 1 \) for the first lane loaded with the axle load, \( n = 2 \) for the second loaded lane, etc. The concentrated axle part of type NA loading is applied in conjunction
with the distributed part and only one axle load must be applied per notional lane.

LOADING PATTERNS
Based on the definition and intensity of type NA loading presented above, two approaches regarding the application of loading are developed, namely so-called "standard patterns" and "critical patterns".

Standard patterns
The standard patterns are based on intuition and engineering judgement, i.e. how an engineer could consider applying type NA loading to yield the most adverse effect if he/she did not have access to influence surface-based software. The proposed standard patterns are illustrated in Figure 3.

In the case of the standard patterns as illustrated in Figure 3, the highest intensity loading is applied towards the bottom edge of the carriageway in the transverse direction. The concentrated axle part of type NA loading is applied towards the left, the centre and the right in the longitudinal direction respectively. A basic torsion load is also included.

When the standard patterns are used, all the loadable parts of the notional lanes are loaded, with a decreasing intensity when the effective loaded length exceeds 36 m.

The standard NA loading patterns provide a systematic procedure for loading a carriageway. It is easy to implement a programmatic procedure to generate standard NA loading patterns for a carriageway with any number of spans and any number of notional lanes.

Critical patterns
The standard NA loading patterns, as described in the previous section, load all the loadable parts of the notional lanes of a specific span. As a result, certain regions that may provide a relieving effect will also be loaded, which results in less adverse values. To avoid this, a method is presented that generates critical positive or negative loading patterns for a selected mode of failure. The critical loading patterns are based on the influence value, at a chosen location, of a load increment as it moves over the notional lanes. A short, distributed line load segment-pair is moved over all the loadable parts of the notional lanes of the carriageway in steps, as shown in Figure 4.

A critical pattern, as illustrated in Figure 4, is generated as follows:

- A region in the carriageway, as well as a mode of failure, are selected for which the critical pattern should be generated.
- The load increment segment-pair moves step-wise over all the loadable parts of the carriageway, as indicated by the arrows in Figure 4.
- The influence of the moving load increment segment-pair on the selected mode of failure in the chosen design region is documented, specifically:
  - whether its current position causes a positive or negative effect in the selected design region, and
  - the magnitude of the current position's influence on the selected mode of failure.
- The critical positive pattern for the selected region and mode of failure is produced when:

- loads are applied only at locations that caused a positive effect, and
- loads are applied following a decreasing sequence of their influence magnitudes. In this way, the maximum-intensity distributed load is applied at the position where it causes the most adverse effect, while reduced-intensity distributed loads are applied where the adverse effect is less. Similarly, the concentrated axle loads will be placed at the position with the highest influence value in each of the notional lanes that should be loaded.

In order to produce the critical negative loading pattern for a specific region and mode of failure, the same procedure is followed, except that the loads are only applied at locations that caused a negative effect.

A critical NA loading pattern, for a selected mode of failure and a specific design region, will provide the following information:

- Which of the loadable parts of the notional lanes should be loaded to provide the most adverse effect:
  - Which of these parts should be loaded first, i.e. with the highest intensity loading.
  - Which of the remaining parts should be loaded next with decreasing intensities, if applicable, and in what sequence.
  - Which loadable parts of the notional lanes should not be loaded:
    - The parts that should not be loaded would provide a relieving effect, i.e. a less adverse effect, if they are in fact loaded.

Critical NA loading patterns provide the most adverse positive or negative effect in a specific region for a selected mode of failure.

Comparative analyses
Qualitative finite element analyses were performed in which the results yielded by the standard patterns were compared with those of the critical patterns. For the purpose of qualitative analysis, the exact cross-section of the bridge deck does not have to be taken into consideration and the deck was modelled as an equivalent slab with constant thickness. The standard and critical patterns of type NA loading were applied to the carriageway and the values were compared relative to each other for various angles of skew of the bridge deck. The errors between the two sets of results were documented for angles of skew ranging from 0° to 40° in increments of 10°. They were measured for each of the moment resultants in the corresponding critical regions. The percentage error is defined as:

\[
\text{Error} (%) = 100 \times \frac{(v_1 - v_0)}{v_0}
\]
where

\[ v_0 \] represents the value produced by the critical pattern

\[ v_1 \] represents the value produced by the envelope of the standard patterns.

A value from the critical pattern was compared with the corresponding value from the envelope of the standard patterns, i.e. the critical value is compared with the nearest or “best” value from the set of standard patterns. The percentage error, as defined above, should always be negative or zero, otherwise the critical pattern would not be the most adverse pattern.

**INTERPRETATION OF RESULTS**

In order to interpret the results of the comparative analyses, the definition of moment resultants and the selection of critical design regions are now presented.

**Definition of moment resultants**

This paper is concerned only with the flexural behaviour of skew bridges. The concept of flexure corresponds to twisting and bending of the bridge deck under loads, which in turn translates into twisting and bending moment resultants. The moment resultants are defined below:

- **Bending moment vector**: A bending moment vector is defined as \( M_{ij} \), where \( i \) indicates the direction of the moment vector and \( j \) indicates the face on which the vector acts.

The different moment resultants are illustrated in Figure 5.

The bending moments are \( M_{12} \) and \( M_{21} \), and the two twisting moment vectors are \( M_{11} \) and \( M_{22} \). The latter two are of equal magnitude and opposite sign.

**Selected design regions**

For the purpose of this paper three design regions were chosen where each of the moment resultants would be measured and where they typically reach their critical values. The selected design regions and the corresponding moment resultants are marked with an “×” in Figure 6.

The selected design regions, the corresponding moment resultants and the assumed reinforcement layout are described below:

- **Mid-span edge**: The longitudinal bending moment \( M'_{21} \) is measured in this region. As illustrated in Figure 6, the accented axes represent rotated axes which correspond to the angle of skew. Moment resultant \( M'_{21} \) represents the bending moment vector in the rotated \( x'_2 \)-direction, acting on the rotated \( x'_2 \)-face. It is assumed that its corresponding reinforcement will be placed in the rotated longitudinal direction, parallel to the edge of the carriageway.

- **Mid-span centre**: The transverse bending moment \( M_{12} \) is measured in this region. Moment resultant \( M_{12} \) represents the bending moment vector in the \( x_2 \)-direction on the \( x_2 \)-face. The transverse bending moment has to be resisted by the reinforcement in the transverse direction, as shown in Figure 6.  

---

**Figure 5** Moment resultant definition

**Figure 6** Selected design regions and corresponding moment resultants

**Figure 7** Percentage error of the standard patterns relative to the critical patterns for the twisting moment in the bottom-right obtuse corner
Obtuse corner: The twisting moment $M_{11}$ is measured in this region. In general, the transverse and longitudinal bending moments also reach significant values in the obtuse corner. These results are not within the scope of this paper, i.e. only the twisting moment will be considered in the obtuse corner.

RESULTS

The results of the detailed comparative analyses are presented for a typical single-span, simply-supported carriageway. For critical and standard loading patterns the results are compared as the angle of skew of the carriageway increases from 0° to 40°. A carriageway with a span length of 15 m and an effective width of 10 m, e.g. a typical inner span of a multi-span bridge deck, was considered.

Twisting moment

The twisting moment in the bottom right obtuse corner of a skew carriageway (see Figure 6) is considered here. The corresponding critical NA loading patterns and resulting twisting moment contours are illustrated in Table 1.

The following can be observed from Table 1 regarding the twisting moment in the obtuse corner:

- To produce the critical NA loading pattern for the twisting moment in the bottom right obtuse corner, the top right acute corner of the carriageway should not be loaded.
- It can be seen from the load indices that the highest intensity loading should be applied towards the bottom edge of the carriageway in the transverse direction and towards mid-span in the longitudinal direction.

The comparative errors of the standard patterns relative to the critical patterns are presented in Figure 7. Both application methods of type NA loading were considered (see Figure 2), namely line loads spaced 1.9 m apart in combination with point loads, as well as uniformly distributed loads in combination with knife-edge loads over the full width of a notional lane. The errors are measured for angles of skew ranging from 0° to 40° in increments of 10°.

The critical NA loading patterns illustrated in Table 1 are not obvious. It is difficult to determine which parts of the notional lanes should be loaded and which parts provide a relieving effect, i.e. should not be loaded. Another difficulty, once the parts that should be loaded have been identified, is determining the position where the highest intensity loading should be applied.

It can be seen in Figure 7 that the standard loading patterns do not yield an acceptable result for the twisting moment in the obtuse corner. In the case of the standard patterns, all the loadable parts of a notional lane are loaded for a single-span carriageway (see Figure 3). This explains the large errors that occur between the standard values and the critical values, since areas that provide a relieving effect are also loaded. The largest error occurs at a 0° angle of skew, when the obtuse corner has not formed. At a 0° angle of skew, the values produced by the standard patterns are 21% and 30% smaller than those produced by the critical NA loading patterns for the line loads and distributed loads respectively. The errors decrease gradually up to an angle of skew of 30°. After 30° the errors increase again. As the angle of skew increases, the density of the twisting moment plots in the region of the obtuse corner also increases, as illustrated in Table 1. This indicates that the gradient of the twisting moment in the region of the obtuse corner increases as the angle of skew increases.

<table>
<thead>
<tr>
<th>Angle of skew (°)</th>
<th>Critical NA loading pattern</th>
<th>Resulting moment plot (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0°</td>
<td></td>
<td>-1.61E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.24E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-8.75E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-5.07E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.39E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.29E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.97E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.65E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.33E5</td>
</tr>
<tr>
<td>20°</td>
<td></td>
<td>-2.51E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.04E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.57E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.1E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6.36E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.67E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.02E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.71E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.24E5</td>
</tr>
<tr>
<td>40°</td>
<td></td>
<td>-2.86E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.26E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.66E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.06E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.57E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.44E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.46E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.35E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.95E5</td>
</tr>
</tbody>
</table>
Transverse bending moment
The transverse bending moment is measured at mid-span at the centre of the carriageway (see Figure 6). The critical NA loading patterns and resulting bending moment contours for this case are presented in Table 2.

The following can be observed from Table 2 regarding the critical NA loading patterns for the transverse bending moment at mid-span, at the centre of the carriageway:
- To produce the critical NA loading pattern, both the acute corners of the carriageway should not be loaded.
- The highest intensity loading should be applied towards the centre of the carriageway and towards the bottom edge of the carriageway.

Another observation from Table 2 is the concentration of the transverse bending moment in the obtuse corner when the angle of skew increases, even though the loading patterns were generated to produce critical values at mid-span and not in the obtuse corner. At a 40° angle of skew, the magnitude of the transverse bending moment in the obtuse corner exceeds its magnitude at mid-span.

The comparative errors of the standard patterns relative to the critical patterns are shown in Figure 8. It can be seen that the standard NA loading patterns do not yield acceptable results for the transverse bending moment at mid-span, at the centre of the carriageway, especially at higher angles of skew.

Longitudinal bending moment
The longitudinal bending moment under consideration here is rotated to coincide with the angle of skew. The rotated longitudinal bending moment is measured at mid-span, at the bottom edge of the carriageway, i.e. M’y, shown in Figure 6. The corresponding critical NA loading patterns and resulting bending moment contours are illustrated in Table 3.

It is clear from Table 3 that the critical NA loading patterns for the longitudinal bending moment at mid-span, at the edge of the carriageway, are as expected for a single-span deck:
- All the loadable parts of the notional lanes are loaded.

Table 2 Critical NA loading patterns and resulting contours of the transverse bending moment at mid-span at the centre of the carriageway

<table>
<thead>
<tr>
<th>Angle of skew (°)</th>
<th>Critical NA loading pattern</th>
<th>Resulting moment plot (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0°</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>−3.63E3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.19E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.74E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.29E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.99E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.94E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.05E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.2E5</td>
</tr>
<tr>
<td>20°</td>
<td></td>
<td>−4.33E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−2.28E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−2.28E3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.82E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.87E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.92E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.97E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.21E5</td>
</tr>
<tr>
<td>40°</td>
<td></td>
<td>−1.18E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−8.91E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−6.05E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−3.19E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−3.33E3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.53E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.39E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.25E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.11E5</td>
</tr>
</tbody>
</table>
The second-largest moment resultant is the twisting moment in the obtuse corner, reaching a value of 55% of the longitudinal bending moment at mid-span. The twisting moment resultant increases rapidly as the angle of skew increases.

- The transverse bending moment at mid-span, at the centre of the carriageway, remains relatively constant as the angle of skew increases, with a relative value of approximately 16–20%.

**CONCLUSIONS AND RECOMMENDATIONS**

Critical NA loading patterns, complying with TMH7, were presented and compared with simplified, standard NA loading patterns for increasing angles of skew of a bridge deck. Only the flexural behaviour of the deck was investigated.

**Critical normal traffic loading**

It was shown that the critical NA loading patterns are not obvious, particularly for the twisting moments and transverse bending moments as the angle of skew increases. It is difficult to determine which parts of the carriageway should be loaded and which parts provide a relieving effect and should not be loaded. Another difficulty is determining the loading sequence, i.e. how to vary the loading intensity over the parts that should be loaded.

The results presented in this paper provide some guidelines regarding the area that should be loaded and the loading sequence. However, in general, specialised software is required to find the critical loading patterns, even for the single-span case considered here. It can be expected that for continuous multi-span carriageways the loading arrangements on the adjacent and alternate spans will not follow any clear pattern as the angle of skew increases.

**Standard patterns and critical patterns**

Simplified standard NA loading patterns were introduced. However, they did not provide good approximations of the twisting and transverse bending moment resultants in the selected design regions. In fact, large errors of up to 90% of the critical values were recorded. The only moment resultant for which the envelope of the standard patterns produced an acceptable approximation was the rotated longitudinal bending moment at the edge of the carriageway.

It is concluded that the standard patterns may be used to provide an approximation of the longitudinal bending moment at the edge of the carriageway. However, to obtain the critical moment resultants, especially the twisting and transverse moments, specialised software is required in order to perform rigorous distribution analysis.

---

**Table 3 Critical NA loading patterns and resulting contours of the longitudinal bending moment at mid-span at the bottom edge of the carriageway**

<table>
<thead>
<tr>
<th>Angle of skew (°)</th>
<th>Critical NA loading pattern</th>
<th>Resulting moment plot (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0°</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-9.09E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-7.97E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6.85E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.61E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.49E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.37E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.25E5</td>
</tr>
<tr>
<td>20°</td>
<td></td>
<td>-6.87E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-5.84E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.8E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.77E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.74E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.71E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6.81E4</td>
</tr>
<tr>
<td>40°</td>
<td></td>
<td>-5.4E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4.43E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.46E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2.49E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.52E5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-6.81E4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.49E4</td>
</tr>
</tbody>
</table>

---

**Bridge deck behaviour**

In order to gain insight into the overall flexural behaviour of the bridge deck, the different moment resultants described above are compared with each other. The different moment resultant values are normalised relative to the largest moment resultant in each of the selected design regions and then compared as shown in Figure 10.

The following observations regarding the flexural behaviour of the particular bridge deck follow from Figure 10:

- The rotated longitudinal bending moment at mid-span, at the edge of the carriageway, is the dominant moment resultant for all angles of skew.
- The second-largest moment resultant is the twisting moment in the obtuse corner.
- The transverse bending moment at mid-span, at the centre of the carriageway, remains relatively constant as the angle of skew increases, with a relative value of approximately 16–20%.

---

Journal of the South African Institution of Civil Engineering • Volume 55 Number 3 October 2013

93
Flexural bridge deck behaviour

For the case presented in this paper, the longitudinal bending moment at mid-span, at the edge of the carriageway, was the dominant moment resultant for all the angles of skew that were investigated. When the carriageway is not skew, the twisting and transverse bending moments are small relative to the longitudinal bending moment. However, as the angle of skew increases, the twisting and transverse bending moments increase relative to the longitudinal bending moment. Although the design is not necessarily done directly for the twisting moments, e.g. the incorporation of the twisting moments in the Wood and Armer (Wood 1968) design values, the twisting and transverse bending moments deserve more design consideration at larger angles of skew.

The k-factor of TMH7

The k-factor of TMH7 is a correction factor used to compensate for the effects of partial loading of parts of the influence lines (see Section 2.A.2.2 of TMH7). The complication arises from the fact that high-intensity loading of part of an influence line may result in a more severe effect than that caused by loading the whole base of the influence line at a lower intensity. With NA loading and an influence line base that exceeds 36 m in length, the challenge is to determine whether a more severe effect will be obtained if the whole base is loaded at reduced intensity, or whether just the peak of the influence line should be loaded at the full intensity of 36 kN/m. Figure 11 illustrates this problem.

TMH7 introduces certain correction factors, called k-factors, to compensate for the effects illustrated in Figure 11. The k-factors depend on the tails of the influence lines and can be difficult to apply in practice. The procedure presented in this paper, with the incremental generation of critical NA loading patterns, provides a solution whereby the k-factor is no longer required. With the incremental generation of critical patterns, the area with the most severe effect, i.e. the peak of the influence line, will be loaded with the full intensity of type NA loading. The parts with decreasing influence values are loaded with correspondingly lower intensities, when these become applicable. The solution is illustrated in Figure 12.

Although the procedure illustrated in Figure 12 is in contradiction to the definition of a separate loadable part (Section 2.6.3.2.1 of TMH7 states that a separate part is that continuous length of notional lane that has an entirely positive or negative effect on the member being analysed), the solution presented here eliminates the need for the k-factor. The procedure considers not only the positive or negative effect, but also the specific magnitude of the positive or negative effect, i.e. there is true optimisation of the information provided by the influence surface.

Recommendations for future work

In order to keep this paper detailed and focused, its scope was limited to critical normal traffic loading for flexure of skew bridges and only a single-span deck was considered. Multi-span, continuous decks were also investigated but are not reported on here. However, additional investigations still have to be done and they are briefly discussed below.
Additional types of traffic loading

Normal (NA) loading was investigated in this paper. Similar studies should be conducted for abnormal (NB) and super (NC) traffic loads.

Additional information for bridge designers

Future research should include additional information that is useful to bridge designers, including the determination of the following:

- The impact on the dimensioning moments, using Wood, Armer, Kuyt or the French equations, specifically since most slabs are isotropic and the cracked stiffness may be different in the transverse and longitudinal directions.
- The impact of elastomer bearings, specifically on hogging and torsion in the corners.

Code considerations

The current South African specification of TMH7 was written in 1981 and was last revised in 1988. By contrast, the Eurocode has undergone more recent upgrades. A comparative calibration analysis can be performed between the Eurocode and TMH7, similar to the comparative analyses performed in this investigation between the standard patterns and the critical patterns. Proposed adjustments to TMH7 can also be tested and compared with the current TMH7 by using the methods and software presented in this paper.

ACKNOWLEDGEMENTS

The authors wish to thank all the participating bridge engineers who provided valuable input throughout the course of the research. We also gratefully acknowledge the OSP funding from the University of Stellenbosch.

REFERENCES


Structural optimisation has been the subject of intense research for many years. However, the engineering profession seems to be oblivious to the important role of natural structures. Natural structures are shapes created by nature and possess geometrical properties that are resistant to environmental conditions. Engineers of antiquity recognised the importance of natural shapes and incorporated these designs into many political and religious structures. Many of these structures are catenary or funicular in form, and are characterised by a pure compressive stress distribution – an essential requirement for masonry structures. Catenary shapes are robust and economical. These principles are applied to the design and construction of a prototype shell – a domed vault – intended as an alternative structure for low-cost housing. Design principles are described, as well as the construction process.

Figure 1 2010 Earthquake damage in Haiti (Source: Wikipedia commons file)
survived – the palm trees are still standing, without any outward signs of distress. Perhaps we will never work and live in palm trees, but the shape and flexibility of this natural structure is perfectly resistant to earthquakes and significant meteorological occurrences – thus, a structure based on nature is a more appropriate form for the Haitian environment.

Natural structures have, perhaps, evolved over millions of years into optimum shapes that are resistant to local environmental conditions. Would it not be prudent to learn from these structures and emulate them in our designs (Bejan & Lorente 2006)?

Nature possesses and persistently reproduces several universal features. All natural structures have curved surfaces and solid forms are not solid, but composed of cellular or tubular shell-like substructures. As an example, Figure 2 is a cross-section of a dinosaur femur bone and Figure 3 is the cross-section of a Black Walnut leaf. In these examples (fauna and flora) the structure and substructure of the natural forms are curve-shaped and composed of shells. The curvature of the shapes is directly related to the flow of stress, dictated by the imposition of environmental conditions. The curvature of the femur bone constitutes resistance to the actions of walking, running, jumping, etc. Leaves, which are miniature solar panels engaged in the act of photosynthesis, are designed to resist wind flutter, rain, snow and the invasion of insects. Each shape takes on a suitable form to optimise resistance to environmental forces and to ensure the species’ ultimate survival.

**STRUCTURAL EFFICIENCY OF RESISTING FORCES**

An important concept in comprehending the strength of curved structures is to understand how loads are efficiently carried by structures. To illustrate this concept, a simple example is given in Figure 4.

The maximum bending stress ($\sigma_b$) of a beam subjected to a point load (Figure 4a) is expressed by Equation 1.

$$\sigma_b = \frac{3PL}{2bh^2}$$

where

$L$ = length  
$h$ = depth of the beam  
$b$ = width

If the same load is resisted by an axial member, the maximum compressive stress ($\sigma_c$) (Figure 4b) in the column is given by Equation 2.

$$\sigma_c = \frac{P}{bh}$$

If the bending stress in the beam ($\sigma_b$) is divided by the column stress ($\sigma_c$), the ratio of the stresses ($r$) is given by Equation 3.

$$r = \frac{3L}{2h}$$

The value ($L/h$) is the span-to-depth ratio, which usually falls within prescribed values (SANS 10100-1 2000). Table 1 is a list of stress ratios for a variety of simple load and beam support conditions. The stress ratio represents the increase in stress compared with that of an axially loaded member. As indicated in the first row of Table 1, the stress in a simply supported beam subjected to a point load is 24 times higher than that in an axially loaded member. Similarly, the stress increase in other beam load/support configurations ranges from 12 to 18 times the stress in a column.

Resisting forces by bending is highly uneconomical compared with axial resistance. Avoiding bending, however, is not always feasible with our current practice of designing box-shaped linear structures. Table 1 demonstrates that greater efficiency is achieved by directing the flow of force along the axis of the member. The axis, however, does not necessarily have to be straight.
A stress will flow along the axis of a curved member or along the curvature of a shell if the shape is catenary. The catenary is a natural form in which stresses will flow in pure compression or tension. Unlimited variations of the catenary are also possible by varying the loads and corresponding shape.

**FORGOTTEN LESSONS OF OUR FOREFATHERS**

The discovery of natural structures and their importance seems to have occurred around the 17th century arising from the work of Robert Hooke (O’Dwyer 1999), the same person who developed the theory of elasticity (i.e. Hooke’s Law). Hooke, a celebrated scientist, discovered the significance of a hanging chain and its relationship to structural forms. He recognised that a hanging chain is in complete tension, without bending and shears. If the chain is locked and inverted, the stress in the chain is reversed and in pure compression. The importance of this discovery is that loads may be carried in pure compression, eliminating shears and bending.

Strangely, Hooke’s discovery was written in the margins of his scientific works – seemingly a fleeting thought. Translated from Latin, “... as hangs the flexible line, so but inverted will stand the rigid arch.” One of the first notable applications of Hooke’s arch theory was applied to St Peter’s Basilica in Rome. The dome exhibited serious cracking and the safety and stability of the structure was questioned. In 1743 Pope Benedict XIV appointed Giovani Poleni to assess the structure (Lopez 2005). Poleni hung 32 weights on a chain to represent the self-weight of the dome walls. The resulting profile of the chain represented the natural thrust line, which was inverted and superimposed on a scaled cross-section of the dome walls (see original drawings in Figure 5). Since the chain profile fitted within the walls of the dome, the structure was deemed stable and safe for occupation. According to Poleni, if the thrust line fell outside the walls, tension cracking would be expected and the structure would be unstable. He concluded that the source of cracking was the inferior materials used in construction, rather than a mismatch of the thrust line. Nearly three centuries later, St Peter’s Basilica remains stable as predicted, substantiating Poleni’s analysis.

The application of the catenary shape to structural forms was taken further by Culmann and Stevin (Block 2006) who developed a graphical analysis method, referred to as the force polygon method (in later years, called the “tip-to-tail” method) (Firmage 1980). The graphical analysis method was the dominant method of structural analysis until the early 1900s. For this reason (until recently), graphical methods were taught in most engineering programmes. Figure 6 illustrates the force polygon of an arch. The weighted diagonal arrows of the force polygon represent the reactions at the base, and the weighted vertical arrow is the total vertical load of the arch. The lines fanning from the left-hand point of the polygon represent the internal thrusts of the free-body segments. The arch may be randomly discretised or the free-body segments may be actual masonry blocks. By satisfying the equilibrium of each segment of the arch, a catenary shape forms automatically.

The catenary shell is, perhaps, the most time enduring shape known to humankind. The Pantheon in Rome is nearly two millennia old, yet proudly remains intact and serviceable. Although the inner surface of the dome is a hemisphere, the secret of the Pantheon is its outer profile, which is catenary, thus permitting the flow of stress in pure compression (Goshima 2011). Although Hooke is credited with the discovery of the catenary, evidence suggests that knowledge of the catenary may have been known much earlier by the Romans, the master builders.

Historically, the catenary dome was used only in a few notable structures, advocated by antiquity’s most prominent engineers-architects, who recognised and understood the structural importance of these shapes. By far the majority of shells were hemispheres. These shapes, however, frequently cracked due to hoop tensions at mid-height and bending and shears at the base (stress characteristics of a hemisphere). For this reason, hemispherical domes were often clad with sheets of copper or tiles to make the shells water-tight, cover unsightly cracks and improve durability.

Unfortunately, the practice of implementing natural forms seems to have faded in the
early 1900s – an incomprehensible digres-
sion in structural design. Utilising natural
forms should be a logical choice, especially
in the design of low-cost housing for which
structural and material efficiency is of para-
mount importance.

APPLICATION OF NATURAL SHAPES
TO HOUSING STRUCTURES

The South African Department of Housing
(Sexwale 2009) has reported a serious
dearth in low-cost housing. The deficiency
is approximately 2 million homes, exclu-
sively among the low-income sector of
society. The key issue is affordability, for
the recipients, who are largely unemployed
or underemployed. Alternative housing
forms and new construction materials are
therefore necessary, since existing building
forms are unsatisfactory in quality and price.
Shell structures, which are natural forms,
are ideally suited as the shape provides both
superior strength and economy of materials.

Although not covered in this paper,
shells are climatically appealing and energy
efficient, since their internal temperatures
are known to be more evenly distributed – a welcome and important attribute for
low-cost housing. This benefit, however, is
only fully realised if the shell is insulated
(Wilson 2005).

The shape investigated was the catenary
vault, which may be described as two cate-
nary barrel shells intersecting at right angles.
A vault was chosen for practical reasons – vertical walls are constructed on the four
open sides of the shell, allowing for finishes
and furniture that are more suited for box-
shaped homes.

The shape of the vault is more complex
than that of a catenary arch or dome. The
catenary mathematical formula is two-
dimensional and can only be inscribed on
the shell diagonally, from support to sup-
port, and along the open edges of the shell
(Figure 7); the shape is therefore undefined
between the inscribed catenary curves. The
mathematical expression of the catenary was
developed by Gregory (1706), in response to a
competition.

\[ y = a \cosh \left( \frac{x}{a} \right) \]  \hspace{1cm} (4)

where the symbol \( a \) is a constant, defining
the amplitude and width of the curve.

The force polygon method is also only suited
to two-dimensional shapes or shapes that
have rotational symmetry about a vertical
axis (e.g. cupola or conical shapes). For
these reasons, physical working models are
implemented to define the surface of the
shell using a hanging model, as illustrated
in Figure 8. The model was constructed of
plaster of Paris medical bandages, hung to
a specified width:height ratio. The model
was then photographed and the coordinates,
defining the shape of the shell, were meas-
ured. The resulting shape was then scaled
to match the size of the full-scale version.
Fabric models must be relatively flat to
prevent, or minimise, the restraining effects
of the weave from defining the shape of the
model incorrectly.

The concept of the hanging model is
similar to the hanging chain, but in three
dimensions. When the hanging model is
inverted, the stresses are transformed into
pure compression, which is ideally suited to
masonry structures. Most important, the
hanging model is an appropriate method for
defining the shape of the shell in the unde-
fined regions (Figure 7).

A three-dimensional chain model was
also constructed (Figure 9) to emulate the
modelling methods of antiquity. However,
this method was found to be more difficult
to work with. It was almost impossible to
define the coordinates from photographs and
therefore this method was not used in the
design. Designers of the past implemented
hanging chain models to define the shape of
structures. The coordinates were painstak-
ingly determined by means of rulers and
squares. As a consequence, the working
models comprised very few draping chains
and the shapes were simple (i.e. without
openings). However, the simplicity of the
dome models accumulated into highly
complex representations if the design was
composed of multiple shells. Antoni Gaudi's Sagrada Familia in Barcelona is an example of a structural design done by means of three-dimensional chain models (Figure 10) (Wikipedia commons file).

CONSTRUCTION OF THE CATENARY VAULTED SHELL

The construction of the shell consisted of two parts: firstly, the construction of the formwork, using earth tiles to define the shape of the shell; and secondly, the construction of successive layers of tiles to achieve the desired thickness. The tiles are composed of compressed cement-stabilised earth, measuring 290 mm x 140 mm x 20 mm, and having a compressive strength of approximately 7 MPa. The plan dimensions of the vault measured 3 m x 3 m and had an apex height of 2 m. The prototype shell is approximately ½ scale.

Four small pad footings were constructed, which were tied by steel straps along the boundaries. The purpose of the pad footings is to resist the vertical loads, and the straps resist the horizontal outward thrusts of the shell. Both structural elements are absolutely vital to maintain stability and satisfy equilibrium requirements. In a full-scale structure, however, the foundation would be in the form of a strip foundation. A prominent and common feature of low-cost housing is damage caused by uneven settlement. Shell structures are typically lighter and therefore cause less soil stress. A unique feature of a similar dome, constructed in Johannesburg, is the orientation and configuration of the foundations (Talocchino 2005). The foundations were orientated vertically (turned 90°) to provide stiffness – if uneven settlement occurs, the structure is designed to settle without distorting the shell.

Four catenary-shaped guides were placed on top of each foundation, used to construct the four arches along the edges of the shell.
The first layer of tiles is “glued” together using Crystacal R, which is a high-strength, rapid-hardening plaster of Paris. The initial set usually occurs within 1 minute (depending on the temperature), allowing rapid construction of the initial layer of tiles.

Formwork was kept to a minimum to economise on construction. In addition to the edge arches, four Masonite strips were placed diagonally, spanning between the pad footings (Figure 12). The Masonite strips were used as guides to construct the “domed” portion of the vault.

The first layer of tiles is the most important layer since it defines the shape of the shell. The thrust line must match the curvature of the shell and therefore the accuracy of the first layer is of paramount importance. Not only does the first layer define the shape, but it also acts as formwork for successive layers.

Additional layers are then applied to the top surface of the first layer and ordinary mortar is used in-between and along the edges of the tiles. The prototype model has two layers of tiles over the entire surface, but additional tiles are placed near the supports to account for the accumulation of compressive forces at the base of the shell. The thickness of the shell is determined by comparing the estimated stress and theoretical capacities of the compressive and buckling strengths of the shell.

Figure 13 is a photo of the completed shell. The thickness of the shell is 50 mm (two layers of tiles, separated by 10 mm of mortar). The shell is completely unreinforced and the resistance is based on the compressive strength of the tiles. The shell is ultra-thin, but capable of resisting the load of three individuals (see Figure 13), thus demonstrating the incredible strength of catenary shapes.

Since shell structures are optimal and materially efficient, higher accuracy in construction is required to ensure that the constructed shape is within tight tolerances. Supervision is an obvious necessity, but creative construction techniques can minimise construction error and allow the engagement of unskilled labour. For example, the hemispherical shell constructed in Johannesburg used inflatable formwork, in the shape of a balloon. The balloon provided a template and bricks were laid on the surface of the balloon by ordinary bricklayers. After the mortar had set, the balloon was deflated and reused in the construction of other shells. The technique permitted rapid erection of the formwork, a guiding template to ensure construction accuracy and reuse of the formwork.
Shells have declined in popularity, which may have been due to the difficult and time-consuming task of erecting formwork or guiding templates. Furthermore, the relatively low cost of materials and the high cost of labour advocates simplicity and speed of construction, which has taken preference over the optimisation of shapes. The world economies, however, have changed in recent years, causing an increase in the cost of materials and therefore concepts of sustainability must be incorporated into new designs – these factors may spur a revival in shell design.

**CONCLUSIONS**

Nature should be the engineer’s mentor. The shapes and forms we see in nature have evolved over time and are optimally load-resistant structures. Many of these are catenary shapes, which are resistant to environmental conditions. Ironically, architects endeavour to replicate building forms that harmonise with the environment. The correctness of this approach is far more apparent when structural resistance and the economy of materials are considered.

Natural shapes have inspired and enlightened the engineering profession for centuries. Catenary concepts have been incorporated into numerous buildings, the majority of which are religious, with the intention of lasting for eternity (and many of these structures probably will!). They have participated in the test of time and have proved to possess superior strength and resilience. Why natural shapes have become foreign to our profession is an absolute mystery. We have deviated from the obvious and have forgotten the lessons of the past.

The dearth of housing in South Africa dominates our agenda and much of the country’s resources have flowed in this direction. We need housing that is robust and economises on materials – natural forms fit these requirements perfectly. A scaled-down prototype was designed and constructed to determine the constructability and strength of a catenary vault. The catenary vault has demonstrated incredible strength, despite being constructed of ultra-thin layers of compressed earth tiles. No steel was incorporated, yet the shell does not exhibit any surface cracking. This implies that the shape is a natural form (i.e. the thrust line coincides with the centre line of the shell walls) and is therefore a highly efficient structure.

**ACKNOWLEDGEMENT**

The authors would like to acknowledge the generous contribution of Hydraform, who donated the earth bricks used in the construction of the catenary vault.

**REFERENCES**


Wilson, A 2005. A practical design of concrete shells. Monolithic Dome Institute, Italy, Texas, US.
Exploring the flow number (FN) index as a means to characterise the HMA permanent deformation response under FN testing

L F Walubita, J Zhang, A E Alvarez, X Hu

Permanent deformation (PD), or rutting, is one of the common distresses occurring in hot-mix asphalt (HMA) pavements. As part of the HMA mix- and structural-design processes to optimise field performance, various laboratory tests, including the Hamburg wheel tracking test (HWTT) and flow number (FN), have been developed to characterise the HMA mix rutting resistance potential. With this background, this study was conducted to explore the potential of routinely using the FN test to characterise the PD response of Texas HMA mixes as a supplement to the HWTT. Towards this goal, a new PD parameter – the FN index – was developed to differentiate and screen the HMA mixes. The research methodology incorporated a two-phase approach, namely: (1) laboratory testing and (2) field correlations. Overall, the findings indicated that the FN index, computed from the FN test data, has the ability to statistically differentiate the HMA mixes evaluated, as well as promising potential to supplement the HWTT for routine HMA mix-design and screening. In addition, the FN test has a practically reasonable test time (about three hours per specimen) and is cost-effective. However, the inability to readily test thin field cores and the need for field validation with long-term performance data remain some of the key challenges to be addressed with the FN test method.

INTRODUCTION
Permanent deformation (PD) – also termed rutting – is one of the major distresses occurring in hot-mix asphalt (HMA) pavements. While the total PD in the HMA pavement structure can either be due to the HMA mix, subgrade or other structural layers, or contributions from all these, surface rutting is predominantly due to issues with the surface HMA mix rather than PD of the underlying layers – see Figure 1. Therefore, the focus of this study was on the PD of the HMA mix, assumed to be predominantly related to materials selection and HMA mix-design. However, this is not to discount the effects of poor structural design, poor construction practices, environmental effects, and/or excessive traffic loading (Walubita et al. 2012; WsDOT 2011).

As illustrated in Figure 1, the primary mechanism of PD in the HMA mix is shear deformation caused by large stresses in the upper portions of the HMA layers under traffic loading, particularly at elevated temperatures (Walubita et al. 2012). Thus, this type of PD occurs mostly in high shear locations, in particular at intersections where braking, accelerating/decelerating and slow moving traffic take place repeatedly. In addition, the PD distress is accelerated during the summer seasons, where the extreme high pavement temperatures contribute to the instability rutting of the HMA mix, particularly under high traffic loading and/or where softer asphalt binders have been used.

On the pavement structure, rutting typically manifests itself as surface depressions in the wheel paths (Figure 1).

Properly designed HMA mixes, that are identified and screened by appropriate laboratory testing, are thus required to minimise the PD on the pavement surface and/or within the pavement structure when HMA mixes are used as intermediate layers. Rutting is considered a structural failure that undesirably distorts the pavement ride quality, and water pooling after rains often causes vehicle hydroplaning with a high potential for traffic accidents. Also, pavement maintenance or rehabilitation activities are financially straining. Thus, the selection of sufficiently rut-resistant HMA mixes during the HMA mix-design stage, based on appropriate laboratory testing, is crucial (Walubita et al. 2012).

Keywords: hot mix-asphalt (HMA), rutting, permanent deformation (PD), flow number (FN), FN test, FN index, Hamburg wheel tracking test

TECHNICAL PAPER
JOURNAL OF THE SOUTH AFRICAN INSTITUTION OF CIVIL ENGINEERING
Vol 55 No 3, October 2013, Pages 103–112, Paper 940

DR LUBINDA F WALUBITA holds a PhD, an MSc (Eng), and a BSc degree in Civil Engineering from the Texas A&M University, the University of Stellenbosch, and the University of Zambia, respectively. He currently works as a Research Scientist for the Texas Transportation Institute (TTI) of The Texas A&M University System, USA. His area of specialisation is roads, materials, and pavement engineering. His other research interests include nanotechnology, where he currently serves as the Vice-President for the Global Network of NanoTechnology and its Applications to Road Materials and Pavements (GNN-RMP). He is an author of over 100 publications including editing an ASCE-GSP book. Dr Walubita also has substantial industrial experience as a roads engineer in South Africa and Zambia.

Contact details:
TTI – The Texas A&M University System
CE-TTI Bldg, Room 330C, 3135 TAMU
College Station, TX 77843, United States of America
T: +1 979 862 3356, F: +1 979 845 1710, E: lfwalubita@hotmail.com

JUN ZHANG graduated from the University of Nebraska–Lincoln with a Masters degree in Civil Engineering in 2010. He has been working as a Research Associate for the Texas Transportation Institute (TTI) of The Texas A&M University System since 2011. His area of specialisation is pavement engineering and pavement materials.

Contact details:
TTI – The Texas A&M University System
CE-TTI Bldg, Room 330E, 3135 TAMU
College Station, TX 77843, United States of America
T: +1 979 458 5728 E: jpyj@obd@gmail.com

PROF ALLEX E ALVAREZ holds a PhD, an MSc (Eng), and a BSc degree in Civil Engineering from the University of Magdalena, the Faculty of Engineering of the University of Los Andes, and the National University of Colombia respectively. He currently works as an Associate Professor for the University of Magdalena in Santa Marta, Colombia, where he leads the CRC research group. His area of specialisation is pavement engineering, paving materials, and roads.

Contact details:
University of Magdalena
Carrera 12 No 22-08, Santa Marta, Colombia
T: +57 5 430 1292, F: +57 5 430 1292, E: allexalvarez@yahoo.com

PROF XIAODI HU holds a PhD and an MSc (Eng), and a BEng degree in Civil Engineering from the Tongji University, and the Huazhong University of Science and Technology. He worked for almost five years as a Research Associate for the Texas Transportation Institute (TTI) of the Texas A&M University System, USA, and now he works as an Associate Professor for the Transportation Research Center of the Wuhan Institute of Technology, China. His area of specialisation is roads, materials and pavement engineering. His research work is internationally acknowledged and he has authored over 30 publications. Xiaodi also has substantial industrial experience as a roads engineer in China.

Contact details:
Transportation Research Centre
Wuhan Institute of Technology, 695 Xiongchu Avenue, Wuhan, Hubei Province 430073, China
T: +86 1587 9719 4471, E: xiaodi8282@hotmail.com

Keywords: hot mix-asphalt (HMA), rutting, permanent deformation (PD), flow number (FN), FN test, FN index, Hamburg wheel tracking test
Currently existing HMA PD-related tests

Various laboratory tests are currently in use to characterise the HMA mix PD response, including field PD performance prediction. These tests include the Hamburg wheel tracking test (HWTT), asphalt pavement analyser (APA), dynamic modulus (DM), repeated load permanent deformation (RLPD), flow time (FT) and flow number (FN) (AASHTO 2001; Abdallah & Nazarial 2011; Archilla et al. 2007; Goh et al. 2011; Mohammad et al. 2006; Walubita et al. 2012; Witczak et al. 2002; Witczak 2007; Zhou & Scullion 2001). Walubita et al. (2012) comparatively evaluated the HWTT, DM and RLPD tests and concluded that each test has its own merits and demerits, and that the onus is to be cautious as to which test to use depending on the specific engineering needs. Although there are inherent issues of high sample confinement and the inability to readily capture the HMA mix shear properties, the overall conclusions were that, of the three tests that were evaluated (Walubita et al. 2012), the HWTT is the most practical test for daily routine HMA mix-design and screening. The RLPD and DM tests were found to be better suited for structural design applications, such as generating input data for mechanistic-empirical models (Walubita et al. 2012). In fact, the Texas Department of Transportation (TxDOT) currently uses the HWTT, as Test Procedure designation Tex-242-F (TxDOT 2011), for routine HMA mix-design and screening in the laboratory, as well as an indicator of field rutting performance.

Study objective and scope of work

With the above background in mind, this study was undertaken to explore the potential of routinely using the FN test to characterise the PD resistance of Texas HMA mixes as a supplement to the standard HWTT test method (TxDOT 2011). Various Texas HMA mixes were comparatively tested in both the FN test and HWTT, and also related to in situ field performance (for some selected mixes). The observed advantages and disadvantages of the FN test in comparison to the HWTT are also discussed in the paper.

In terms of the paper organisation, following this introduction is a description of the FN test, along with the FN analysis models and output parameters, and the HWTT. The experimental design is then presented, followed by the laboratory test results and analysis and preliminary correlations with field performance data. The paper then concludes with a comparison of the laboratory test methods and a summary of key findings and recommendations.

THE FLOW NUMBER (FN) TEST METHOD

The FN is one of the laboratory PD tests that show promise for rutting performance evaluation of HMA mixes, which can be used as a supplement to the HWTT (Archilla et al. 2007; Goh et al. 2011; Mohammad et al. 2006; Witczak et al. 2002; Witczak 2007; Zhou & Scullion 2002; Zhou & Scullion 2003; Zhou et al. 2004). As shown in Figure 2, the FN test involves application of a specific vertical compressive (dynamic) stress level to measure the HMA accumulated vertical deformations as a function of time or load cycles. As Figure 2 shows, the Universal Testing Machine (UTM-25) was used to conduct the FN test.

In this study, the FN test was conducted at 50°C (122°F) for temperature consistency with the HWTT test – which is subsequently discussed – and to closely simulate the average Texas high summer pavement surface temperatures (when HMA mix rutting is generally more critical). A compressive repeated Haversine stress-controlled loading mode at 1 Hz (0.1 seconds loading plus 0.9 seconds rest period = 1 cycle) was applied using the UTM (see Figure 2). Based on several trial tests by these researchers at 50°C, the applied stress was selected as 207 kPa (30 psi) in magnitude with zero confinement stress for all the mixes evaluated in this study. The FN test was set to terminate at 10 000 load cycles or after accumulation of 30 000 micro-strains, whichever came first (i.e. after about 3 hours of testing time or 166.7 minutes to be exact).

With these test parameters, a single FN test on an HMA mix cylindrical specimen – 100 mm (4 inches) diameter by 150 mm (6 inches) in height – lasted for at most three hours. The HMA mix specimens were temperature-preconditioned for a period of approximately three hours prior to testing, with the temperature monitored via a thermocouple probe inserted inside a dummy HMA mix specimen also placed in the same temperature chamber as the test specimens.

Data analysis models and output parameters

During FN testing, the primary output data include the load (stress), number of load cycles (or load cycles), deformation (strain) and time per load cycle. Based on a plot of accumulated permanent strain versus load cycles, the following PD parameters are generated and used as indicators of the HMA mix rutting resistance potential:
### Table 1 FN data analysis models

<table>
<thead>
<tr>
<th>#</th>
<th>Item/Parameter</th>
<th>Model</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>General relationship between the accumulated permanent strain and the number of load cycles</td>
<td>$\varepsilon_p = aN^b$</td>
<td>$\varepsilon_p$ is the accumulated permanent strain due to dynamic vertical loading, $N$ is the number of load cycles to produce $\varepsilon_p$, and $a$ and $b$ are regression constants that depend on the material and stress state conditions.</td>
<td>Archilla et al 2007, Zhou et al 2004</td>
</tr>
<tr>
<td>2</td>
<td>Probabilistic distribution (Weibull) model for the relationship between $\varepsilon_p$ and $N$</td>
<td>$N = \gamma [1 - e^{-\beta/F}]$</td>
<td>$\beta$, $\gamma$, and $\alpha$ are the probability distribution and shape parameters. The parameter $\gamma$ has the simple interpretation of being the maximum number of load cycles that the specimen would last if the testing machine could apply an arbitrary deformation to the sample (i.e. the number of load cycles at which the rate $d\varepsilon_p/dN \to 0$).</td>
<td>Archilla et al 2007</td>
</tr>
<tr>
<td>3</td>
<td>Predicted permanent strains ($\varepsilon_p^{\text{Predicted}}$)</td>
<td>$\varepsilon_p^{\text{Predicted}} = \frac{1}{\beta} \times [-\ln \left(1 - \frac{N}{\gamma}\right)]^{\alpha}$</td>
<td>$\varepsilon_p^{\text{Predicted}}$ is the predicted accumulated permanent strain as a function of $N$, where $\gamma$, $\alpha$, and $\beta$ are as previously defined.</td>
<td>Archilla et al 2007</td>
</tr>
<tr>
<td>4</td>
<td>Flow number (FN, cycles)</td>
<td>$FN = \gamma [1 - \exp \left(\frac{1}{a} - 1\right)]$</td>
<td>$FN$ = flow number or number of load cycles at the onset of tertiary zone; at which $d^2\varepsilon_p/dN^2 = 0$</td>
<td>Archilla et al 2007</td>
</tr>
<tr>
<td>5</td>
<td>Accumulated permanent strain at tertiary flow (or per-manent strain) at tertiary flow ($\varepsilon_p(F)$, microns)</td>
<td>$\varepsilon_p(F) = \frac{1}{\beta} \times \left[1 - \left(\frac{N}{\gamma}\right)^{1/\alpha}\right]$</td>
<td>$\varepsilon_p(F)$ = accumulated permanent strain at the onset of tertiary flow, i.e. at $d^2\varepsilon_p/dN^2 = 0$.</td>
<td>Archilla et al 2007</td>
</tr>
<tr>
<td>6</td>
<td>Time to tertiary flow ($t(F)$; minutes)</td>
<td>$t(F) = \frac{FN}{60}$</td>
<td>$t(F)$ = time at the onset of tertiary flow (based on a loading frequency of 1 Hz) or time count in minutes at $d^2\varepsilon_p/dN^2 = 0$.</td>
<td>–</td>
</tr>
<tr>
<td>7</td>
<td>FN index (micro-strains/cycle)</td>
<td>$FN\text{ Index} = \frac{\varepsilon_p^{\text{Predicted}}(F)}{FN}$</td>
<td>Derived composite parametric ratio that simultaneously incorporates the strain at tertiary flow $\varepsilon_p(F)$, and flow number (FN) at tertiary flow.</td>
<td>–</td>
</tr>
</tbody>
</table>

![Graphical illustration of the FN (TxDOT 2004)](image)

- Flow number (FN), in cycles
- Accumulated permanent strain (or permanent strain) at tertiary flow ($\varepsilon_p(F)$), in microns
- Time to tertiary flow ($t(F)$), in minutes

The respective analysis models for computing these parameters are listed in Table 1. The FN is defined as the number of load cycles for which the slope of the curve of accumulated permanent strain versus load cycles is minimum (Archilla et al 2007; WsDOT 2011) or as the number of load cycles at which tertiary flow (i.e. tertiary zone) begins (Figure 3). As shown in Figure 3, tertiary flow is differentiated from secondary flow by a distinctive departure from the linear relationship between the cumulative permanent strain and number of load cycles in the secondary zone. Thus, the “PD failure” of the HMA mix was defined as the point of onset of tertiary flow.

#### The flow number (FN) index concept

In addition to the traditional FN (cycles), $\varepsilon_p(F)$, and $t(F)$ parameters, the FN index was introduced (Table 1) as an exploratory concept to fully capture the HMA mix PD response and is defined as the ratio of $\varepsilon_p(F)$ to FN (cycles) at tertiary flow. Theoretically, lower FN index values are related to HMA mixes that are more stable and resistant to PD. The opposite should also theoretically hold. As a supplement and/or surrogate to the individual classical PD parameters, the basic idea was to investigate if a simple composite parameter or a derivative (the FN index) that combines the classical individual PD parameters, provided superior and/or more distinctive differentiation and screening capabilities of the HMA mixes.

#### THE HAMBURG WHEEL TRACKING TEST (HWTT)

In Texas, the HWTT is the standardised routine test used for characterising both the rutting resistance potential and stripping susceptibility (i.e. moisture damage potential) of HMA mixes in the laboratory (TxDOT 2011). The standard HWTT test parameters that were used are 703 N (158 lbs) vertical loading, applied at a rate of 52 passes per minute in a 50°C (122°F) water bath; with 150 mm (6-inch) diameter by 62.5 mm (2.5-inch) in height paired specimens loaded up to 20 000 HWTT load passes for about 6 hours 25 minutes (about twice the FN test time) (TxDOT 2011). The primary output data is the HMA mix rut depth as a function of the number of load passes. The terminal HWTT rutting failure criterion in the State of Texas is 12.5 mm (0.5-inch) rut depth, i.e. $R_u = 12.5$ mm (TxDOT 2011; Zhou & Scullion 2001).

#### EXPERIMENTAL DESIGN

Six Texas HMA mixes were evaluated, including crack attenuating mixtures (CAM),
permeable friction course (PFC), Type B (coarse-graded), Type D- and F-dense graded mixes, and stone matrix asphalt (SMA) mixes; Table 2 presents the corresponding mix design characteristics. For each HMA mix, a minimum of three replicate samples were molded from HMA plant-mix materials and tested per each test type. With the exception of the PFC mix samples that were molded to a final density of 80±2%, all the other HMA samples were molded to a final target density of 93±1% as specified by the TxDOT standards (TxDOT 2004).

The categorisation of mix PD resistance in the last column of Table 2 was based on the rutting resistance performance and stiffness of the mixes from previous laboratory testing (i.e. HWTT, RLPD and DM) and historical field performance observations (Walubita et al 2012). However, this categorisation should not be taken as a standard, but was merely used as a reference guide for this study.

LABORATORY TEST RESULTS AND ANALYSES

This section presents the results and the corresponding data analyses, based on a minimum of three replicate samples (or sample sets) per mix per test type. However, it should be noted that these test results pertain only to the HMA mixes and the laboratory test conditions defined in this study. Therefore, the overall findings and conclusions may not be exhaustive.

FN test results

Figure 4 shows an example of a plot of the accumulated permanent strain and strain rate (slope) determined from the FN test data using the models listed in Table 1 (i.e. Equation 3) and the MS Excel spreadsheet optimisation technique based on minimising the sum of square error method (Archilla et al 2007).

Table 3 presents the FN test results – computed for each HMA mix as exemplified in Figure 4 – and the corresponding statistics expressed in terms of the mean, standard deviation (Stdev), and coefficient of variation (COV) values. In addition, Figure 5 shows a comparison of the computed FN test parameters. This comparison suggests that the \( ep_F \) parameter exhibits no trend and will thus be unable to effectively differentiate and screen the HMA mixes. On the other hand, both the FN (cycles) and \( t_F \) parameters are showing the theoretically expected opposite trend to that of the FN index, except for the deviation in the trend by the CAM mix; i.e. the higher the FN (cycles) and \( t_F \) parameters in magnitude, the lower the FN Index.

Theoretically, the lower the FN index in magnitude, the more resistant to PD the HMA mix is. Thus, the ranking of the HMA mixes based on the FN index would be as follows: SMA (best) → Type F → Type B → Type D → CAM (second poorest) → PFC (poorest). As indicated in Table 3, two specimens of the SMA mix even lasted up to 10 000 load cycles (i.e. FN index < 0.67). The SMA gap gradation and internal structure provide a very good stone-on-stone contact condition that is responsible for this mix’s excellent PD resistance performance. By contrast, the poorer ranking performance

<table>
<thead>
<tr>
<th># HMA mix</th>
<th>Aggregate gradation</th>
<th>Mix-Design</th>
<th>AV (%)</th>
<th>Highway where used</th>
<th>Category of HMA Mix PD resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 CAM</td>
<td>Fine-graded (9.5 mm NMAS)</td>
<td>7.0% PG 64-22 + Igneous/limestone</td>
<td>7±1%</td>
<td>SH 121 (Paris)</td>
<td>Poor</td>
</tr>
<tr>
<td>2 Type B</td>
<td>Coarse-graded (19 mm NMAS)</td>
<td>4.6% PG 64-22 + Limestone + 30% RAP</td>
<td>7±1%</td>
<td>IH 35 (Waco)</td>
<td>Good</td>
</tr>
<tr>
<td>3 Type D</td>
<td>Fine-graded (9.5 mm NMAS)</td>
<td>5.1% PG 64-22 + Quartzite + 20% RAP</td>
<td>7±1%</td>
<td>US 59 (Atlantia)</td>
<td>Good</td>
</tr>
<tr>
<td>4 Type F</td>
<td>Fine-graded (9.5 mm NMAS)</td>
<td>7.4% PG 76-22 + Sandstone</td>
<td>7±1%</td>
<td>US 271 (Paris)</td>
<td>Good</td>
</tr>
<tr>
<td>5 PFC</td>
<td>Open-graded (19 mm NMAS)</td>
<td>6.0% PG 76-22 + Igneous/limestone</td>
<td>20±2%</td>
<td>SH 121 (Paris)</td>
<td>Good</td>
</tr>
<tr>
<td>6 SMA</td>
<td>Gap-graded (19 mm NMAS)</td>
<td>6.0% PG 76-22 + Limestone</td>
<td>7±1%</td>
<td>IH 35 (Waco)</td>
<td>Very good</td>
</tr>
</tbody>
</table>

Legend: CAM = crack attenuating mix; PFC = permeable friction course; SMA = stone matrix asphalt; NMAS = nominal maximum aggregate size; RAP = reclaimed asphalt pavement material; PG = performance grade

Figure 4 Accumulated permanent strain and strain rate as a function of load cycles

Figure 5 Graphical comparisons of the FN test parameters
### Table 3 FN test results – PD parameters and statistics

<table>
<thead>
<tr>
<th>#</th>
<th>HMA Mix (Hwy)</th>
<th>HMA samples after testing</th>
<th>ID#</th>
<th>FN (cycles)</th>
<th>t(F)(min)</th>
<th>εp(F)</th>
<th>FN index, εp(F)/FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Type B (IH 35)</td>
<td>Sample#1 1 239</td>
<td>1</td>
<td>20.7</td>
<td>8 058</td>
<td>6.50</td>
<td>6.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 1 550</td>
<td>1</td>
<td>25.8</td>
<td>5 074</td>
<td>3.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 1 945</td>
<td>1</td>
<td>32.4</td>
<td>6 595</td>
<td>3.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean 1 578</td>
<td>1</td>
<td>26.3</td>
<td>6 576</td>
<td>4.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev 354</td>
<td>1</td>
<td>5.9</td>
<td>1 492</td>
<td>1.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV 22.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Type D (US 59)</td>
<td>Sample#1 1 485</td>
<td>1</td>
<td>24.7</td>
<td>12 034</td>
<td>8.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 960</td>
<td>960</td>
<td>16.0</td>
<td>8 787</td>
<td>9.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 1 205</td>
<td>1 205</td>
<td>20.1</td>
<td>6 962</td>
<td>5.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean 1 217</td>
<td>1</td>
<td>20.3</td>
<td>9 261</td>
<td>7.68</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev 263</td>
<td>1</td>
<td>4.4</td>
<td>2 569</td>
<td>1.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV 21.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>CAM (SH 121)</td>
<td>Sample#1 1 374</td>
<td>1</td>
<td>22.9</td>
<td>18 025</td>
<td>13.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 1 258</td>
<td>1 258</td>
<td>21.0</td>
<td>20 374</td>
<td>16.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 1 501</td>
<td>1 501</td>
<td>25.1</td>
<td>22 078</td>
<td>14.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean 1 378</td>
<td>1</td>
<td>23.0</td>
<td>20 159</td>
<td>14.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev 122</td>
<td>1</td>
<td>2.1</td>
<td>2 035</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV 8.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Type F (US 271)</td>
<td>Sample#1 5 074</td>
<td>5 074</td>
<td>84.6</td>
<td>13 952</td>
<td>2.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 4 583</td>
<td>4 583</td>
<td>76.4</td>
<td>13 138</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 2 760</td>
<td>2 760</td>
<td>46.0</td>
<td>17 440</td>
<td>6.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean 4 139</td>
<td>4 139</td>
<td>61.2</td>
<td>15 289</td>
<td>3.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev 1 219</td>
<td>1 219</td>
<td>21.5</td>
<td>3 042</td>
<td>2.44</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV 29.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>PFC (SH 121)</td>
<td>Sample#1 1 035</td>
<td>1 035</td>
<td>17.3</td>
<td>37 761</td>
<td>36.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 1 055</td>
<td>1 055</td>
<td>17.6</td>
<td>24 158</td>
<td>22.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 806</td>
<td>806</td>
<td>13.4</td>
<td>17 239</td>
<td>21.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean 931</td>
<td>931</td>
<td>15.5</td>
<td>26 386</td>
<td>27.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev 176</td>
<td>176</td>
<td>2.97</td>
<td>10 441</td>
<td>8.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV 18.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SMA (IH 35)</td>
<td>Sample#1 10 000</td>
<td>10 000</td>
<td>166.7</td>
<td>5 425*</td>
<td>&lt;0.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#2 10 000</td>
<td>10 000</td>
<td>166.7</td>
<td>5 425*</td>
<td>&lt;0.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sample#3 5 527</td>
<td>5 527</td>
<td>92.1</td>
<td>5 168</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean &gt;8 509</td>
<td>&gt;8 509</td>
<td>&gt;142</td>
<td>5 339</td>
<td>&lt;0.67</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stdev N/A</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>COV N/A</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Micro-strains measured at 10 000 cycles 166.7 minutes
of the coarse-graded Type B mix compared to the fine-graded Type F mix, may partially be attributed to a lower soft-grade PG 64-22 asphalt-binder used in the Type B mix. Note that as per Texas standards the Type B is typically designed as an intermediate-base mix with coarse-aggregate gradation predominately for providing rutting resistance and PD support to the pavement structure (TxDOT 2004).

In addition, it should be noted that while the TxDOT specification (TxDOT 2004) calls for use of PG 76-22 asphalt-binder for all the CAM mix-designs, the contractor mistakenly used a lower soft-grade PG 64-22 asphalt-binder on the SH 121 project. This could have partially contributed to this mix’s fifth ranking (or second poorest) performance based on the FN index. In the field (i.e. highway SH 121 project, Table 2), however, and as discussed subsequently, the 1.5 inch thick cover of the PFC surfacing mix, overlaying the 2 inch thick CAM layer, is theoretically expected to mitigate the potential PD problems of the CAM layer.

The inferior performance of the PFC mix (i.e. poorest in the ranking) was partly due to its high total AV content (i.e. 20±2%) and the fact that the FN test was conducted in an unconfined mode. Considering the fact that the PFC mix develops stone-on-stone contact for rutting resistance, similar to the SMA, and has generally performed well in the field (McDaniel et al. 2010), this result may suggest that the unconfined FN test is not ideal for the high AV content PFC mixes. These mixes should probably be tested in a confined FN test loading configuration.

**Discrimination and screening of HMA mixes based on the FN test parameters**

Figure 6 provides an assessment of the potential of the FN parameters to screen and discriminate the PD resistance of the mixes. A concept of discriminatory ratio (DR) was used, where the FN parametric value (i.e. FN (cycles), t(F), εp(F), and FN index) of a good mix is divided by that of the corresponding value of a poor mix. The larger the DR value in magnitude, the greater the difference between the mixes and the better the FN parameter to discriminate mixes. For this analysis, and based on the results in Figure 5 and Table 3, the SMA mix was adapted to be the best mix in terms of PD resistance and the CAM the poorest. The PFC mix was excluded from this analysis on account that the unconfined FN test appears to be unsuitable for these high AV mixes.

The DR values shown in Figure 6 indicate that the FN index provides a more distinctive differentiation and ranking capability of the mixes compared to the other parameters evaluated (i.e. high DR values in magnitude for the FN index). In particular, the ratio difference of the SMA-CAM (i.e. best versus poorest) with a DR value of 22 is visibly distinct. The εp(F) parameter on the other hand, with a DR of 1.0, shows no difference between the SMA versus Type B, and the Type F versus CAM, which is not the case with the other parameters evaluated. That is, the εp(F) parameter failed to sufficiently capture the differences in PD performance between these HMA mixes. Under this scenario (εp(F) parameter), the design engineer would not be able to readily differentiate the mixes when faced with a choice for design recommendations and may even end up selecting an inferior mix, because the laboratory PD performance prediction is hardly different.

**Statistical analysis of the FN test parameters**

Analysis of variance (ANOVA) and Tukey’s Honestly Significant Differences (HSD) multiple comparison procedure (Tukey 1953) were used to statistically investigate the ability of the FN test parameters to differentiate the PD resistance of the HMA mixes. The results of these analyses, at a 95% confidence level, are summarised in Table 4.

The statistical interpretation of the results in Table 4, for example the last column for the FN index, is as follows: the CAM mix has the highest FN index value and is statistically listed in Group A (poorest in terms of the PD resistance), whereas the SMA – with the lowest FN index value – is categorised as Group C (best in terms of the PD resistance). The Type F, B and D mixes on the other hand, have statistically indifferent FN index values that lie in-between Groups A and C, and are subsequently listed in the same Group B.

At a 95% confidence level, Table 4 shows that only the FN index is able to statistically differentiate the SMA (best) and the CAM (poorest) from the other mixes. Statistically, the FN index is inferring that the Type F, B and D mixes have insignificantly different PD resistance properties and that SMA and CAM mixes have significantly different PD resistance properties. By contrast, the FN (cycles) and t(F) parameters are unable to capture any statistical difference in the PD resistance potential among the Type F, Type B, Type D and CAM mixes, but are able to single out the SMA as being significantly different. The εp(F) parameter, on the other hand, considers the SMA to be statistically similar to the Type B and D mixes, while the Type F and CAM mixes are categorised as being statistically different. Thus, only the FN index is able to provide a reasonable statistical differentiation of the mixes.

For comparative studies of this nature, the εp(F) parameter would normally be

---

**Table 4 ANOVA and Tukey’s HSD test analyses**

<table>
<thead>
<tr>
<th>HMA Mix</th>
<th>FN (cycles)</th>
<th>t(F)</th>
<th>εp(F)</th>
<th>FN Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type F</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Type B</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
</tr>
<tr>
<td>Type D</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
</tr>
<tr>
<td>CAM</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>SMA</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>C</td>
</tr>
</tbody>
</table>

**Figure 6** Discriminatory ratios computed for the FN test parameters

![Figure 6](image-url)
analysed and interpreted in conjunction with FN (cycles) and/or t(F) or vice versa. Otherwise, one of the parameters must be held constant. That is, to meaningfully compare these mixes using the \( \varepsilon_p(F) \) parameter for instance, they must be evaluated at the same FN (cycles) or t(F) level. As evident in Table 3, the problem is that tertiary failure occurs at different FN (cycles) and \( \varepsilon_p(F) \) values for different mixes. Therefore, this logic would not apply unless the tertiary failure criterion is ignored. As supported by the preceding results, the FN index, which is a function of \( \varepsilon_p(F) \) and FN (cycles), takes care of this quandary.

Looking at the preceding analyses and discussions, it is evident that while some parameters may give a similar mix performance ranking (i.e. FN [cycles] versus FN index), the FN index exhibits statistically superior discrimination and screening capabilities for the HMA mixes evaluated in this study. However, evaluating more mixes and exploring means to minimise variability in the determination of this parameter will further serve to substantiate the FN index concept.

**FN test parameters and HMA tertiary flow occurrence**

As indicated in Table 3 and in the preceding discussions, two specimens (samples # 1 and 2) of the SMA mix lasted up to 10 000 load cycles without any tertiary failure, i.e. flow. For the FN test conditions prescribed in this study, if flow or tertiary failure does not occur within the specified 10 000 cycles after 166.7 minutes of FN testing time, the following two options are proposed:

a. The FN index should be calculated at 10 000 cycles as a function of the actual measured micro-strains at 10 000 cycles (166.7 minutes) and the final FN index result should be reported as being less than this calculated value, i.e. “FN index < Calculated Value”. For the two SMA samples #1 and 2 in Table 3, the FN index calculated at 10 000 cycles corresponding to the measured micro-strains of 5 425 at 10 000 cycles is 0.54; so, the FN index results would be reported as “FN index < 0.54” (see Table 3).

b. Perform extrapolative data analysis to estimate the flow parameters using the models listed in Table 1, and then compute the FN index. In the case of the SMA sample #2 for instance, the extrapolated “flow parameters” would be as follows: FN = 544 326 cycles; t(F) = 9 072.1 minutes; \( \varepsilon_p(F) \) = 18 713 micro-strains. So, the estimated and corresponding FN index would be “FN index = 0.03 (extrapolated)”. However, extrapolation inherently introduces some uncertainties in the results obtained (i.e. there is no complete certainty to indicate that sample #2 would actually have reached flow at 544 326 cycles and 9 072.1 minutes). Therefore, these researchers recommend adapting the former approach in situations where flow does not occur within the 10 000 cycles test conditions. Thus, the FN index results for the SMA mix would be reported and interpreted as follows:

1. FN index (sample # 1) < 0.54
2. FN index (sample # 2) < 0.54
3. FN index (sample # 3) = 0.94
4. FN index (SMA) < 0.67

**Comparison of FN test and HWTT results**

The average HWTT results based on three replicate test sets per mix type are shown in Figure 7, and the resistance to PD of the HMA mixes ranked as follows: Type D (4.36 mm) → SMA (4.61 mm) → Type F (5.45 mm) → PFC (7.60 mm) → Type B (12.90 mm) → CAM (18.00 mm; poorest). The difference in the ranking compared to the FN test results is partially attributed to the differences in the loading configuration and high sample confinement in the HWTT setup, unlike in the unconfined FN test. Even the high AV content PFC mix outperformed the Type B mix in the HWTT, which is not the case with the unconfined FN test. The possibility of moisture damage (i.e. stripping of the Type B mix) could have been another factor, with the inflexion point seemingly occurring after 10 000 HWTT load passes in Figure 7. In either case, however, the CAM mix still remains at the bottom of the ranking. Lower asphalt-binder PG grade, high asphalt-binder content, and fine aggregate gradation (Table 2) could be some of the contributing factors for this particular result.

Like for the FN index, a DR analysis of the HWTT data also exhibited a distinctive difference in the laboratory rutting performance between the SMA and CAM mix, with a DR value of 4.0. However, the DR trends for the other mix comparisons did not correlate with the FN test results – possibly on account of the differences in the test loading configuration and sample confinement as previously discussed. The extreme HWTT sample confinement may be over-scoring the true PD performance of some of these mixes. As lately experienced in the State of Texas, some HMA mixes that passed the HWTT in the laboratory are failing in the field. These field failures could be related to poor screening during the mix-design process, partially due to this high specimen confinement. Excluding the PFC mix, the HMA mix ranking comparison is as shown in Table 5.

From Table 5, both test methods rank the CAM as the least PD-resistant mix, which like the Type B mix, would have been

---

**Table 5** HMA mix ranking comparison

<table>
<thead>
<tr>
<th>HMA mix</th>
<th>FN index (micro-strains/cycles)</th>
<th>HWTT rut depth (mm)</th>
<th>FN index ranking</th>
<th>HWTT ranking</th>
<th>HWTT decision level</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMA</td>
<td>0.67</td>
<td>4.60</td>
<td>1</td>
<td>2</td>
<td>Accept</td>
</tr>
<tr>
<td>Type F</td>
<td>3.98</td>
<td>5.50</td>
<td>2</td>
<td>3</td>
<td>Accept</td>
</tr>
<tr>
<td>Type B</td>
<td>4.39</td>
<td>12.90</td>
<td>3</td>
<td>4</td>
<td>Reject</td>
</tr>
<tr>
<td>Type D</td>
<td>7.68</td>
<td>4.36</td>
<td>4</td>
<td>1</td>
<td>Accept</td>
</tr>
<tr>
<td>CAM</td>
<td>14.67</td>
<td>18.01</td>
<td>5</td>
<td>5</td>
<td>Reject</td>
</tr>
</tbody>
</table>

---

**Figure 7** HWTT graphical results
rejected under the HWTT screening criterion (i.e. rut depth greater than 0.5 inches). As evident in Table 5, the values of the HWTT rut depths for the SMA, Type F and Type D are statistically indifferent and could hypothetically be equated to the FN index ranking. The Type D mix, although having a lower asphalt-binder PG grade (PG 64-22; see Table 2) compared to its companions SMA and Type F (with PG 76-22), is composed of 20% RAP that adds to its PD resistance properties.

**PRELIMINARY CORRELATIONS WITH FIELD PERFORMANCE DATA**

As predicated by both the FN and HWTT laboratory test results, the 1½ inch thick Type D surfacing mix on the US 59 highway has performed satisfactorily after being in service for over two years, while subjected to an average daily truck traffic (ADTT) of 1 500 vehicles at an average vehicle speed of 72.6 mph. As shown in Figure 8, the measured average surface rutting in 2012 was only 0.19 inches; thereby substantiating the well-averaged surface rutting in 2012 was only 0.19 inches; thereby substantiating the service life of the pavement (i.e. continued conventional traffic loading, increase in traffic volume, overloading, subsequent summers, etc). However, the ongoing long-term performance monitoring of these highway sections will aid to further verify and validate these laboratory test results.

It should be noted that PD of HMA mixes, unlike other distresses, most often occurs in the early life of the pavement just after construction under traffic densification, particularly under extreme temperatures. Nonetheless, this is not to indicate that this distress will not occur in the design and service life of the pavement (i.e. continued conventional traffic loading, increase in traffic volume, overloading, subsequent summers, etc). However, the ongoing long-term performance monitoring of these highway sections will aid to further verify and validate these laboratory test results.

**COMPARISON OF THE LABORATORY TEST METHODS**

Table 6 provides a subjective itemisation of the characteristic attributes of the two rutting test methods based solely on the HMA mixes evaluated in this study and on the authors’ experience with these test methods.

---

**Table 6 HWTT and FN test methods – merits and demerits**

<table>
<thead>
<tr>
<th>Test</th>
<th>Advantages and applications</th>
<th>Limitations and challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>HWTT</td>
<td>1. Simplicity and practicality.</td>
<td>1. Cannot readily generate HMA material properties such as modulus for structural design and mechanistic-empirical analyses.</td>
</tr>
<tr>
<td></td>
<td>2. Can readily test both laboratory made samples and field cores.</td>
<td>2. High sample confinement in molds during testing that may at times negatively impact the test results and rutting performance of the HMA mix.</td>
</tr>
<tr>
<td></td>
<td>3. Reasonable test time (≤ 8 hours).</td>
<td>3. Inability to sufficiently capture the shear resistance characteristics of the HMA mix.</td>
</tr>
<tr>
<td></td>
<td>4. Repeatability and low variability in results (COV ≤ 10%) (Zhou &amp; Scullion 2002).</td>
<td>4. Test was run at a single temperature (50°C), so there is need to explore multiple temperatures that are reflective of field temperatures.</td>
</tr>
<tr>
<td></td>
<td>5. Rutting and moisture damage (stripping) assessment.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6. Applicable for daily routine mix-designs.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7. Applicable for HMA mix screening and acceptance.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8. Predicts performance and provides reasonable correlation to field performance.</td>
<td></td>
</tr>
<tr>
<td>FN test</td>
<td>1. Reasonable test time (&lt; 3 hours per sample).</td>
<td>1. Sample fabrication process is both laborious and long.</td>
</tr>
<tr>
<td></td>
<td>2. Multiple data outputs (i.e. FN, t(F), εp(F), FN index).</td>
<td>2. Cannot readily test thin field cores (must use prismatic specimens but only for layer thickness equal to or greater than 50 mm).</td>
</tr>
<tr>
<td></td>
<td>3. HMA mix material properties for structural design.</td>
<td>3. Problematic maintaining LVDT studs at high temperatures.</td>
</tr>
<tr>
<td></td>
<td>4. HMA mix rutting performance prediction.</td>
<td>4. Unconfined FN test setup is not ideal for porous mixes like the PFC mix, i.e. allows for lateral or horizontal defamation (bulging of the specimens).</td>
</tr>
<tr>
<td></td>
<td>5. Reasonable variability and repeatability (an average COV of 33% based on this study).</td>
<td>5. Test was performed at a single temperature (50°C), so there is need to explore multiple temperatures that are reflective of field temperatures.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires experienced operator.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Requires UTM or MTS equipment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Still requires validation with field data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Variability can still be optimised further.</td>
</tr>
</tbody>
</table>
Table 7 Example of variability in the HWTT test results

<table>
<thead>
<tr>
<th>HMA mix = Type D (US 59)</th>
<th>Sample ID#</th>
<th>HWTT rut depth (mm) @ 20 000 load passes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample#1</td>
<td>4.60</td>
</tr>
<tr>
<td></td>
<td>Sample#2</td>
<td>4.19</td>
</tr>
<tr>
<td></td>
<td>Sample#3</td>
<td>4.29</td>
</tr>
<tr>
<td>Avg</td>
<td></td>
<td>4.36</td>
</tr>
<tr>
<td>COV (%)</td>
<td></td>
<td>4.85</td>
</tr>
</tbody>
</table>

Overall, while the HWTT is the simplest, most practical and readily applicable for routine daily mix-design and screening, its major challenges include the adaptability to generate multiple HMA material properties (e.g. modulus) and high specimen confinement that tends to over-score the PD resistance performance of the mixes. The FN shorter test time, as compared to that of the HWTT, means that the test is both cost-effective and applicable for daily routine use, particularly with the FN index parameter that exhibited potential to sufficiently discriminate and screen mixes. Inability to readily test thin field cores and the need for field validation are some of the key challenges associated with the FN test. In addition, the unconfined FN test setup appears unsuitable for testing PFC mixes.

Compared to the FN variability and repeatability, with an overall average COV of 33% in the test results (Table 3), the HWTT exhibited statistical superiority with COV values less than 10% in the test results; see the example in Table 7 (Walubita et al 2012; Zhou & Scullion 2002).

As evident in Table 3, some of the HMA mixes (Type B, Type F and PFC) have FN parameters and statistics with COV values that are unacceptably on the higher side (i.e. greater than 30% in the case of the FN index). Although HMA, due to its visco-elastic nature, is generally associated with high variability at high test temperatures such as 50°C (particularly for unconfined tests like the FN), this high variability in Table 3 is primarily due to some outliers that may warrant exclusion from the overall analysis of the test results. Based on the FN index parameter in Table 3, Sample #1 (Type B), Sample #3 (Type D), Sample #3 (Type F), and Sample #1 (PFC) would be considered as outliers. If these outliers are discarded from the analysis, the statistics would be as shown in Table 8, which is considered to be reasonably acceptable and comparable to the HWTT. These statistics are also consistent with the AASHTO TP 79-12 recommendations for FN testing using the Asphalt Mixture Performance Tester (AMPT) (AASHTO 2012).

Based on the test methods and HMA mixes evaluated, the key findings, conclusions and recommendations drawn from this study are summarised as follows:

- The FN index, computed from the FN test data, offers promising potential as a parameter to use for differentiating and screening in the laboratory the resistance to PD of HMA mixes during the HMA mix-design stage. This fact and the shorter test time of the FN test, as compared to the HWTT, indicates that the FN test (i.e. based on the FN index computation) offers promise for routine HMA mix-design in the laboratory as a surrogate and/or supplementary rutting test to the HWTT. However, caution must be exercised to watch out for outliers as these have the potential to statistically distort the final FN index results. Nonetheless, there is still a need to test more HMA mixes to supplement and validate these results/findings.

- The conventional parameters computed based on the FN test – including the FN (cycles), \( t(F) \) and \( \varepsilon_p(F) \) – as individual parameters do not provide an effective, nor statistically significant, differentiation and screening of resistance to PD for the HMA mixes that were evaluated in this study. Therefore, application of these parameters for routine HMA mix-design and screening of PD resistance should be approached with caution. Unconfined FN testing should not be applied to PFC mixes as this tends to underestimate the potential PD resistance of these mixes, due predominately to their high AV content nature.

Table 8 Statistics of the FN index results after discarding the outliers

<table>
<thead>
<tr>
<th></th>
<th>SMA (SH 121)</th>
<th>Type F (US 271)</th>
<th>Type B (IH 35)</th>
<th>Type D (US 59)</th>
<th>CAM (SH 121)</th>
<th>Type F (US 271)</th>
<th>PFC (SH 121)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg</td>
<td>3.33</td>
<td>8.63</td>
<td>14.67</td>
<td>2.81</td>
<td>22.15</td>
<td>&lt; 0.67</td>
<td>N/A</td>
</tr>
<tr>
<td>Stddev</td>
<td>0.08</td>
<td>0.74</td>
<td>1.54</td>
<td>0.08</td>
<td>1.06</td>
<td>0.08</td>
<td>N/A</td>
</tr>
<tr>
<td>COV (%)</td>
<td>2.49%</td>
<td>8.61%</td>
<td>10.49%</td>
<td>3.02%</td>
<td>4.79%</td>
<td>4.79%</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 9 Ranking of the HMA mix based on the FN index parameter

<table>
<thead>
<tr>
<th></th>
<th>SMA (SH 121)</th>
<th>Type F (US 271)</th>
<th>Type B (IH 35)</th>
<th>Type D (US 59)</th>
<th>CAM (SH 121)</th>
<th>PFC (SH 121)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FN index ranking</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Table 3 FN index</td>
<td>&lt; 0.67</td>
<td>3.98</td>
<td>4.39</td>
<td>7.68</td>
<td>14.67</td>
<td>27.20</td>
</tr>
<tr>
<td>Table 8 FN index</td>
<td>&lt; 0.67</td>
<td>2.81</td>
<td>3.33</td>
<td>8.63</td>
<td>14.67</td>
<td>22.15</td>
</tr>
<tr>
<td>Table 3 FN index</td>
<td>N/A</td>
<td>61.30%</td>
<td>41.70%</td>
<td>22.50%</td>
<td>10.49%</td>
<td>30.61%</td>
</tr>
<tr>
<td>Table 8 FN index</td>
<td>N/A</td>
<td>3.02%</td>
<td>2.49%</td>
<td>8.61%</td>
<td>10.49%</td>
<td>4.79%</td>
</tr>
</tbody>
</table>

CONCLUSIONS AND RECOMMENDATIONS

In order to compare two PD-performance predictive laboratory tests, namely the FN test and HWTT, six types of HMA mixes were evaluated. The laboratory test data was accordingly supplemented with some limited field performance data from in-service highways where the same HMA mixes were used.
For baseline comparison with the HWTT, whose samples are tested under a tightly confining mold (i.e. no lateral or horizontal deformation), future FN tests should also be conducted in confined loading mode with no lateral deformation (i.e. bulging) of the samples. In general, test conditions should, as far as possible, be close to one another in comparative studies of this nature. Furthermore, conducting the confined FN testing will also help to further validate the laboratory performance of HMA mix sample confinement for future studies.

For the highway sections evaluated, the FN index exhibited a promising correlation with field performance data. However, this was only limited data and as such, long-term performance monitoring with more field sections is strongly recommended to further validate the findings of this study and to establish some screening criterion/threshold values for the FN index, as well as future applications of the FN index for pavement design and analysis.

Overall, this study showed the necessity of caution when selecting HMA rutting tests; depending on the specific needs, each test method has its own merits and demerits. In general, the following are some of the key challenges associated with selecting the appropriate laboratory rutting test: sample fabrication, simplicity and practicality of the test, cost-effectiveness, reasonable test time, applicability for routine HMA mix-design and screening, ability to generate multiple data, and correlation with field performance.

Specifically for the FN test, some of the primary challenges to be addressed in future research studies include the following:

- HMA mix sample configuration and preparation
- HMA mix sample confinement for further comparisons with the HWTT
- Exploring multiple test temperatures
- Addressing and minimising variability in the FN index computations
- Exploring the computation of the FN cycles to a predefined plastic strain limit through modification of regression constants in the strain-cycle model (Equation 1).
- Evaluating and testing more HMA mixes
- Validation with field performance data.

**ACKNOWLEDGEMENTS**

The authors thank TxDOT and the Federal Highway Administration (FHWA) for their financial support and all those who helped during the course of this research work. In particular, a special word of thanks goes to the following individuals: Jacob Hoefnner (TTI), Jason Huddleston (TTI), Abu N M Faruk (TTI), Dr Geoffrey S Simate (University of the Witwatersrand) and Charles Mushota (RDA).

**DISCLAIMER**

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

**REFERENCES**

Abdallah, I & Nazarial, S 2011. Strategies to improve performance of PFC mixes. With more test data, other options would be to explore the use of FN (cycles) to a predefined plastic strain limit (or set value) for easy comparison with the HWTT results.

For the highway sections evaluated, the FN index exhibited a promising correlation with field performance data. However, this was only limited data and as such, long-term performance monitoring with more field sections is strongly recommended to further validate the findings of this study and to establish some screening criterion/threshold values for the FN index, as well as future applications of the FN index for pavement design and analysis.

Overall, this study showed the necessity of caution when selecting HMA rutting tests; depending on the specific needs, each test method has its own merits and demerits. In general, the following are some of the key challenges associated with selecting the appropriate laboratory rutting test: sample fabrication, simplicity and practicality of the test, cost-effectiveness, reasonable test time, applicability for routine HMA mix-design and screening, ability to generate multiple data, and correlation with field performance.

Specifically for the FN test, some of the primary challenges to be addressed in future research studies include the following:

- HMA mix sample configuration and preparation
- HMA mix sample confinement for further comparisons with the HWTT
- Exploring multiple test temperatures
- Addressing and minimising variability in the FN index computations
- Exploring the computation of the FN cycles to a predefined plastic strain limit through modification of regression constants in the strain-cycle model (Equation 1).
- Evaluating and testing more HMA mixes
- Validation with field performance data.

**ACKNOWLEDGEMENTS**

The authors thank TxDOT and the Federal Highway Administration (FHWA) for their financial support and all those who helped during the course of this research work. In particular, a special word of thanks goes to the following individuals: Jacob Hoefnner (TTI), Jason Huddleston (TTI), Abu N M Faruk (TTI), Dr Geoffrey S Simate (University of the Witwatersrand) and Charles Mushota (RDA).

**DISCLAIMER**

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

**REFERENCES**

Abdallah, I & Nazarial, S 2011. Strategies to improve performance of PFC mixes. With more test data, other options would be to explore the use of FN (cycles) to a predefined plastic strain limit (or set value) for easy comparison with the HWTT results.

For the highway sections evaluated, the FN index exhibited a promising correlation with field performance data. However, this was only limited data and as such, long-term performance monitoring with more field sections is strongly recommended to further validate the findings of this study and to establish some screening criterion/threshold values for the FN index, as well as future applications of the FN index for pavement design and analysis.

Overall, this study showed the necessity of caution when selecting HMA rutting tests; depending on the specific needs, each test method has its own merits and demerits. In general, the following are some of the key challenges associated with selecting the appropriate laboratory rutting test: sample fabrication, simplicity and practicality of the test, cost-effectiveness, reasonable test time, applicability for routine HMA mix-design and screening, ability to generate multiple data, and correlation with field performance.

Specifically for the FN test, some of the primary challenges to be addressed in future research studies include the following:

- HMA mix sample configuration and preparation
- HMA mix sample confinement for further comparisons with the HWTT
- Exploring multiple test temperatures
- Addressing and minimising variability in the FN index computations
- Exploring the computation of the FN cycles to a predefined plastic strain limit through modification of regression constants in the strain-cycle model (Equation 1).
- Evaluating and testing more HMA mixes
- Validation with field performance data.
SUBMISSION PROCEDURES AND REQUIRED FORMAT

| Hard copies: Hard copies should only be submitted if the author is unable to submit electronically, for whatever reason. In such a case only one unstapled copy, together with an electronic copy on CD, should be posted to the Chief Executive Officer of SAICE, Private Bag X200, Halfway House, 1685, or delivered to Building 19, Thornhill Office Park, Bekker Street, Midrand. Hard copies should be printed on one side of A4 paper.
| Electronic submission: Manuscripts should be e-mailed to the editor at 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 not printed in colour. 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).

POLICY REGARDING LANGUAGE AND ORIGINALITY OF SUBMITTED ARTICLES

| Language: Manuscripts should preferably be presented in English, as the journal is distributed internationally. Articles submitted in any of the other official South African languages should be accompanied by an expanded abstract in English.
| Original work: Papers and technical notes must be original contributions. Authors must affirm that submitted material has not been published previously, is not under consideration for publication elsewhere and will not be submitted elsewhere while under consideration by the SAICE Journal Editorial Panel. It is the responsibility of the authors to ensure that publication of any paper in the journal will not constitute a breach of any agreement or the transgression of any law. The corresponding author should confirm that all co-authors have read and approved the manuscript and accept these conditions. Authors are responsible for obtaining permission to publish experimental data and other information that may be confidential or sensitive. Authors are also responsible for obtaining permission from copyright owners when reproducing material that had been published elsewhere.

Guidelines for the preparation of papers and technical notes

The Journal of the South African Institution of Civil Engineering is published bimonthly in April and October. Articles submitted for publication are reviewed by a panel of referees under the guidance of the SAICE Journal Editorial Panel, a sub-committee of the SAICE Editorial Panel. When preparing articles for publication, authors should please take note of the following and comply with the guidelines as set out below.

CLASSIFICATION OF ARTICLES CONSIDERED FOR PUBLICATION

- Technical papers are well-researched, in-depth, fully referenced technical articles not exceeding 6 000 words in length (excluding tables and illustrations). Related papers that deal with ‘softer sciences’ (e.g. education, social upliftment, etc) are accepted if they are of a technical nature and of particular interest to the civil engineering profession. The latter type of paper will be subject not only to peer-review by civil engineers, but also to review by non-engineering specialists in the field covered by the paper.
- Technical notes are short, fully referenced technical articles that do not exceed 2 000 words. A typical technical note will have limited scope often dealing with a single technical issue of particular importance to civil engineering. An article incorrectly submitted as a technical paper can, with the author’s agreement, be re-categorised by the Panel as a technical note.
- Reviews and case studies are considered for publication as either technical papers or technical notes on condition that they are the original work of the author and will assist the reader with the understanding, interpreting or applying of the subject under review.
- Discussion on published articles is welcomed up to six months after publication. The length of discussion contributions is limited to 1 500 words. Where appropriate, discussion contributions will be subject to the normal reviewing process and will be forwarded to the authors of the original article for reply.

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

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-drafted figures will not be accepted.

Mathematical expressions: Equations should be presented in a clear form which can be easily read by non-mathematicians. Symbols should be typed using the Times New Roman symbol set. Equations should be in italics. Each equation should appear on a separate line and be numbered consecutively.

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
Headings 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.

Photographs 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.