ISSN 1991-1696

SAIEE

Africa Research Journal

Research Journal of the South African Institute of Electrical Engineers

Incorporating the SAIEE Transactions

September 2014
Volume 105 No. 3

www.saiee.org.za

SAIEE AFRICA RESEARCH JOURNAL

(SAIEE FOUNDED JUNE 1909 INCORPORATED DECEMBER 1909)
AN OFFICIAL JOURNAL OF THE INSTITUTE

ISSN 1991-1696

Secretary and Head Office
Mrs Gerda Geyer
South African Institute for Electrical Engineers (SAIEE)
PO Box 751253, Gardenview, 2047, South Africa
Tel: (27-11) 487-3003
Fax: (27-11) 487-3002
E-mail: researchjournal@saiee.org.za

Published by
SAIEE Publications (Pty) Ltd, PO Box 751253, Gardenview, 2047,
Tel. (27-11) 487-3003, Fax. (27-11) 487-3002,
E-mail: SAIEEPublications@saiee.org.za

Additional reviewers are approached as necessary

ARTICLES SUBMITTED TO THE SAIEE AFRICA RESEARCH JOURNAL ARE FULLY PEER REVIEWED PRIOR TO ACCEPTANCE FOR PUBLICATION

The following organisations have listed SAIEE Africa Research Journal for abstraction purposes:
INSPEC (The Institution of Electrical Engineers, London); 'The Engineering Index' (Engineering Information Inc.)

VOL 105 No 3
September 2014

SAIEE Africa Research Journal

LANDMINE DETECTION BY MEANS OF GROUND PENETRATING RADAR: A MODEL-BASED APPROACH

P.A. van Vuuren *

* *School of Electrical, Electronic and Computer Engineering, North-West University, Private Bag X6001, Potchefstroom, 2520, South-Africa. E-mail: pieter.vanvuuren@nwu.ac.za*

Abstract: The presence of landmines poses a worldwide humanitarian problem. Often, these mines are difficult to detect with metal detectors. Ground penetrating radar (GPR) is a promising technology for the detection of landmines with low metal content. Automatic landmine detection typically consists of two steps, namely preprocessing (or clutter removal) and classification. In this paper the clutter removal algorithm consists of a nonlinear frequency domain filter followed by principal component based filtering. Principal component analysis is performed in the frequency domain to build a background model for the clutter. The latter model is removed from the observed data in the log-frequency domain in order to preserve the phase component of the spectrum. Finally, the data is normalized and transformed to the time domain. The results presented in this paper show a marked improvement in the ability to remove general background clutter. Classification is performed on the basis of the prediction performance of neural network time-series models of the various classes of GPR responses. The classification system can correctly identify the position of metal anti-tank (AT) mines. It can also recognize specific examples of low metal AT and (anti-personnel) AP mines, but does have a low generalization ability for such mines.

Key words: Landmine detection, ground penetrating radar, neural networks, principal component analysis, clutter removal

1. INTRODUCTION

Landmines and other explosive remnants of war pose a debilitating threat to innocent civilians long after hostilities have ceased. World-wide, 4191 people (of which 75 % were civilians) were killed or injured during 2010 by landmines and other explosive remnants of war [1]. Each one of these casualties is one too many. Furthermore, their mere suspected presence in certain regions prevents agricultural activities, inhibit people's freedom of movement and hampers reconstruction and development of war-torn societies [1]. Clearance of contaminated areas is a costly exercise. The effectiveness of mine clearance operations obviously hinge on accurate detection of landmines.

Landmine detection is a vexing problem due to the low metal content of some anti-tank and anti-personnel mines. Furthermore, contaminated areas often are littered with other general metallic debris, causing a high false alarm rate in metal detectors [2]. Metal detectors are therefore unsuited to detect the broad range of landmines that can be encountered in practice.

Ground penetrating radar (GPR) is a proven technology for subsoil research [3]. The radar reflections received by a GPR system are a function of all subsoil dielectric changes. As such, GPR systems can in principle detect a wider range of mines compared to metal detectors. Whether this is true in practice remains to be seen.

GPR scans are typically performed from a moving platform such as an armoured vehicle. In order to cover as large an area of ground as possible, GPR antennas are often arranged in arrays of identical antennas. GPR antenna arrays make it possible to reconstruct a three-dimensional image of the subsurface features. This 3D image is commonly known as a C-scan. The axis along which the GPR antennas are arranged is called the *cross-track* direction, since the entire array is moving in the *down-track* direction. The radar reflections obtained by a single GPR antenna at a specific (x, y) position form a 1-dimensional signal, known as an A-scan. A B-scan refers to a set of A-scans in either the down-track (x) direction or in the cross-track (y) direction.

In this paper a stepped-frequency, continuous-wave (SFCW) GPR antenna system is used to obtain a GPR C-scan. At each (x, y) position a sequence of GPR sinewaves is generated with a constant amplitude and stepwise increase in frequency. In this manner the A-scan at position (x, y) can (after some processing) be interpreted as the averaged frequency response of the ground to a wide band of radar frequencies (100 MHz to 1.8 GHz).

The term commonly used in the GPR literature for various phenomena that obscure the presence of valid targets (landmines) is *clutter*. Clutter increases the false alarm rate of a landmine detection system [4]. This is because clutter often dominates the observed data (especially in the frequency domain) [5].

Some researchers define clutter as all phenomena in an A-scan that remain constant over a large section of the corresponding down-track B-scan [6]. Such phenomena include antenna effects [7], certain instances of layers in the soil with differing electromagnetic properties [8], as well as the ubiquitous air-surface reflection (also referred to as *ground bounce*) [4]. Other forms of distracting phenomena are however more transient in nature (e.g. rocks and roots [8] as well as soil roughness resulting in diffuse scattering [7]).

In landmine detection via GPR the preprocessing step is therefore primarily occupied with the removal of as much clutter as possible to facillitate successful landmine detection. Clutter removal is then followed by the next step which typically involves feature extraction and classification of feature vectors.

The data obtained by a SFCW GPR array can be interpreted as the averaged frequency response of the soil at a particular location. In the time domain this corresponds approximately to the impulse response of the soil. In this paper, the first in a two-part series, clutter removal is performed in the frequency domain. In the second paper (Landmine detection by means of ground penetrating radar: a rule-based approach) clutter removal is performed in the time domain. Classification is performed in the time domain in both papers. A model-based approach is followed to perform classification of GPR data in this paper, whereas a rule-based classifier is presented in the second paper.

This paper contributes to both the clutter removal stage and the final classification stage of the GPR landmine detection problem. A novel algorithm is developed in section 2 for the removal of clutter from frequency domain GPR data. After the resulting A-scans are transformed from the frequency to the time domain, classification is performed in a manner novel to the GPR literature by means of time-series models derived for various classes of data. Artificial neural networks are used for these time-series models. Section 3 discusses the design of the time-series models. The performance of the clutter removal algorithm is presented in section 4, while the ability of the time-series based classifier to detect a variety of mines is the topic of section 5. The paper is concluded in section 6.

2. CLUTTER REMOVAL ALGORITHM

2.1 Previous work performed on clutter removal

The fundamental objective of clutter removal is to accentuate the radar reflections made by landmines while at the same time diminishing the scattering due to soil features. Numerous clutter removal techniques exist, but they can all be classified into the following three categories.

- *Time or range gating techniques*. The simplest response to ground bounce is to merely discard the relevant sections of an A-scan [5]. This is known as time gating or range gating. The computational cost of time gating is very low. Unfortunately, this technique only removes a portion of the reflections due to ground bounce, since ground bounce reflections also occur deeper underground [9]. Furthermore this technique is only applicable to time-domain data. Lastly, time gating can't be used with mines that are buried shallowly [9].
- *Classification-based techniques*. This is a loose collection of techniques that avoid explicit clutter removal in a variety of ways. One example entails extracting features that are insensitive to the presence of clutter [8]. A related approach involves performing statistical hypothesis testing on features based on linear prediction models of the data [10]. Such techniques highlight the blurry boundary between clutter removal and prescreening algorithms [11].
- *Background removal techniques*. Most clutter removal techniques essentially involve modelling the background clutter and subtracting it from the raw GPR data. This is most commonly accomplished one B-scan at a time (either in the cross-track or down-track direction). Models for the background clutter range from simplistic (e.g. the mean value at each depth in the down-track direction [6]) to first principles models based on the various physical processes responsible for clutter formation [7]. In fact, any technique that explicitly models the background clutter resorts under the banner of background removal techniques, whether via black-box system identification [9] or principal component analysis of time-series data [6].

Of the three above mentioned approaches, background removal techniques represent the best balance between accuracy and simplicity and is therefore further pursued in this paper.

Before delving into the details of the improved clutter removal algorithm we'll first pause a moment at a few fundamental assumptions concerning the process by which GPR reflections occur. GPR reflections occur when the electromagnetic properties (predominantly the permittivity) of the medium through which the radar signal is travelling change abruptly. These changes in permittivity occur at the boundaries of different objects (e.g. pebbles, roots, tunnels and mines) as well as at the boundaries of different soil layers. Some of these variations (specifically changes in the soil composition) can occur quite gradually. Most of the variations can only be regarded as random events and can't be predicted in advance. In fact, it is virtually impossible to construct an accurate and practically useful first principles model for the clutter in GPR data [12]. The only practical solution for building a model of the background scattering is therefore by means of adaptive numerical methods.

A moving average (or median) filter is a prime contender for a model of the background scattering, since it depends on minimal assumptions regarding the data and is inherently adaptive [6]. Typically, a background model is obtained by implementing a separate moving average filter at each depth index. The drawback of such an approach is however that the time axis of the A-scans is only an inaccurate estimation of depth. This is due to the fact that the propagation velocity of the radar signal is dependent on the dielectric properties of the soil, which in turn is variable [13], [12].

The safest option is therefore to base a background model on an entire B-scan, either in the down-track or the cross-track direction. An entire B-scan can be regarded as an image (i.e. a vertical slice through the ground). Interpreting GPR data as images opens up the possibility of modelling the background clutter by means of principal components [14]. Principal component based clutter removal has been applied successfully in both the time- and frequency domain. The calculation of the principal components can also be performed recursively rendering an adaptive clutter removal algorithm [6].

2.2 Algorithm overview

The clutter removal process presented in this paper consists of a nonlinear frequency domain filter followed by principal component analysis also conducted in the frequency domain. The aim of the principal component analysis is to build a model of the background clutter. Clutter removal then entails that the background model is subtracted from the observed GPR spectra. Finally, the data is normalized prior to being transformed to the time domain for further processing by feature extraction algorithms. The entire procedure is as follows:

1. Removal of narrow-band disturbances via nonlinear filtering. During this step spurious peaks situated in the vicinity of 945 MHz are removed.
2. Form a background model by means of frequency domain principal component analysis and subtract it from the original spectrum in the logarithmic domain. Calculations are furthermore performed in the log-domain in order to prevent the phase component of the spectrum being amplified.
3. Normalize the resultant spectra.
4. Application of the inverse Fourier transform to obtain decluttered time-domain A-scans.

Each of the above mentioned steps will now be discussed in greater detail.

2.3 Spike removal via nonlinear filtering

The observed GPR spectra contain a number of distinctive peaks. These peaks have the following properties:

- They only occur at certain frequencies.
- The specific frequencies of these ”spikes” exhibit a small variation from one A-scan to the next, but remain relatively constant throughout the entire C-scan.
- The specific ”spike” frequencies are a function of the bandwidth of the stepped-frequency continuous wave signal of the antenna array.
- A large variation occurs in the amplitude of the peaks.
- A small variation occurs in the spectral width of the peaks.
- The peaks are visible in both the real and imaginary components of the spectrum.
- The peaks often (but not always) cause abrupt jumps in the phase component of the spectrum.

As an example, the magnitude and phase spectra of an A-scan are shown in figure 1. The peaks are clearly discernible in the magnitude spectrum and the corresponding jumps in the (unwrapped) phase spectrum are also quite obvious.

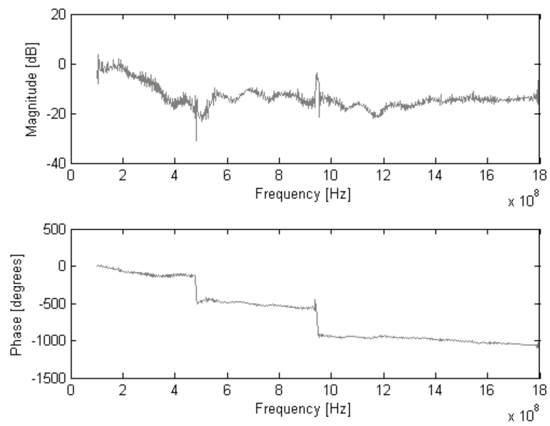

Figure 1: Frequency domain A-scan showing disturbance peaks

The semi-deterministic character of the spikes indicate that they are an artefact of the measurement system. Notch filters can be applied in the frequency domain to filter these peaks. This approach however unavoidably introduces its own artefacts to the spectrum.

An alternative approach is to employ a nonlinear filter to the observed spectra. The basic idea is to transform the data contained in a small window surrounding the spike in such a fashion to ensure that the resultant statistical distribution conforms to some desired distribution. More specifically, the ”desired” distribution doesn’t contain any outliers. This calls for a saturation-type nonlinearity to ensure that ”reasonably-valued” data is left unhindered, while data points with large values are gradually limited to some upper (or lower) value. One option for this nonlinearity is the hyperbolic tangent function illustrated in figure 2.

The hyperbolic tangent filter is applied as follows:

1. Extract a small portion of the spectrum surrounding a spike. The width of this window obviously has

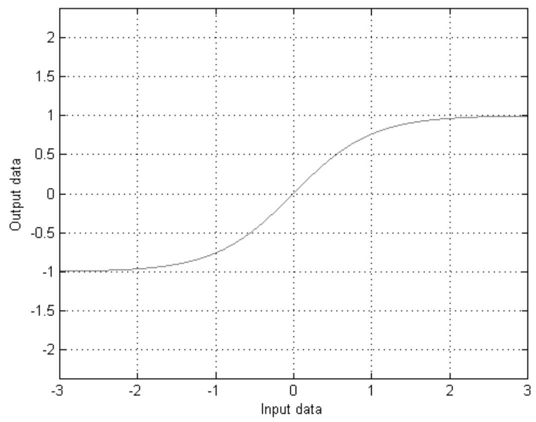

Figure 2: Hyperbolic tangent filter

an effect on the response of the filter. If it is too narrow, there won't be any outliers in the data in addition to the fact that the underlying distribution will be nonstationary due to the dominating effect of the spike. If, on the other hand, the window is too wide, then larger trends in the local spectrum will cause the data to also be nonstationary. In the end, the width of this window was chosen at 41 samples, since such a window only infrequently contained a nonstationary data sequence. (The stationarity of a data sequence can be tested by means of a run-test [15].)

2. Limit the amplitudes of the data in the window by means of the hyperbolic tangent function. This operation is performed separately on the real and imaginary components of the spectrum [6], since the data distributions of the latter spectra correspond more closely to the desired normal distribution than the magnitude and phase spectra.

In order to filter data at any general position and spread, it first has to be normalized to be commensurate with the input domain of the definition of the hyperbolic tangent function (which is approximately the unit domain). This normalization is done as follows:

$$y = \tanh\left(\frac{2x - 2\bar{x}}{\sigma}\right), \tag{1}$$

where x is the original data, $\bar{x}$ is the estimate of the position of the distribution (e.g. the median in the case of an outlier insensitive estimate) and σ is an estimate of the spread of the data (e.g inter-quartile range).

To ensure that the filtered data has a distribution similar to the original data x, the filter output has to be modified as follows:

$$y = \frac{\sigma}{2}\tanh\left(\frac{2x - 2\bar{x}}{\sigma}\right) + \bar{x}. \tag{2}$$

As an example, the hyperbolic tangent filter was applied to the A-scan of figure 1. The spikes are completely removed by the hyperbolic tangent filter with no perceptible effect on the rest of the spectrum, as can be seen in figure 3.

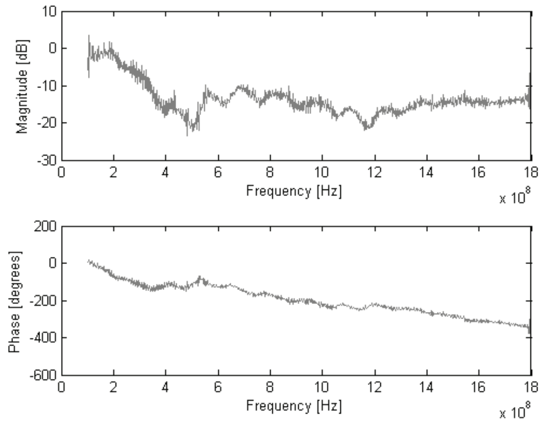

Figure 3: Frequency domain A-scan after spike removal

The hyperbolic tangent filter has two limitations. First, the data within the window should be stationary. Secondly, the data in the window should have an approximately normal distribution (at least it shouldn't be skewed). Both of these limitations can be addressed by a judicious choice of the width of the filter window.

2.4 Modelling background clutter by means of principal component analysis

It is well-known that image compression can be performed by principal component analysis [14]. In this technique the main features of an image can be reconstructed from its dominant eigenvalues (also known as the image's principal components).

The background clutter over a certain window of data in the down-track direction can therefore be modelled by means of the first few principal components of the B-scan contained in the window. These principal components are obtained by means of singular value decomposition (which is applicable to rectangular matrices). If the B-scan at position y_0 contained in a certain window of down-track data is represented by matrix $\mathbf{X}$, the singular value decomposition can be expressed as:

$$\mathbf{X} = \mathbf{USV}^{\mathrm{T}}, \tag{3}$$

where $\mathbf{U}$ is an orthogonal matrix containing the left singular vectors, $\mathbf{V}$ is an orthogonal matrix containing the right singular vectors and $\mathbf{S}$ is a diagonal matrix whose off-diagonal inscriptions are zero and whose main diagonal elements comprise the singular values of $\mathbf{X}$ arranged in descending order.

By attempting to reconstruct the original matrix $\mathbf{X}$ with only the largest few singular values and corresponding singular vectors, a matrix is obtained that contains the main characteristics of the original matrix, without some of the detail. The resultant reconstructed matrix is in other words a model of the background in the window of GPR data.

Clutter removal is then performed by subtracting the background model from the observed data. This is equivalent to a high-pass filter and typically results in the edges and other sharp features in an image being accentuated [14].

2.5 Background removal in the log-frequency domain

Unfortunately, the subtraction operator inherent in the above mentioned high-pass filter inevitably causes an enormous amplification in the phase spectrum of the data, as can be seen in the following short analysis.

Say that the original value of a frequency-domain A-scan at a specific frequency (ω) is given by: $z_{\text{original}}(j\omega) = a + jb$ and that the value of the background model at the same position and frequency is: $z_{\text{background}}(j\omega) = c + jd$. Then the value of the decluttered A-scan at the frequency of interest is given by:

$$\begin{aligned} z_{\text{decluttered}}(j\omega) &= (a-c) + j(b-d) \\ &= \sqrt{(a-c)^2 + (b-d)^2} \ldots \\ &\quad \ldots \exp\left[j\tan^{-1}\left(\frac{b-d}{a-c}\right)\right]. \quad (4) \end{aligned}$$

By focussing on the phase component of (4), it is clear that its phase will become very large if there is little difference between the original A-scan and the background model (which is most often the case).

As the above analysis shows, subtraction of *any background model* causes an unacceptable increase in the phase component. Clutter removal by means of subtraction of a background model is furthermore flawed since it inherently assumes an additive model for the process that gives rise to the observed data. In this approach (which is rather similar to a time-series model) the observed GPR signals ($S_o(j\omega)$) can be modelled as the sum of surface reflection effects ($S_s(j\omega)$), antenna effects ($S_a(j\omega)$), ground scattering ($S_g(j\omega)$) and possible target scattering ($S_t(j\omega)$) as follows [8]:

$$S_o(j\omega) = S_s(j\omega) + S_a(j\omega) + S_g(j\omega) + S_t(j\omega). \quad (5)$$

Improved clutter removal however requires an improved model for the process that gives rise to the observed data. A more accurate model makes use of the fact that the input signal generally is known. Under the (admittedly simplistic) assumption that the process of generating GPR reflections is linear and time-invariant, a measured GPR A-scan ($S_o(j\omega)$) can be modelled as the convolution of the input signal ($S_i(j\omega)$) with the various above mentioned processes [7]. In the frequency-domain convolution becomes multiplication, which means that the observed spectrum can be modelled as follows:

$$\begin{aligned} S_o(j\omega) &= S_i(j\omega) \times G_s(j\omega) \times G_a(j\omega) \times G_g(j\omega) \ldots \\ &\quad \ldots \times G_t(j\omega) \\ &= S_i(j\omega) \times G_b(j\omega) \times G_t(j\omega), \quad (6) \end{aligned}$$

where all of the effects that can be regarded as background clutter have been condensed into a background model $G_b(j\omega)$. (Note that in (6) the various processes are modelled as transfer functions ($G(j\omega)$), rather than mere observed signals ($S(j\omega)$) as in (5).)

A time-domain version of the target response can be obtained from (6) by means of a deconvolution filter [16], [17]. Equivalently, the spectrum of the target response can be obtained from (6) by division:

$$\begin{aligned} G_t(j\omega) &= \frac{S_o(j\omega)}{S_i(j\omega) \times G_b(j\omega)} \\ &= \frac{G_o(j\omega)}{G_b(j\omega)}, \quad (7) \end{aligned}$$

where $G_o(j\omega)$ represents the measured frequency response, which is essentially the measured spectrum in a stepped-frequency continuous-wave GPR system.

Practical implementation of (7) is best performed in the log-domain to avoid division by zero problems. By taking the natural logarithm on both sides of (7) and processing the magnitude and phase components of the spectra separately, the background model can be removed from the observed frequency response as follows:

$$\begin{aligned} \ln|G_t(j\omega)| + j\theta_t(\omega) &= \ln|G_o(j\omega)| - \ln|G_b(j\omega)| \ldots \\ &\quad \ldots + j\theta_r(\omega) - j\theta_b(\omega). \quad (8) \end{aligned}$$

The final frequency response of the target can then be obtained by taking the exponential function on both sides of (8) to obtain (7).

2.6 Frequency-to-time conversion

As we'll see in section 3, classification of GPR signals is performed in the time domain. The inverse fast Fourier transform (IFFT) can be used to perform the transformation from the frequency domain to the time domain.

Before the IFFT can be applied, the decluttered spectra have to be normalized. Normalization is typically employed in pattern recognition systems to ensure that the numerical ranges of the various features are commensurate [18]. The simplest form of normalization entails subtracting the mean value and dividing by the standard deviation, as in (9).

$$G_n(j\omega_k) = \frac{G_t(j\omega_k) - \overline{G_t(j\omega)}}{\sigma}, \quad (9)$$

where ω_k refers to a specific frequency index in the spectrum, $\overline{G_t(j\omega)}$ is the mean value of the entire A-scan and σ is the standard deviation of the same spectrum.

The most important aspect of (9) is that it removes the bias level from the spectrum. This is important, since the bias component in a spectrum causes an impulse (Dirac delta) to appear in the equivalent time-domain signal.

Finally, the IFFT can be applied. This is performed according to the procedure outlined in [16]. First, the spectrum is multiplied with a Kaiser window (with a coefficient of $1.32 \times \pi$). Next, the signal is zero-padded until the total length is 4096 samples. Finally, the signal is mirrored and the IFFT applied. As an example, the equivalent time-domain A-scan of the scan shown in figures 1, 3 and 7 is shown in figure 4. This A-scan clearly shows the response of a metal anti-tank (AT) mine.

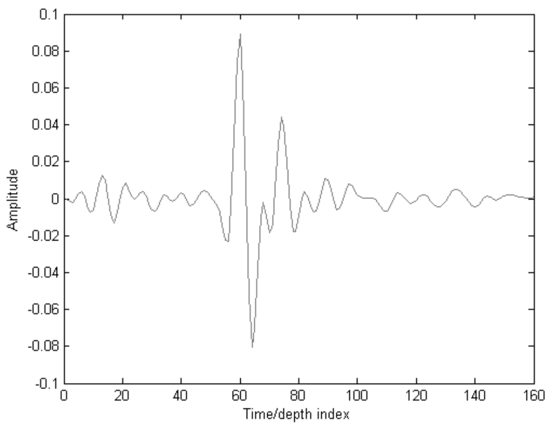

Figure 4: Time-domain A-scan after clutter removal

3. CLASSIFIER DESIGN

Feature extraction is arguably the most important step in any classification system. If the features can easily discriminate between the different classes (i.e. the classes form clearly separable clusters in the feature space), then classification is a mere formality. If however the features result in severely overlapping clusters, classification by even the most advanced classifiers is all but impossible.

Most of the GPR landmine classifiers in the literature are based on features that describe some or other discriminating property in the observed data. In such an approach the classifier isn't concerned with the fundamental processes that produce the observed signals, but only how to summarize the observed data in such a manner that the different classes can be detected.

Within the above mentioned feature extraction paradigm there are many examples of quite successful classification systems. Some researchers prefer an image-analysis approach in which features are extracted from horizontal B-scans that mimic those features a human operator would focus on in order to detect mines [4]. These features try to capture the essence of the concentric circles and hyperbolas in B-scans that betray the presence of buried objects. Others favour using features distilled from the energy density spectra of A-scans [5], [8].

There are however researchers that attempt to model the responses of buried objects to GPR signals in a much more principled manner. As an example Roth *et al.* derived expressions for the impulse response of a low metal content mine under idealized soil conditions [19]. A model-based approach is also quite useful to obtain a background model for the observed B-scans with the aim of clutter removal [7]. Inevitably these authors have to make numerous simplifying assumptions in order to obtain tractable models.

As previously mentioned, in practice a GPR landmine detector is faced with a staggering array of unknowns including: the exact location of different soil layers; the frequency of occurence and size of rocks and pebbles; the water content of the soil (which may change as a function of location and weather); the position, depth and orientation of mines; and lastly, the dielectric properties of sediment and rocks (which is influenced by a huge number of factors that are difficult or impossible to measure beforehand [17]).

In the light of the above mentioned uncertainties it should be clear that first-principles modelling of all possible landmine responses is unfeasible. The only practical alternative is so-called black-box system identification. One such approach is to model the observed GPR data by means of parameterized LTI models. Accurate models for the various classes of buried objects imply that the parameter vectors of the models are also unique and therefore suitable candidates for classification features. Unfortunately from extensive simulation experiments, classical parameterized models (ranging from ARX models to Wiener-Hammerstein models) exhibit both poor generalization and discrimination ability on GPR data as compared to artificial neural networks. (The *generalization ability* of a model refers to its ability to model other instances of the same system that weren't included in the data set upon which the model's parameter were estimated. On the other hand, a classifier's *discrimination ability* is its ability to distinguish between different classes.)

One of the few remaining options to obtain a black-box model for the observed time-domain A-scans is by means of artificial neural networks.

3.1 Detail design

A full exposition of the vast field of artificial neural networks is beyond the scope of this paper (consult e.g. [18]). Certain classes of artificial neural networks (especially multilayer perceptrons and radial basis function networks) can be viewed as essentially parameterized nonlinear function approximators. Consequently artificial neural networks (henceforth referred to as neural networks) shouldn't be regarded as fundamentally different to any other parameterized nonlinear model structure.

Radial basis function networks (RBFNs) possess a number of advantages over multilayer perceptrons (MLPs), namely:

- Training of RBFNs typically entails clustering (for the hidden layer neuron centre-points) and least-squares approximation (for the output nodes). This process is much faster than the backpropagation and nonlinear minimization required for MLP training.
- Due to the limited extent of each basisfunction in the hidden layer, RBFNs don't commit huge extrapolation errors as is the case with MLPs. (MLPs approximate a function by means of hyperplanes, while RBFNs do the same job with Gaussian kernels.)
- The mapping performed by a RBFN can be interpreted in terms of a basisfunction expansion, which is slightly more insightful than the essentially unknown mapping done by a MLP.

In this paper, RBFNs are used as function approximators. More specifically, a model is estimated for each class of landmine (e.g. metal mines, low metal AT and anti-personnel (AP)). The neural network is trained to form as a nonlinear autoregressive (AR) time-series model for the time-domain A-scan of the particular mine in question. In an AR-model the current sample value of the signal is modelled by a linear combination of previous sample values of the signal. Mathematically, an AR-model can be expressed as follows:

$$y(t) = \sum_{i=1}^{N} a_i y(t-i). \tag{10}$$

An AR-model can be duplicated by a neural network by arranging the inputs of the network in a so-called tapped-delay line. This entails that the previous sample of the signal is used as the first input to the network, the sample prior to the previous sample is used as the second network input and so on. The output of this network is then taken as the current sample value of the signal.

The neural network architecture described above functions as a single step-ahead predictor. In other words, the current and previous sample values of the signal are used to predict the value of the forthcoming sample. This can obviously be extended to a k-step ahead predictor in which the network predicts the value of the signal k-sample values into the future on the basis of the current and previous samples. If a separate neural network is trained to perform as a k-step ahead predictor for each class of landmine, then these models can be used to classify a new (unknown) A-scan. This is done by simply allocating the A-scan to the class whose neural network has the best prediction performance. In this manner both feature extraction (represented by the Gaussian basisfunctions of a RBFN) and classification can be performed by the same neural network.

The training set of each class model comprised a 3×3 neighbourhood of A-scans in the vicinity of a known mine position. Training was performed on these nine A-scans until either a maximum number of neurons were reached or a minimum error was reached. (Training of the standard Matlab® RBFN entails automatic positioning of the basisfunctions. These basisfunctions are added incrementally at each iteration. This is followed by a least squares solution for the linear mapping between the current basisfunctions and the future sample presented at the output of the network.)

4. RESULTS: CLUTTER REMOVAL

4.1 Clutter removal performance measures

Successful clutter removal algorithms are characterised by the following attributes:

- Low computational cost. This is important with real-time implementation in mind.
- Quality. High fidelity reproduction of target signals with only a minimal presence of undesired clutter. Quite often in the literature the quality of a clutter removal technique is "measured" in terms of a subjective assessment of the resulting A- and B-scans. An objective comparison of the quality various clutter removal algorithms however requires a quantitative measure of the performance of the respective algorithms.

One quantitative measure of the quality of the clutter removal process is the signal-to-clutter ratio (SCR) defined in [20]. The SCR is a measure of the relative spread of GPR data at a specific depth in the cross-track direction and assumes that all antennas in the array are equal, that the soil properties in the cross-track direction is constant and that landmines only occur infrequently. Although originally defined for cross-track B-scans, it can also be applied to down-track B-scans.

The SCR is quite useful to determine how much clutter has been removed by a particular algorithm. Unfortunately this measure can only be used to study the effect of a clutter removal algorithm one A-scan at a time. It would be much more convenient if a measure could be defined that allows a clutter removal algorithm to be evaluated on the basis of its performance on an entire C-scan. One such measure is based on correlation.

The presence of ground-bounce and gradually changing soil characteristics results in a high degree of cross-correlation between adjacent A-scans in all directions. Background removal would therefore ideally cause a large decrease in the general degree of cross-correlation between neighbouring A-scans, with the only exceptions occuring at the locations of buried objects (e.g. landmines). (The latter conclusion is under the assumption that the occurence of landmines can be viewed as statistical outliers as discussed in the appendix to this paper.) The quality of a clutter removal algorithm can therefore be assessed by calculating the average degree of peak cross-correlation between each A-scan and its eight immediate neighbouring A-scans.

Figure 5 shows the average degree of correlation of neighbouring A-scans after application of a time-domain principal component based clutter removal algorithm (which merely entails application of the algorithm described in section 2.4 on time-domain GPR data). This figure should be interpreted as a top-down view of an area of ground which has been scanned with a ground-penetrating radar. Areas showing high degrees of correlation may be indicative of interesting subsurface objects or soil layers. Superimposed on this figure are the positions of known mines (indicated with black squares). Clearly, there is room for improvement since large areas of ground are shown to have highly correlated A-scans.

Clutter removal performed via the procedure described in section 2 has a marked improvement on the average degree of correlation in the data. As figure 6 shows, the positions of metal mines are now clearly visible (a significant improvement on figure 5).

Individual A-scans are also much more informative after the log-frequency domain clutter removal process. This can be clearly seen by comparing the A-scan obtained after spike filtering (figure 3) with the A-scan at the same position after background removal (figure 7). The existence of potentially informative subbands can be clearly seen in both the magnitude and phase spectra.

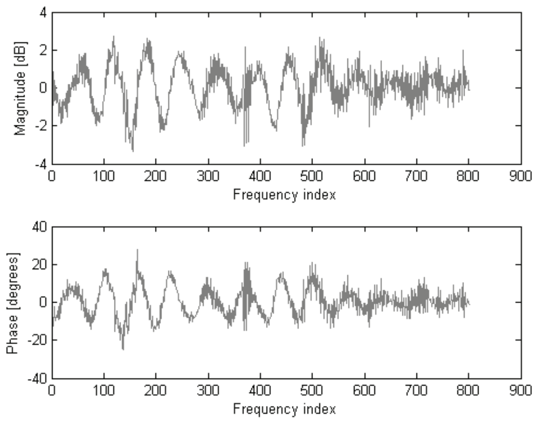

Figure 7: Frequency domain A-scan after clutter removal

4.2 *Descriptive power of the decluttered data*

One of the vexing problems of data preprocessing is that the process occurs rather blindly. Educated guesses are made and algorithms applied in the hope that useful information can be distilled from the raw data. Only after the final classification stage has been completed, can it be seen if the various preceding stages were successful. It would therefore be of great advantage if a quantitative measure could be developed with which the descriptive and discriminative power of a dataset can be assessed prior to feature extraction.

The traditional approach to solve the above mentioned problem is to cluster the data and measure the ratio of the inter-cluster distance to the cluster size [18]. In this manner a measure can be found of the class separability in the data. Such a clustering-based approach is however limited to data consisting of a limited number of features and is definitely not suited to raw data.

A viable option is to implement a primitive correlation-detector (also known as a matched filter) [21]. Here the objective is to determine whether there are other A-scans in a C-scan that are similar to a given A-scan. The difference between this correlation-detector and the correlation-based measure of the previous section is that the entire C-scan is analysed for potential similarities with a given A-scan and not only the A-scan's immediate neighbourhood. In this manner the data can be analyzed at different stages of the system to determine the separability of the classes.

As an example, the correlation-detector was applied to the decluttered time-domain data to determine whether there are any other A-scans similar to an A-scan at the position of a high metal AT mine. The results are shown in figure 8. Clearly, the A-scans of the two examples of high metal AT mines are quite similar to each other. Another encouraging result from figure 8 is that the high metal AT A-scans are quite different from the other A-scans in the data. Such a result predicts that good classification results can be obtained on the data.

Figure 9 however paints a different picture altogether. This figure shows the peak cross-correlation with an A-scan at the position of an AP mine. If the scale of the colour bar is taken into consideration, figure 9 tells us that the A-scans of AP mines aren't correlated with any other A-scans in the C-scan (despite the fact that there are two known examples of M14 mines in the dataset). Similar analyses on the other low metal mines indicate that plastic mines are all but invisible in the current decluttered time-domain data. Similar problems have been reported elsewhere in the GPR literature [22].

5. RESULTS: MINE DETECTION

All of the results in this paper were obtained on a small dataset consisting of seven landmines buried in an artificially constructed sandpit. This set of mines consists of the following specific types: metal AT (2 examples), M14 (2 examples), TMA-3 (1 example), VSHCT (1 example), No. 8 (1 example). Over time a meerkat family also made the sandpit their home, with their burrows quite conspicuous in the GPR reflections. A GPR antenna array

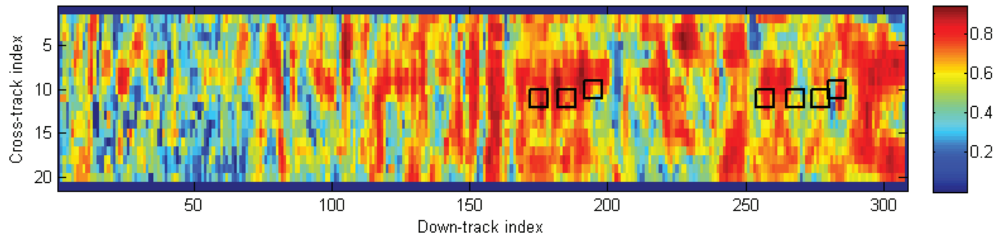

Figure 5: Degree of average correlation after time-domain principal component based filtering

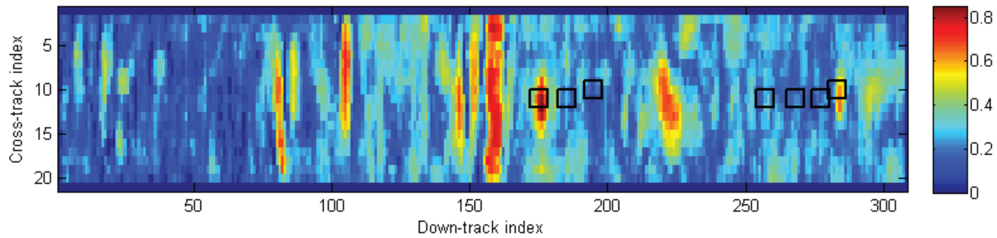

Figure 6: Degree of average correlation of magnitude spectra after log-frequency domain processing

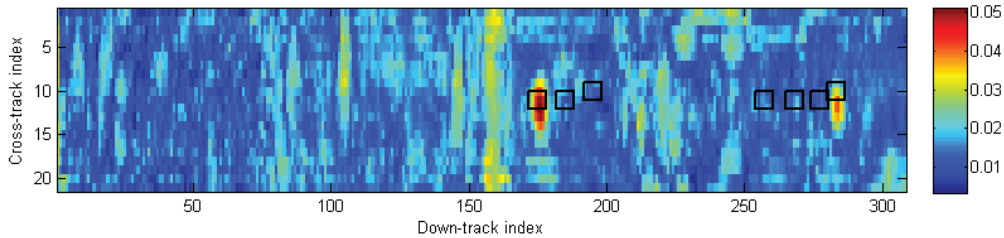

Figure 8: Peak cross-correlation with an A-scan at the position of a high metal AT mine

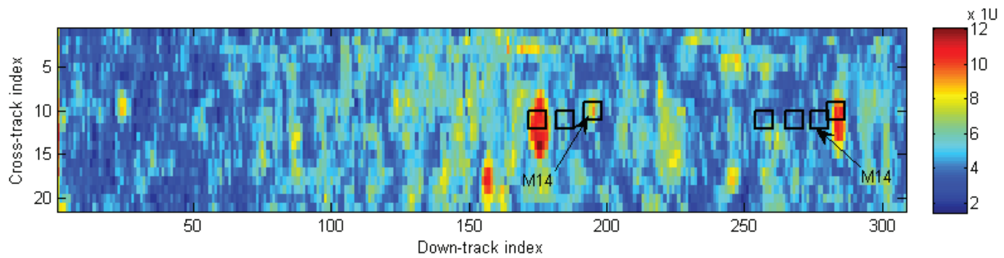

Figure 9: Peak cross-correlation with an A-scan at the position of an AP mine

(model V1821) from 3D-radar was used to obtain the GPR data. This GPR antenna array is a general-purpose instrument designed for applications ranging from airport and road inspection to archeology and military purposes. Most of the GPR studies in the literature are obtained with impulse type GPR, whereas the instrument used in this research is a SFCW radar.

5.1 RBFN parameter optimization

Optimizing a RBFN to serve as an accurate time-series model involves that optimal values for the following parameters have to be found:

- Prediction horizon. This refers to the value of k in a k-step ahead predictor.
- Length of the tapped delay line. The length of the tapped delay line determines how many previous samples are taken into consideration to form a prediction of a future sample. (This parameter is the same as the order of a classical AR-model.)
- Width of the basis functions.
- Maximum network size. This indicates the maximum number of basis functions that can be placed during training.

These parameters were optimized with the eventual objective of the neural networks models in mind, namely to serve as classifiers of GPR A-scans. Consequently, the RBFN parameters were chosen in such a fashion that the resultant network exhibits a good discrimination ability coupled with a good generalization ability. Both of these properties can be inferred from the peak prediction performance of a network on a 3×3 neighbourhood of A-scans. These A-scans were chosen from different locations in the available data representing known classes of data.

The discrimination ability of a network is revealed by the difference in the prediction performance between data from the training class and data from other classes. The generalization ability, on the other hand, corresponds to the degree of similarity in the prediction performance of the network on data from similar classes (e.g. two different instances of metal AT mines).

Optimizing the prediction horizon: The peak prediction performance of the RBFN predictor is set out as a function of the prediction horizon in table 1. This table was obtained by fitting a nonlinear mapping between 15 previous samples of an A-scan and a sample k steps into the future. The networks for the different prediction horizons were all estimated on a 3×3 neighbourhood of A-scans belonging to one example of a metal AT mine. Furthermore, their basis functions have a spread of 0.5 times the amplitude range of the entire A-scan and a maximum network size limited to 200 hidden neurons. The prediction prediction performance of the RBFN predictor on its estimation dataset is given in the first column of table 1. The other columns contain the peak prediction performance of the same predictor applied to data of other classes.

Table 1: Prediction performance [%] of a RBFN as a function of the prediction horizon

k	Metal AT	Low metal AT	Ground
1	95.49	80.99	82.70
2	96.30	57.80	63.84
3	93.94	31.83	47.43
4	90.76	-32.48	-0.29
5	87.95	-22.33	-11.26
6	86.15	-37.24	-37.46

From table 1 it is clear that the best combination of discrimination and generalization is attained by a RBFN with a prediction horizon of four.

Optimizing the rest of the RBFN parameters: The remaining three parameters of the RBFN (the length of the tapped delay line, the spread of the basisfunctions and the maximum network size) were optimized in a similar fashion as the prediction horizon. Consequently, we'll focus on the most important observations gained during the optimization process without delving into the detail results.

With respect to the length of the delay line, it was found that the best combination between discrimination and generalization is obtained by using the previous 13 samples to form a prediction. Short delay lines generally result in poor discrimination ability, while the generalization ability reduces if the delay-line which is too long (i.e. the network is forced to perform an overfitting).

Concerning the spread of the basis functions, it was found that a scaling factor of 0.3 times the range of the A-scan amplitude values results in the best discrimination and generalization ability. If the spread is too small, then the network can't generalize at all. If, on the other hand, the spread is too large, the network loses the ability to discriminate between classes. Lastly, it was found that the best marriage of discrimination ability and generalization ability is obtained by a RBFN with a maximum of 200 basis functions.

In summary, the optimized parameter values for a RBFN model for GPR A-scans are given in table 2.

Table 2: Optimized parameter values for a RBFN

Parameter description	Value
Prediction horizon	4
Length of the tapped delay line	13
Spread of the basisfunctions	0.3
Maximum number of basisfunctions	200

5.2 *Prediction performance of the RBFN on various classes of mines*

The above mentioned RBFN was used to form time-series models for the following classes of GPR targets: metal AT mines, low metal AT mines (namely: VSHCT, TMA-3 and No. 8), AP mines (M14), meerkat tunnels and clean soil. Within the general class of clean soil two subclasses were modelled: one inside the sandpit in which the landmines were buried and the other (red ground) outside the sandpit.

The peak prediction performance of the various RBFN class-models are given in table 3. The prediction performances in this table were measured in 3×3 neighbourhoods at the locations indicated in the columns of the table. Each row represents a different model (estimated on a particular class). Ideally, this table should resemble a diagonal matrix, with a few exceptions. One would for example expect that a model estimated on one instance of a metal AT mine should perform well on both examples of metal AT mines in the dataset.

The results in table 3 generally indicate that the individual models can discriminate well between their A-scans of their class and other classes. Unfortunately only the models fitted to the high metal AT mines exhibited useful generalization ability. The lack of generalization ability of

Table 3: Peak prediction performance [%] of RBFN models trained on different classes

	Metal AT(1)	Metal AT(2)	VSHCT	No.8	TMA-3	M14(2)	M14(1)	Tunnels	Sandpit
Metal AT(1)	89.77	44.48	13.61	12.31	17.73	12.37	24.22	5.72	-4.44
Metal AT(2)	38.68	81.69	-31.19	-19.15	-24.13	10.95	15.55	-47.54	-10.67
VSCHT	12.77	11.65	74.13	16.13	20.88	22.36	35.03	15.72	37.72
No. 8	8.51	14.97	31.61	81.94	30.00	32.46	27.67	26.76	21.24
TMA-3	10.87	12.69	26.64	33.82	80.07	27.42	33.54	31.16	26.95
M14(2)	10.96	9.26	32.19	17.44	35.55	72.54	37.31	25.83	25.49
M14(1)	12.18	11.91	28.71	7.96	33.70	24.61	74.83	23.72	27.25
Tunnels	18.01	17.88	27.39	21.57	30.42	26.11	44.28	79.32	30.54
Sandpit	14.64	8.58	34.05	25.74	-2.35	28.01	37.06	2.32	81.39

the models estimated on the AP mines and examples of clear ground will be investigated in greater depth shortly.

By applying any of the models mentioned in table 3 to an entire C-scan, it is possible to form a plot of the prediction performance of the particular model as a function of the cross-track and down-track indices. Figure 10 is an example of the prediction performance of a RBFN model estimated on one example of metal AT mines. From this figure it is clear that the model only performs well on two areas. The latter areas coincide with the positions of metal AT mines. Figure 10 therefore shows that the RBFN model estimated on metal AT mines can both discriminate and generalize well.

The generalization ability of the RBFN models estimated for high metal AT mines, however, isn't shared by the models for the various low metal mines. As an example, figure 11 shows the prediction performance of the model for TMA-3 mines. Clearly the model can be used to recognise A-scans from its estimation dataset. The lack of duplicate versions of this specific type of mine in the dataset makes it difficult to come to general conclusions, but it doesn't seem as if the model has the capability to generalize well to other types of low metal ATs.

Some of the other low metal AT and AP mines (e.g. VSHCT and M14) aren't clearly distinguishable from the soil in which they are buried. This can be clearly seen from the prediction performances of the respective models trained on these classes of mines. As an example, figure 12 shows the prediction performance of the RBFN model estimated on one of the M14 mines. The first observation that can be made from this figure is that the model can't generalize well to include the other example of the M14 mine. More interesting however is the fact that a relatively large number of A-scans in the immediate vicinity of the training set are also modelled quite well by the RBFN. This suggests that the "M14-model" is actually to a large extent a model for the specific soil in which the mine was buried.

5.3 *Classification performance of the RBFN-based classifier*

The complete procedure for the entire classification system is as follows:

1. Train the individual RBFN models on examples of their respective classes. This training phase occurs offline before the system has to sweep an area for mines. Consequently, the time constraints for training are less stringent than during the final classification stage.
2. When an area has to be sweeped, the following steps have to be followed:
   (a) Perform clutter removal. At present, this algorithm is performed one down-track B-scan at a time.
   (b) Determine the prediction performance of each RBFN model in the classifier's database. This step is performed for each A-scan in the cross-track B-scan.
   (c) Determine the model with the best prediction performance. This will form the system's hypothesis of the identity of the A-scan.
   (d) Only allocate the identity of the best model to a particular A-scan if the prediction performance of the model is better than a minimum threshold (e.g. 50 %). In this manner extrapolation errors can be avoided and control exercised over the false alarm rate of the system.
   (e) Repeat the entire process for the next cross-track B-scan.

The final classifier only consisted of five models. One trained on one example of high metal AT mines. The low metal AT class was represented by a model estimated on a No. 8 mine, while only one of the two M14 mines was used to obtain a model for the AP mine class. Lastly, meerkat tunnels and clear ground were also represented by one model each.

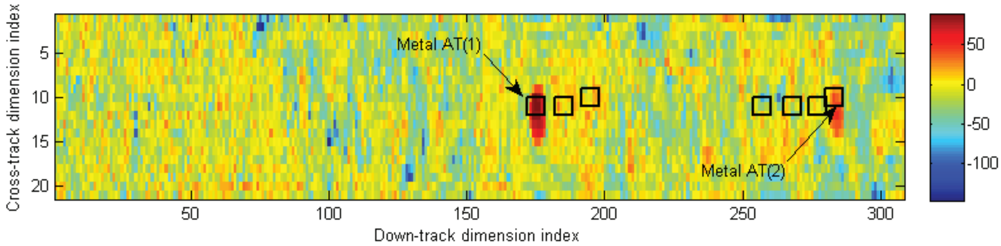

Figure 10: Prediction performance [%] of the RBFN trained on one example of metal AT mines

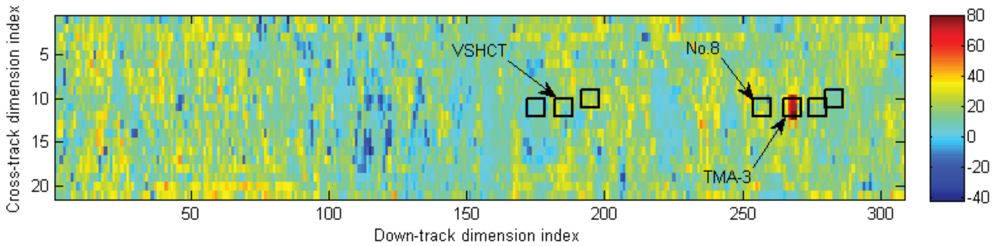

Figure 11: Prediction performance [%] of the RBFN trained on the TMA-3 low metal AT mine

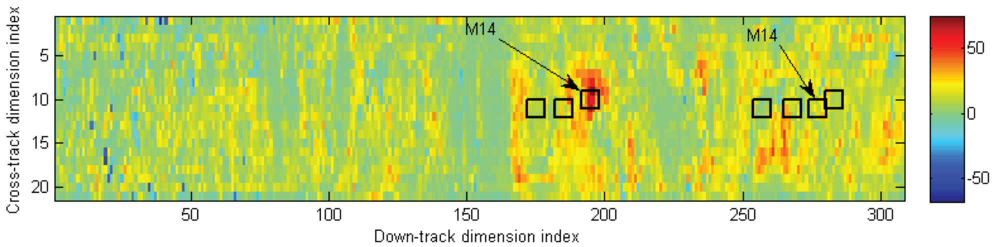

Figure 12: Prediction performance [%] of the RBFN trained on one of the M14 mines

The classification performance of the above mentioned system is reported in figure 13. This figure shows a horizontal slice of the time-domain GPR data at a ”depth” of 6.1 nanoseconds. Classification was performed with a minimum threshold of 50 % and the classification decisions are superimposed on the time-domain GPR image. The symbols used in figure 13 convey the following meaning:

- Red dots indicate high metal AT mines (as identified by the system).
- Blue stars indicate low metal AT mines (as identified by the system).
- Green stars represent AP mines (as identified by the system).
- Magenta dots represent meerkat tunnels (as identified by the system).
- Yellow squares indicate the true position of high metal AT mines.
- Cyan squares show the true location of low metal AT mines.
- Black squares represent the true position of AP mines.

The results in figure 13 show that the system can recognize both instances of high metal AT mines, but has problems generalizing between the various low metal mines in addition to a high false alarm rate for low metal AT mines. A positive feature of the system is that it can correctly distinguish between high metal mines and meerkat tunnels, which is important since meerkat tunnels often resemble high metal AT mines quite closely.

The small size of the dataset and its relative uniqueness (due to the use of a commercial SFCW GPR antenna array) coupled with the realism of the test setup makes comparison with the existing literature quite difficult. It does however seem as if the detection system presented in this paper is quite capable of accurate detection of high metal AT mines, but does struggle with low metal mines.

6. CONCLUSION

In this paper a new clutter removal algorithm for GPR data is presented as well as a novel application of neural network-based time-series models for landmine detection. The system does seem promising for the detection of metal

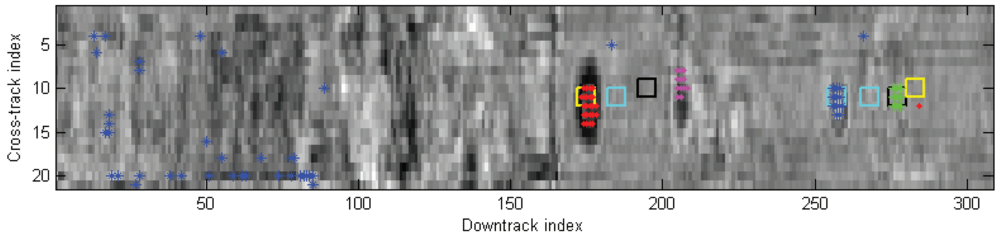

Figure 13: Classification results for a decision threshold of 50 %

AT mines. The algorithm's inability to reliably detect low metal mines is a source of concern, since this defect does limit its practical applicability for humanitarian demining operations. Similar problems have however been recently reported elsewhere in the GPR literature [22].

The large variation in the dielectric properties of both the landmines (ranging from AT to AP mines) as well as the soil in which they are buried, leads us to the conclusion that it seems unlikely that one single sensor and classifier will be able to detect all mines in every practical situation. Combining (or fusing) the recommendations of different classifiers operating on the outputs of different sensors (e.g. GPR and electromagnetic induction sensors) has recently been reported on in the literature [23] and [22]. The results obtained by this approach seem promising enough to warrant additional work, but significant improvements will still have to be made to solve the mine detection problem in practice. The next paper describes alternative approaches to both clutter removal and classification.

7. APPENDIX: MINES REPRESENT OUTLIERS IN GPR DATA

The clutter removal task can be rephrased as removing everything from the data that isn't due to the presence of a landmine. The choice of clutter removal techniques is heavily influenced by the frequency of occurrence of a landmine in the field. With the exception of the border between North and South Korea, the "Cordon Sanitaire" in Zimbabwe represents one of the minefields in the world with the highest density of landmines [24]. This 25 meter wide corridor has an average density of 5,500 mines per kilometer, which can for the sake of argument be approximated as a density of 0.22 mines per square meter (under the assumption of an uniform distribution of landmines). A model V1821 GPR antenna array from 3D Radar has an effective scan width of 1.575 meter [25], which would result in the occurrence of one landmine every 2.886 meter of down-track movement. This area corresponds to 606 A-scans (if down-track samples are separated by 10 cm). Consequently, it is safe to assume that landmine signatures represent outliers (at most one in 606 A-scans) in field-measured data.

REFERENCES

[1] *Landmine monitor 2011*. Canada: Mines Action Canada, October 2011.

[2] K. Schreiner, "Landmine detection research pushes forward, despite challenges." *IEEE Intelligent Systems*, vol. 17, pp. 4–7, March/April 2002.

[3] L. Robledo, M. Carrasco, and D. Mery, "A survey of land mine detection technology," *International Journal of Remote Sensing*, vol. 30, no. 9, pp. 2399–2410, May 2009.

[4] P. Gader, W.-H. Lee, and J. N. Wilson, "Detecting landmines with ground-penetrating radar using feature-based rules, order statistics, and adaptive whitening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2522 – 2534, 2004. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2004.837333

[5] J. N. Wilson, P. Gader, W.-H. Lee, H. Frigui, and K. Ho, "A large-scale systematic evaluation of algorithms using ground-penetrating radar for landmine detection and discrimination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2560 – 2572, August 2007. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2007.900993

[6] S. Tjora and E. Eide, "Evaluation of methods for ground bounce removal in GPR utility mapping," in *Tenth International Conference on Ground Penetrating Radar*, 2004.

[7] O. Lopera, E. C. Slob, N. Milisavljevic, and S. Lambot, "Filtering soil surface and antenna effects from GPR data to enhance landmine detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 707 – 717, March 2007. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2006.888136

[8] K. Ho, L. Carin, P. D. Gader, and J. N. Wilson, "An investigation of using the spectral characteristics from ground penetrating radar for landmine/clutter discrimination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 4, pp. 1177 – 1191, 2008. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2008.915747

[9] A. Van der Merwe and I. Gupta, "A novel signal processing technique for clutter reduction in GPR measurements of small, shallow land mines," *IEEE transactions on geoscience and remote sensing*, vol. 38, no. 6, pp. 2627–2637, November 2000.

[10] K. Ho and P. Gader, "Alinear prediction land mine detection algorithm for hand held ground penetrating radar," *IEEE transactions on geoscience and remote sensing*, vol. 40, no. 6, pp. 1374–1384, June 2002.

[11] V. Kovalenko, A. G. Yarovoy, and L. P. Ligthart, "A novel clutter suppression algorithm for landmine detection with GPR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3740 – 3750, 2007. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2007.903694

[12] D. Daniels, *Surface-penetrating radar*. London: The Institution of Electrical Engineers, 1996.

[13] P.-A. Sandness, *3DR, 3d-Radar data format: version 1.3*, 3d-radar, December 2009.

[14] R. Gonzalez and R. Woods, *Digital image processing*. Reading, Massachusetts: Addison-Wesley, 1992.

[15] J. Bendat and A. Piersol, *Random data: analysis and measurement procedures*. New York: Wiley-Interscience, 1971.

[16] B. Karlsen, J. Larsen, K. Jakobsen, H. Sø rensen, and S. Abrahamson, "Antenna characteristics and air-ground interface deembedding methods for stepped-frequency ground penetrating radar measurements," in *In Proc. of SPIE, AeroSense 2000*, vol. 4038, April 2000.

[17] J. Lester and L. Bernold, "Innovative process to characterize buried utilities using ground penetrating radar," *Automation in Construction*, vol. 16, p. 546555, 2007.

[18] C. Bishop, *Neural networks for pattern recognition*. USA: Oxford university press, 1996.

[19] F. Roth, P. Van Genderen, and M. Verhaegen, "Convolutional models for buried target characterization with ground penetrating radar," *IEEE transactions on antennas and propagation*, vol. 53, no. 11, pp. 3799–3810, November 2005.

[20] X. Feng, M. Sato, Y. Zhang, C. Liu, F. Shi, and Y. Zhao, "CMP antenna array GPR and signal-to-clutter ratio improvement," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 1, pp. 23 – 27, January 2009. [Online]. Available: http://dx.doi.org/10.1109/LGRS.2008.2006634

[21] E. Ifeachor and B. Jervis, *Digital signal processing: a practical approach*. Harlow, England: Addison-Wesley, 1993.

[22] A. C. B. Abdallah, H. Frigui, and P. Gader, "Adaptive local fusion with fuzzy integrals," *IEEE Transactions on Fuzzy Systems*, vol. 20, no. 5, pp. 849 – 864, 2012. [Online]. Available: http://dx.doi.org/10.1109/TFUZZ.2012.2187062

[23] H. Frigui, L. Zhang, P. Gader, J. N. Wilson, K. Ho, and A. Mendez-Vazquez, "An evaluation of several fusion algorithms for anti-tank landmine detection and discrimination," *Information Fusion*, vol. 13, no. 2, pp. 161 – 174, 2012, bayesian fusion;Borda Count;Context dependent;Decision template;Dempster-shafer;Fuzzy integral;Land mine detection;. [Online]. Available: http://dx.doi.org/10.1016/j.inffus.2009.10.001

[24] *Landmine monitor report 2009: toward a mine-free world. Special ten-year review of the mine ban treaty*. Canada: Mines Action Canada, October 2009. [Online]. Available: www.lm.icbl.org/lm/2009

[25] "V-series multi-channel antenna arrays for GPR: Mkt-ds-v-series-040810v1," 2010.

THE INFLUENCE OF MAJOR EXTERNAL AND INTERNAL EVENTS ON THE CULTURE OF AN ENGINEERING ORGANISATION

W. Theron[*], L. Pretorius[**] and K.-Y. Chan[***]

[*] *Master's student, Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa E-mail: willie@wazat.co.za*
[**] *Professor, Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa E-mail: leon.pretorius@up.ac.za*
[***] *Senior Lecturer Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa E-mail: alice.chan@up.ac.za*

Abstract: A Case study company that was set up as a project where the technical focus, activities and behaviour set the initial culture is considered in this research. Over a period of 11 years the Case study engineering organisation was exposed to many influences in the electrical utility industry that now give lead to questions such as: How did events influence the engineering culture and how did the culture change over time? Engineering organisations are subjected to external and internal events which are not always within their control. These include technological changes, economical changes or new competition, change in ownership, business focus or technical leadership. The ability to absorb such events is not only a function of the organisation's technology infrastructure, availability of funding or skills, but also of the organisational culture prevailing at the time. The objective of the research is to determine how eight events impacted the culture of an engineering organisation over a period of six years. The results show that the culture is indeed influenced by events, with an indication that the different work areas within the organisation experienced the cultural changes differently. The employees that worked for the organisation six years or longer also experienced the changes differently from those that were only employed for the last five years of the organisation's life. These results may assist the understanding of the impact that events may have on an organisation and allow early risk mitigation to counter undesirable culture forming.

Keywords: organisational culture, change management, event impact, engineering organisation.

1. INTRODUCTION

All engineering organisations are subjected to external and internal events, which are not always within their control. External events can include technological changes, global or local financial impacts, market shifts or new competition entering the market. Some of these events may be initiated by a black swan event as defined by Taleb [1]. On the other hand internal events can include critical skills shortages, change in ownership, and change in business focus or change in technical leadership. These events will all test the resilience of the organisation. The ability to absorb such events is not only a function of the organisation's technology infrastructure, knowledge management or availability of funding or skills, but is also driven by the culture prevailing in the organisation at the time.

Understanding how a major event may impact the culture of an engineering organisation, may make it possible to prepare the organisation in advance of the event with the intent to stay on course or to adjust course to be able to absorb the event.

1.1 Theory and research review

The research topic implies that organisational culture is investigated as part of the research. Schein [2] offers the following definition for organisational culture: "A pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems".

Brown as quoted by Manetje et al [3] defines organisational culture as: "the pattern of beliefs, values, and learned ways of coping with experience that have developed during the course of an organisation's history, and which tend to be manifested in its managerial arrangements and in the behaviours of its members".

Both these definitions suggest that culture is developed over time, but do not imply how an organisation may behave should an event be imposed on the organisation.
A large number of change management models have been developed, but when approaching change management

within an organisation it cannot be based on a model produced "from elsewhere" as shown by Oxtoby [4]. Therefore the question can be asked how effective changes or events can be managed within a new organisation when no clear culture has been established on how events can be absorbed and managed.

According to O'Reilly [5] organisational commitment from a motivational perspective is the "individual's psychological bond to the organisation, including the sense of job involvement, loyalty and belief in the values of the organisation". These levels of individual motivation were explored as part of the research to determine the impact the major events had on the organisation's commitment level.

The attitudinal definition of organisational commitment according to the attitude-behaviour model of Eagerly and Chaiken as quoted by Solinger et al [6] indicates that organisational commitment is a combination of an employee's attitude and response toward the work experiences and perceptions regarding the organisation, and the employee's attitudes and personal traits that lead to a positive or negative emotion and behaviour (e.g. leaving or staying). On the other hand the definition of organisational commitment by Storm et al [7] indicates that the individuals, through their own actions, develop a commitment by involving themselves in commitment behaviour.

1.2 Problem statement

The definition of organisational culture implies that a pattern of going about doing business is established over time and will continue to be used as long as it is regarded as being valid. However, it is not clear how an event will influence an organisation's culture and indirectly influence the operational outcome.

When reading and talking about engineering achievements, practise or the role of engineering in society, words such as innovation, team work, product delivery, product improvement and related terminologies are used. It is widely accepted that the achievements of an engineering organisation depends on all of these and many more. However, engineering organisations prefer predictability in outcome (minimum risk) and therefore huge effort is spent on ensuring processes are in place, rigorous project plans are established, teams are "aligned" and good governance is used to ensure progress continues as planned. The role that engineering culture may play in the predictability of outcome is not clear. Furthermore, can the culture be managed to improve the outcome? Is the culture merely a reflection of the-way-we-did-the-project-to-date status?

The definition of organisational culture by Brown as quoted in [3] implies that a pattern of going about doing business is established over time and will continue to be used as long as it is regarded as being valid.

The preliminary investigation suggests that it is not clear if it can be proactively determined that an event will influence an organisation's culture and how it will influence the operational outcome.

The following questions support the problem statement:
- Which organisational culture characteristics were present in the case study project?
- How did the external factors and events influence the engineering culture?
- How did the engineering culture change over time?
- What was the impact of the major events on the cultural characteristics of the Case study company?
- To what extent did teams or individuals disassociate or associate themselves from the "new" organisation culture that was established after each event?

1.3 Research objective

The research objective in this paper is to determine what influence an event may have on an organisation's culture. The study uses the history of the Case study engineering organisation to explore the cultures that existed over an 11 year period and how these were shaped by external and internal events. The intent of the research is not to determine whether the culture could have been manipulated to reach a different outcome, neither to develop a new model, but rather to measure the outcome of the events in terms of the impact if had on the engineering organisation's culture.

1.4 Importance of the research problem

According to Ries [8], start-up endeavours rely on valid learning experiences, experimentation with possibilities, short iterative product releases, measuring progress, and obtaining customer feedback as soon as possible. Schein [9] reports that the founder in a start-up endeavour embeds the culture not necessarily by explicit actions and that the embedment process to establish a culture is mainly via a "teaching" process. This "teaching/learning" process of culture formation takes place when, in its simplest form, someone must propose a solution to a problem the group faces.

Without a strong established culture any major event (or even not so major) may pose a risk to the project that can influence the outcome of the project. The results from this research may benefit new start-up organisations on how to approach the change management process when they are exposed to major external events.

2. CURRENT MODELS AND DISCUSSION

2.1 Current models

The literature survey suggests that there may be knowledge gaps not yet addressed in the socio-cultural aspects of organisational behaviour, in particular the cultural change mechanisms present when an organisation is subjected to an external event.

The organisational life cycle model (Mintzberg as quoted in [10]) describes the four phases of an organisation's life as follows: formation, development, maturity and decline. The model suggests that the culture grows from the initial start of the organisation with the behaviour of the key role players setting the initial culture in place, but that over time the culture as described by the senior management can differ from the culture as described by the lower levels in the workforce. The model further suggests that organisation demise is suffered once politics play the most important role in the organisation, unless renewal efforts are made to enable continuation.

Three distinct cultures are identified in an organisation, i.e. the corporate culture, the professional culture and the social culture [10]. The model suggests that the social culture will determine how conflict between the managers and professionals will be approached and resolved. The social culture includes the common values within the organisation.

The culture of a professional organisation is the essence of its competitive advantage [10]. The model indicates that when the vision of an organisation is changed then risk may be induced that reduces the competitive advantage of the organisation. When the deep-seated organisational culture is disturbed then either culture drag or culture precession takes place [11], and the change may take a totally different and unexpected route than initially intended, therefore having the risk of impacting the competitive advantage.

The control theory model for the organisational transfer function implies that should any of the areas that define organisational culture be subject to change, then culture is influenced. The model further implies that culture can be measured and managed indirectly via these four areas, i.e. People, Management Systems, Technology and Organisational Structure [12]. The diversity of the make-up of these areas will have an impact on the common language available to facilitate change and have an implication for the stability or instability of the organisation during and after this change.

According to Denison et al [13] a strong organisational culture can be associated with increased organisational effectiveness. They further indicated that the aspects of organisational culture most critical to success included: empowering employees; having a team orientation; having a clear strategic direction and intent; and possessing a strong and recognizable vision.

2.2 Discussion on models

The current models as reported address the following aspects of organisational culture:

- The formation of organisational culture,
- The composition of organisational culture,
- The life cycle of organisational culture,
- The possible impact of a changed vision on competitive advantage due to the reaction of the existing organisational culture towards the change.
- The lack of a universal language within an organisation may result in instability during organisational changes or events.
- The culture of an organisation can be used to strengthen its agility and resiliency.

The models indicate that when an event or change is imposed on an organisation, the ability of the organisation to absorb the event can be influenced by its culture. It is not clear from the models whether the outcome of an event on the organisation can be predicted.

The following hypothesis is put forward for testing:

> Major events do impact the culture of an engineering organisation.

The models suggest that as organisations mature and are subject to events, the culture of an organisation can change. It is hoped that some indicators from the research can be used in similar engineering endeavours to manage culture proactively for success.

3. RESEARCH METHODOLOGY

The selected research design and methodology are discussed in this section. The strategy is provided. The rationale for the choice of design tools is presented and a critical analysis of the tools as selected is done.

3.1 Research strategy

According to Welman et al [14] professional groups such as market researchers have established explicit codes of conduct to which members should adhere. Although some practices can be considered ethical, it may still offend the respondents and the data collected may still generate unpleasant repercussions. Therefore participants should take part freely and prior consent obtained, in particular for interviews.

It is generally accepted that there is more confidence in causal relationships from true experimental research, than would be the case from non-experimental research [14]. If the phenomena being studied are orderly or regular then it would be possible to deduce some relationships between the variables in a non-experimental approach. During this research the intent is to study the behaviours of engineering staff over a period of time, and in particular when the organisation was subjected to a number of events. All the events were experienced by all staff at the same time within the engineering organisation and would therefore comply with the requirement of 'orderly and regular'.

The purpose of the research was to uncover facts, relationships and causations. Therefore the selection of the tools should preferably not support subjective information and specific care should be taken to obtain objective facts. The research strategy is based on a non-experimental hypothesis testing approach, and more specifically a field research approach consisting of a survey.

The target group for the survey should include responses from not only a few work areas within the engineering group but as many work areas as possible, including other entities related to the Case study company.

3.2 Choice of tools argued

A survey questionnaire can be used to obtain biographical details, typical behaviour, attitudes, opinions, beliefs and convictions from respondents [14].

The target group consisted of ex-Case study company staff members and members of other organisations closely involved with the company, such as the local electrical utility Client Office. Since it was not possible to determine the probability that any specific respondent would be included in the target group, a non-probability sampling method was used. Early warnings were sent to the target group to provide them with the purpose of the research and to identify their desire to take part in the research [14]. This sampling method has the advantage that the respondents with specific feelings and opinions about the research will then take part in the survey.

The target group can be regarded as a special target group if they may have a common loyalty [14]. Furthermore, if the target group is offered the opportunity to "tell their story", as experienced at the Case study company, it will bolster the response rate.

The survey considered the attitudes and responses of the respondents towards the events as experienced in the engineering organisation. A list of questions was provided to the respondents on which they responded on a five point Likert Scale. The aspects of organisational culture as reported by Denison et al [13] were used to formulate the survey questions to determine the change in culture as perceived by the respondents. The dimensions address an organisation's adaptability, mission, consistency and involvement. The semantic differential scale as developed by Osgood, Suci and Tannenbaum as reported in [14] was used in the survey. In addition to the survey questions the respondents were requested to provide limited biographical information, their qualifications, position at the Case study company, date when employment started and departure from the organisation.

A free text section was added to the survey questionnaire in which each respondent was requested to add feedback related to the events as experienced. With the ability to provide feedback in private and on their own time, it was expected that the respondents would express their true feelings and opinions without concern that censure might be placed on information provided.

3.3 Critical analysis of tool selected

The risk in using a survey tool is that it lends itself to deliberate deception and the with-holding of information from the respondents [14]. Deliberate deception is difficult to detect, but the responses from an individual can be compared with others in the same organisational or peer group to assist with the detection of deception. With-holding of information can be addressed by forcing the respondents to respond to all questions.

Measurement reactivity may also occur where respondents remember their previous responses and wittingly alter these or unwittingly respond differently to follow-up questions based on the new questions put to them. The potential effect of measurement reactivity cannot be determined by speculation and should be measured empirically [14].

3.4 Questionnaire design

The questionnaire used eight events that occurred over a six year period and fourteen questions per event were formulated for evaluation by the respondents. The questionnaire ends with a request to the respondents to provide comments and personal observations on how they have experienced these and any other events that influenced the Case study organisation's culture. The introduction to the questionnaire requires the respondents to position themselves in terms of skills, knowledge and work area within the organisation.

List of case events: The Case study organisation was subjected to many external and internal events over the period 1999 to 2010. The sample group that could still report on events prior to 2005 was very small, and had the additional risk that it could be regarded as a selective sample. Therefore the events prior to 2005 were excluded from the survey. Other events that were excluded were those that impacted only smaller areas and not the bigger group; and events and decisions that took place at levels beyond the visibility of the staff, typical at board and governmental level.

The purpose of the study was explained to the respondents as: to investigate the impact that events had on the culture of the organisation in general and the engineering organisation in particular.

The following eight events as listed in Table 1 were selected and regarded as being prominent enough to allow most respondents to recollect their related experience:

Table 1: List of Events

Event number	Year	Description of the Event
1	2005	Until early 2005 all staff was employed as contractors / consultants to work on the project. By April 2005 all staff was appointed as permanent employees with permanent employee benefits but with less flexibility in terms of working hours.
2	2005	Appointment of the CEO. The focus shifted from an engineering project to building a new energy development organisation, in particular expanding the staff functions and the non technical activities.
3	2006	Appointment of the Project Director. A project team and structures were formed, with a distinct style of project management.
4	2006	Appointment of an EPCM. The EPCM started to work on Balance of Plant designs and provided some of the project management services.
5	2006	Stop Work Order issued by the National Nuclear Regulator in October. The Stop Work Order was only lifted in early 2008.
6	2007	Departure of the General Manager: Engineering in November.
7	2008	Appointment of the new General Manager: Engineering in mid 2008 and the adoption of the Indirect Cycle plant in Oct/Nov 2008 as prime product offering.
8	2009	Announcement is made in November 2009 that no further funding will be received from government and restructuring may be on the cards.

Table 2: Organisational Traits and Survey Questions

Organisational Trait	Survey question
Involvement is defined as the building of human capability, ownership and responsibility.	• How did the event impact your ability to be in control of your work output? • How did the event impact your sense of loyalty to the organisation? • How did the event impact your sense of loyalty to your own team?
Consistency provides a central source of integration, coordination and control, and helps organisations develop a set of systems that create an internal system of governance based on consensual support	• How did the event impact the commonality in techno speak / techno language in your work environment? • How did the event impact interpersonal conflict? • How did the event impact interdepartmental conflict? • Was the vision(s) clearly communicated (verbal and walk the talk) by management? • To what extent did the event contribute to the organisation becoming a political organisation?
Adaptability is the ability to perceive and respond to the environment, customers, and restructure behaviours and processes that allow them to adapt.	• How did the event impact your ability to make or influence work related decisions? • Did the event change the participative climate in the areas where you were involved?
Mission relates to the defining of a meaningful long-term direction for the organisation. The mission tells employees why they are doing the work they do, and how the work they do each day contributes to the why	• Did the event have an impact on the sub-cultures (silos)? • How did the event impact the company values?

Formulation of the questions: According to Denison et al [13] a strong organisational culture can be associated with increased organisational effectiveness. The study by Denison et al [13] indicates that the aspects of organisational culture most critical to success included: empowering employees; having a team orientation; having a clear strategic direction and intent; and possessing a strong and recognizable vision. The questions presented in Table 2 were also formulated to reflect the four organisational culture traits as defined in [15].

The events were evaluated by the respondents by indicating their preference on the questions on a five point Likert Scale, graded between two opposites. Not all the questions were applicable to all the events and only the questions applicable to a particular event were put forward to the respondents.

Two questions were used as additional indicators whether there was a change in culture or not. These questions asked the respondent to make an assessment whether a change has taken place or not. The one question asked whether culture change had taken place after the event, and the other asked whether an important item related to culture had changed after the event, in this case the vision of the organisation.

Free text section: The questionnaire ended with a request to the respondents to provide comments and personal observations on how they experienced these and any other events that influenced the Case study organisation culture, and in hindsight what they thought could have been done to direct the culture. Since six of the eight events included in the questionnaire occurred prior to the end of 2007, the free text section allowed those respondents that became involved with the Case organisation only after 2007 with the opportunity to reflect on the culture they expected prior to appointment and their actual experience after being appointed.

4. RESULTS

When the Case organisation indicated in January 2010 that restructuring was imminent, a contact list was created by the organisation on which staff that either took early voluntary retrenchment packages starting in March 2010 or were eventually laid off by October 2010 could add their contact details. This list contained 391 names of which 252 provided e-mail addresses. A further 70 e-mail addresses were obtained via discussion groups on internet, making up the total of 322 to whom the survey questions were distributed on 8 August 2012. The mail delivery system reported a permanent failure on 12 of the 322 addresses and it is not known how many of the remainder of the 310 e-mail messages as sent actually reached the intended recipients. A total of 76 completed questionnaires were received between 9 and 31 August 2012. Results received after this date were not included due to time restrictions placed on the completion of the research. The results received on all questions are assumed to be random, unbiased and approximately normally distributed.

The respondents were grouped into staff that was employed for 5 years or less and staff that was employed for longer than 5 years. The following two durations of employment groups were identified:

- The staff employed for 5 years or less, identified as **L5 Group**.
- The staff employed for more than 5 years, identified as **M5 Group**.

The **All Group** includes the responses received from both the L5 and M5 Groups, combined into one group. Refer to Table 3 for a summary of the employment information of the respondents.

Table 3: Employment Information of the Groups

Characteristic	Group Identifier		
	L5	M5	All
Percentage of total	47.4	52.6	100%
Average Duration of Employment	3.8	8.5	6.2
Standard Deviation on Duration of Employment	1.0	2.1	2.9
* Average Start Date	2006.2	2000.8	2003.4
Standard Deviation on Start Date	1.1	0.6	3.2
* Average Departure Date	2010.0	2009.3	2009.6
Standard Deviation on Departure Date	0.6	1.2	1.0

* The questionnaire did not allow the respondents to add a specific month on which their employment started or ceased, only the year value. Therefore the year fractions as provided in the results do not imply a specific month of the year.

4.1 Summary of results

The results for the three groups are provided in Tables 4, 5 and 6 and are interpreted as follows:

Column A: "Cultural change as indicated by adjectives". Note that these results only indicate whether a change in culture has taken place or not.

- "Yes" – a cultural change has taken place after the event.
- "Inconclusive" – it is not clear from the results whether a cultural change has taken place after the event or not.

Column B: "Cultural change as indicated by high scoring". Note that these results indicate whether there was a strong or a weak change in culture.

- "Yes" – a strong indication that a cultural change was experienced by the group after the event.
- "Low" – a weak indication that a cultural change was experienced by the group after the event.

Column C: "Positive / Negative Counts". Note that these results indicate whether the culture was strengthened or weakened by an event.

- "Positive" – the number of positive counts indicate that the culture has strengthened (positive) after the event.
- "Negative" – the number of negative counts indicate that the culture has weakened (negative) after the event.

- "Inconclusive – the number of positive and negative counts are equal and it is not clear from the results whether the culture has weakened (negative) or strengthened (positive) after the event.

Table 4: Summary of Results – All Respondents Combined (All Group)

Event	Cultural change as indicated by adjectives A	Cultural change as indicated by high scoring B	Positive / Negative Counts C
1	Yes	Yes	Negative
2	Inconclusive	Yes	Inconclusive
3	Inconclusive	Yes	Negative
4	Yes	Yes	Negative
5	Yes	Yes	Negative
6	Yes	Yes	Negative
7	Yes	Yes	Negative
8	Yes	Yes	Negative

With reference to the All Group, Table 4, the results based on measured adjectives indicate that only events 2 and 3 were inconclusive regarding cultural change. The experience as indicated by the high scoring indicated that every event resulted in a cultural change. The group indicated that organisational culture was weakened by all the events, except for event 2, as shown by the positive/negative counts. Event 2 was inconclusive on whether it strengthened or weakened the culture.

Table 5: Summary of Results for Respondents Employed for 5 Years or Less (L5 Group)

Event	Cultural change as indicated by adjectives A	Cultural change as indicated by high scoring B	Positive / Negative Counts C
1	Events 1 and 2 sample sizes too small		
2			
3	Inconclusive	Low	Negative
4	Yes	Low	Negative
5	Yes	Yes	Negative
6	Inconclusive	Yes	Negative
7	Inconclusive	Yes	Inconclusive
8	Yes	Yes	Negative

With reference to the L5 Group, Table 5, the sample sizes of events 1 and 2 were too small and were excluded from the results. The results for L5 based on measured adjectives indicate that events 3, 6 and 7 were inconclusive regarding cultural change. The experience as indicated by the high scoring indicated that the events resulted in a cultural change. The group indicated that organisational culture was weakened by all the events, except for Event 7, as shown by the positive/negative counts. Event 7 was inconclusive on whether it strengthened or weakened the culture.

Table 6: Summary of Results for Respondents Employed for more than 5 Years (M5 Group)

Event	Cultural change as indicated by adjectives A	Cultural change as indicated by high scoring B	Positive / Negative Counts C
1	Yes	Yes	Negative
2	Inconclusive	Yes	Inconclusive
3	Inconclusive	Yes	Negative
4	Yes	Yes	Negative
5	Yes	Yes	Negative
6	Yes	Yes	Negative
7	Yes	Yes	Negative
8	Yes	Yes	Negative

The results for the M5 Group as provided in Table 6 are the same as for the All Group. The noticeable differences between the L5 and M5 Groups are the results for Events 6 and 7. Those employed for 5 years or less indicated that the cultural changes for these two events were more inconclusive than for those employed for more than 5 years.

In summary an indication of the impact on the organisational culture by the events as measured on Groups L5 and M5 is provided in Figure 1: Impact on Culture, where the mostly negative impact except for event 2 and partially for event 7 group L5 is notable.

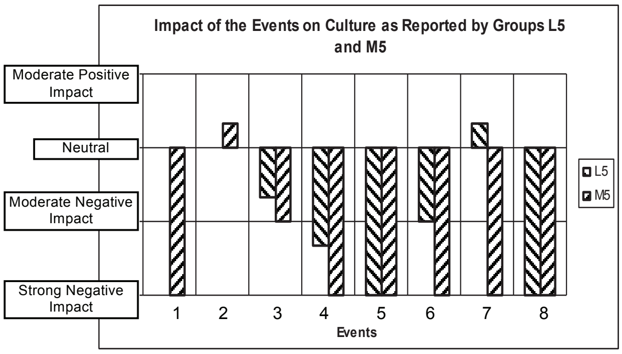

Figure 1: Impact on Culture

5. CONCLUSIONS AND RECOMMENDATIONS

The hypothesis tested stated: "Major events do impact the culture of an engineering organisation". The results show that most of the events have changed the culture of the engineering organisation and in all instances where a change was measured, the culture had weakened. As can be expected, Event 8 had a strong impact on the change in the engineering organisation's culture; it was when the announcement was made that restructuring (and lay-offs by implication) is imminent.

The results from events 1, 4, 5, 6, 7 and 8 support the hypothesis that major events can change an engineering organisation's culture. It is inconclusive whether events 2 and 3 have changed the engineering organisation's culture or not.

All the events, except for Event 2, have weakened the engineering organisation's culture. Although the measurements on Event 2 indicate that the event was inconclusive regarding its ability to change the culture, the indications are that the event tended to weaken the culture. Those employed for longer than 5 years experienced the cultural changes due to events 6 and 7 more distinctly and more negatively than those employed for a lesser number of years.

5.1 Area for further investigation

Engineering organisations are spending focussed and expensive effort to reduce the risk of not achieving their project goals and to ensure an effective workforce. The question can be asked whether the culture can be managed to endure the unforeseen and improve the predictability in outcome.

Further investigation will improve the understanding of how a major event may impact the culture of an engineering organisation. Furthermore if it is assumed that those working longer within an organisation are in a position to express themselves more strongly during or prior to an event and thereby influencing the cultural change, then understanding how these employees may impact the dissemination of the new culture into the organisation is important. The ideal situation will be to prepare the organisation in advance of an upcoming visible and even not so visible event with the intent to stay on course or to adjust course to be able to survive the event.

6. REFERENCES

[1] N.N. Taleb: *The Black Swan: Second Edition: The Impact of the Highly Improbable,* New York, Random House, pp. xxii, 2010.

[2] E.H. Schein: *Organizational Culture and Leadership. 2nd Ed.*, San Francisco: Jossey-Bass, 1992.

[3] O. Manetje and N. Martins: "The relationship between organisational culture and organisational commitment" *Southern African Business Review*, Vol. 13 No. 1, pp. 87-111, 2009.

[4] B. Oxtoby, T. McGuiness and R. Morgan: "Developing Organisational Change Capability" *European Management Journal*, Vol. 20 No. 3, pp. 310-320, 2002.

[5] C. O'Reilly: "Corporations, culture, and commitment: Motivation and social control in organizations" *California Management Review,* Vol. 31, pp. 9-25, 1989.

[6] O.N. Solinger, W. Van Olffen and R.A. Roe: "Beyond the Three-Component Model of Organizational Commitment" *Journal of Applied Psychology,* Vol. 93 No. 1, pp. 70-83, 2008.

[7] L. Storm and G. Roodt: "Die verband tussen Organisasie sosialisering en Organisasie verbondenheid" *South Africa Journal of Industrial Psychology,* Vol. 28 No. 1, pp. 14-21, 2002.

[8] E. Ries: *The Lean Start-up: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses,* Crown Publishing, 2011.

[9] E.H. Schein: "The Role of the Founder in Creating Organizational Culture and Organizational Dynamics" *Periodicals Division, American Management Associations,* pp. 13-28, Summer 1983.

[10] A.P. Sage: "Organizations, Organizational Cultures, Catalysts for Change in Organisations" *International Conference on Systems, Man and Cybernetics,* pp. 720-724, 17-20 October 1993.

[11] C.M.S. Kirke: "Organisational Culture: Can System Designers Ignore it?" *The IEE and MOD HFI DTC Symposium on People and Systems – Who are We Designing for?,* pp. 9-15, 16 to 17 November 2005.

[12] M.A. Friedlander and S.A. Evans: "Influence of organizational culture on human error" *Sixth Conference on Human Factors and Power Plants, 1997. 'Global Perspectives of Human Factors in Power Generation'*, pp. 12/19 – 12/22, 13 June 1997.

[13] D.R. Denison, S. Haaland and P. Goelzer: "Corporate Culture and Organizational Effectiveness: Is Asia Different from the Rest of the World?" *Organizational Dynamics,* pp. 99-109, February 2004.

[14] C. Welman, F. Kruger and B. Mitchell: *Research Methodology, 3rd Ed.*, Cape Town, Oxford University Press, pp. 69, 120, 149, 152, 154, 157, 2005.

[15] D.R. Denison: Understanding the model. Retrieved 10 October, 2012, from http://www.denisonconsulting.com.

A SOUTH AFRICAN PERSPECTIVE OF THE REQUIREMENTS DISCIPLINE: AN INDUSTRY REVIEW

A Marnewick*, JHC Pretorius and L Pretorius*****

** Faculty of Engineering and the Built Environment, University of Johannesburg, Corner Kingsway and University Road, Auckland Park, 2006, South Africa. E-mail: amarnewick@mweb.co.za*
*** Faculty of Engineering and the Built Environment, University of Johannesburg, Corner Kingsway and University Road, Auckland Park, 2006, South Africa. E-mail: jhcpretorius@uj.ac.za*
**** Department of Engineering and Technology Management, Graduate School of Technology Management, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa. E-mail: leon.pretorius@up.ac.za*

Abstract: The requirements discipline is at the heart of systems engineering, software engineering and business analysis. Across these three communities common practices, approaches and techniques are available in literature. However, little application data is available on how practitioners apply these common practices during the requirements engineering process in practise. To generate data on how practise carries out the requirements engineering process, a two-part survey was conducted. The first part of the survey is reported here and investigates how practitioners carry out the requirements engineering process. The survey was completed by individuals involved in practice as requirements practitioners. This survey and its results offer opportunities to increase industry relevance of research outcomes and identified focus areas for practitioners, including software system developers, to exploit and increase their effectiveness during requirements activities.

Keywords: Requirements, requirements engineering process, survey

1. INTRODUCTION

It is irrelevant whether a software solution or system is developed or bought. For any software system to be useful to the users and business, the developers must understand what the software solution or system is intended to achieve for the users and business [1]. The purpose of requirements engineering is to maximise the likelihood that a developed or bought solution or maintenance initiative will deliver solutions that function as desired. The requirements discipline has different origins and different approaches to capture requirements during the solution development due to the relevance of requirements of different communities such as systems engineering, software engineering and business analysis. However, an understanding of the problem is always required before the development of a solution [2].
The importance of requirements engineering is acknowledged by the engineering literature [3-5]. Requirements are quoted as the input to the software or systems engineering process. If the requirements delivered during requirements development is not of high quality, the solution derived from the requirements have a risk to not achieve what it was intended for [6-8]. Even the best requirements practices cannot make up for inaccurate requirements [9].

The research community acknowledges the importance of requirements. However, industry itself is still facing many challenges in practice with the requirements engineering process during the delivery of a solution [10]. Industry reports, surveys and research continuously quote poor requirements as the main contributor to failed projects, along with the cost and rework implications of the requirements errors [9, 11-23].

In order to contribute relevant knowledge to the requirements discipline, research focus areas have been highlighted by Cheng and Atlee [24]. One of these focus areas is the importance of collaborative partnerships between researchers and practitioners to increase the industry relevance of research [24-26]. A second focus area highlighted by Cheng and Atlee [24] was to document empirical research on how well requirements engineering research addresses industrial problems.

In response to the focus areas highlighted, the research reported on in this paper was aligned to increase the industry relevance of the research outcomes; the first step of the research undertaken was to gather data on how practitioners apply existing requirements engineering knowledge in practice. Results existed for some of the international industries as per literature [27-29]. For the South African requirements industry, there is little or no recent research data describing how practitioners execute the requirements process.

This paper's main objective is to explore and document how practitioners also focusing on software engineering execute the requirements engineering process, as little shared knowledge within the South African context is available in research. The first section summarises how literature suggests the requirements engineering process should be executed, after which the research design is explained. The results collected to derive a description of how practitioners execute the requirements engineering process are presented. Finally, key findings are summarised from results.

2. LITERATURE REVIEW

A requirement is a collection of capabilities originating from users and stakeholders (organisational, legislation, and industry standards) that all must be met by the solution to solve the problem or achieve the objective [4]. The stakeholders must be involved for the proposed solution to solve the problem or achieve the objective. It is often assumed that the stakeholders already know what the requirements are at the beginning of a problem solution. As far back as the 1970s, Bell and Thayer [30] cautioned that "*requirements for a system, in enough detail for its development, do not arise naturally. Instead, they need to be engineered and have continuing review and revision*".

When a solution is implemented, different engineering process models are available to use and implement a possible solution as discussed in literature [31-36]. During the problem solution process, the assumptions of problem complexity, the availability of knowledgeable users and the type of the required solution have a direct impact on which of the engineering process models would be most appropriate [37].

The selection of an engineering model will impact on how the requirements will be documented. In a small environment, with a rapidly changing marketplace where people are capable and work collaboratively, agile methods would be appropriate and the requirements will be documented less formally [38]. In an environment where people do not work collaboratively, requirements will have to be documented formally to minimise misunderstandings [38].

Although the method of requirements documentation could vary, based on the selected engineering model, the requirements have to be understood to solve the problem. The goals of the requirements engineering process as described by Pohl [39] are:

- Transforming unclear user needs into a complete system specification;
- Transforming informal knowledge into formal representations;
- A common agreement on the specification from the various personal views.

The term "requirements engineering process" is defined as the systematic process of developing requirements through an iterative co-operative process of analysing the problem, documenting the resulting observations in a variety of representation formats and checking the accuracy of the understanding gained [40, 41]. This definition can be illustrated by a model explained by Wiegers [42] that divides the requirements engineering domain into two domains as illustrated in Figure 1.

The initial domain is requirements development to derive a complete user specification. Once the specification is produced, it must be agreed upon by all stakeholders. These are the baseline requirements that will be used to build the solution. Once the baseline requirements have been established and agreed upon, changes should be managed to ensure that all stakeholders stay in agreement. This is presented as the requirements management domain.

The common activities in the requirements development are elicitation, analysis, specification and validation, and change management [4]. These activities are executed in different forms, namely linear, incremental, non-linear, spiral or adaptive processes as discussed by Van Lamsweerde [37], Wiegers [42], Kotonya and Sommerville [43] and Macaulay [44].

Requirements elicitation is the discovery of knowledge about the problem that must be solved [37]. The literature mentions many approaches or techniques to determine the required knowledge [25, 26, 45]. The choice of elicitation technique depends on resource availability, information required and the types of problems that need to be solved.

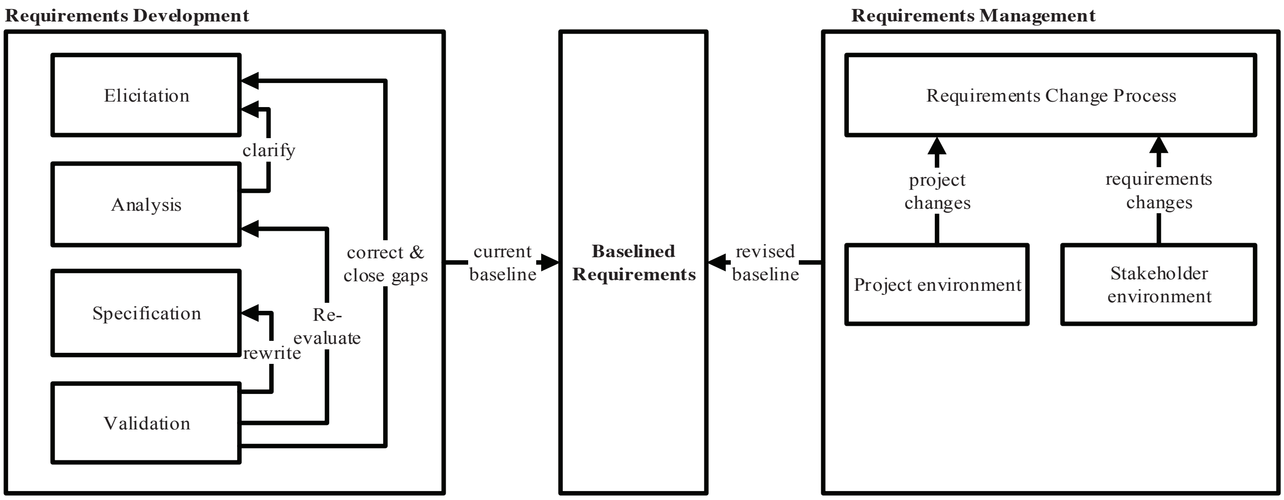

Figure 1: Requirements engineering domain [42]

Guidelines on how to choose the appropriate elicitation technique are provided by Robertson and Robertson [1], Robertson [46], Maiden and Rugg [47] and Zowghi and Coulin [25].

The analysis activity considers all the elicited information and generates a list of potential requirements. The objective of the requirements analysis step is to increase understanding, identify problems and search for inconsistencies in the list of produced requirements [43, 48]. Models are generated to understand requirements. Various modelling techniques or notations are available as discussed by Nuseibeh and Easterbrook [45].

Once all the information has been elicited and requirements have been analysed and modelled, the findings from these two activities should be documented in the specification document. The specification document can be generated in various formats or languages [26, 49]. Many best practice templates are available to provide guidelines on what information should be presented in the specification document [50-52]. Quality elements of a specification have been detailed by standards [51-53].

The output of the requirements development is a set of commonly agreed upon requirements by all stakeholders as presented in the specification document. These form the baseline requirements that will be used to build the solution. The priority of each requirement should be discussed by and agreed upon with the stakeholders to identify the most important requirements with the greatest impact on solving the problem. Techniques that are available are discussed by Berander and Andrews [54], Hansen et al. [26] and Cheng and Atlee [24]. The requirements must be validated for completeness and conflict and should reflect what needs to be done to solve the stakeholders' problem. Once a baseline set of requirements has been agreed upon, requirements must be managed and changes should be analysed based on the impact that they will have on the solution.

Each of the requirements communities also has professional bodies that provide guidelines on how to execute the activities during the requirements engineering process [53, 55, 56].

A summary was presented of knowledge available from literature about the requirements engineering process, the tools, techniques and modelling methods and guidelines. Next an industry review was done to describe how practitioners in South African industries execute the requirements engineering process. This generated knowledge on how the requirements engineering process is executed in practice. This knowledge will be utilised to determine whether the existing requirements engineering knowledge base is actually migrated into practice.

The most appropriate research method had to be used to obtain a description of how the requirements engineering process is executed by practitioners, as well as how practitioners behave during the execution of the process. The research methodology selection is described in the next section.

3. RESEARCH METHODOLOGY

The purpose of the industry review was to provide a description of how the requirements process is performed within the South African requirements engineering community. A survey research method was evaluated as it is characterised as producing quantitative descriptions on some aspects of a studied population, which include the examination of the relationships among different variables [57]. Survey research is appropriate if there is not adequate data available. This is the case with South African requirements practice. The target population of requirements practitioners was accessible and a portion of the data was personal and self-reported data of practitioners [58].

Real-world data was required from as many practising respondents as possible within the requirements discipline in a short time. The strengths of a survey confirmed that a survey would be the appropriate method [57-61]. A rigid systematic approach was followed to ensure that the survey would be conducted rigorously and in an unbiased manner. This process followed is illustrated in Figure 2and has been derived from literature [58, 62, 63].

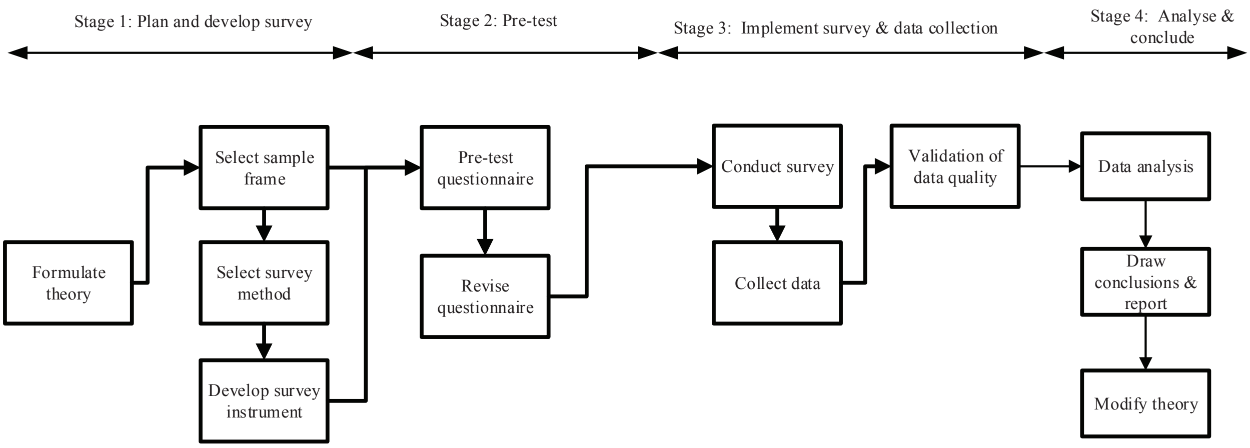

Figure 2: Survey process

3.1 Planning and development of the survey

The objectives of the industry review were to produce a quantitative description of how the requirements process is executed by practitioners, with practitioners belonging to multiple communities. The requirements discipline crosses multiple communities, namely systems engineering, software engineering and business analysis [2]. The review would also determine whether the available tools and methodologies as identified in literature are used or known by the practitioners.

The sample frame refers to how the population was constituted. The factors for consideration to ensure a complete survey design are listed by Fowler [57] and Sapsford [64]. The target population was only practitioners that were responsible for any activities in the requirements engineering process. The requirements practitioners were classified as a difficult-to-reach population as they had not been previously identified [65]. In the case of hard-to-reach populations, snowball, targeted, time space and respondent-driven sampling are suggested to access these hidden populations [66].

Snowball sampling is a very useful methodology to conduct exploratory, qualitative and descriptive research, especially where a high degree of trust is required for initial contact [66]. This sampling makes contact with a small group of relevant people and then uses these contacts to establish new contacts with others [67]. Snowball sampling was selected as the preferred sampling method as the population was difficult to research and a description of the population was the main objective of the industry review [68].

Three different approaches were followed to find respondents within the requirements practitioners' social network and to start the referral chain of the snowball sampling:

- Relationships were established between the researchers and individuals who were either chief information officers to whom requirements practitioners typically reported or individuals responsible for requirements practice. These individuals were contacted prior to the survey to obtain their cooperation. The researcher used multiple individuals with established relationships to start a referral chain within as many organisations as possible.
- A list of individuals practising as requirements practitioners within the industry were known to the researchers. These known practitioners were contacted directly and requested to complete the survey. They were also requested to distribute the survey link to their network of requirements practitioners.
- Two professional organisations were contacted and requested to send the survey to their members. The first professional organisation was Computer Society South Africa, which confirmed that they distributed the survey to all their members that had an interest in requirements practice. The second professional organisation contacted was the International Institute of Business Analysis South Africa. They were requested to send the survey to all their members as their focus is on analysis, which is within the requirements engineering space.

This was done to ensure that the sample frame included a non-homogeneous set of requirements practitioners across the industry and to minimise potential bias which could emerge due to a sample frame that did not fully represent the population [62].

After careful consideration of the advantages and disadvantages of each survey type, the Internet survey type was selected as a collection tool. It enabled the collection of data electronically that would be ready for analysis at a low cost within a very short time. The respondents' confidentiality was also protected. The population coverage would not be impacted by the Internet access as the requirements practitioners normally have Internet access within their work environment.

To explore and document how the practitioners execute the requirements engineering process, data was collected from the first two sections in the survey. The survey questions were based on knowledge collected from previous studies and a literature review [27, 69, 70]. Data was collected to profile the participants and to derive a description of the requirements engineering process. The focus was on the input to the requirements process, requirements activities and the quality of the output of the requirements process.

3.2 Pre-testing
When the questionnaire was completed, it was tested to ensure that it would work under real-life conditions [57]. As it was a self-administered instrument, it was first configured on the online survey platform that would be used to collect the data. Once ready, various pre-tests were done, including survey duration and pilot testing.

3.3 Implementation of survey and data collection
The survey was opened on the Internet via the platform used. The platform service provider was Survey Monkey. The data collection was facilitated by the platform on which the survey was configured. Data was automatically collected in various electronic formats by the platform.

During the design of the questionnaire, elements for good questionnaire design were taken into consideration to address data quality. Additional reliability and validation tests were done to evaluate the survey instrument before implementation.

3.4 Analyse and conclude
Data analysis involved the data collected being reduced and summarised into a usable format and patterns in data being identified [59]. The data analysis followed four stages as described by Quinlan [71]:

- Stage one was to engage in a descriptive analysis of the collected data.
- Stage two was to interpret the data.
- Stage three was to use the results of stages one and two to draw conclusions from the data.
- Stage four was the theorisation stage. In this stage the results from stages one, two and three were used to apply to existing knowledge and produce new theory.

Stages one to three of the data analysis will be discussed in the survey results, as the purpose of this article is to present the application data collected on how practitioners apply common practices during the requirements engineering process.

4. SURVEY RESULTS AND DISCUSSION

A total of 127 responses were received from the requirements practitioners. The main survey results are provided in a graphical format for summary purposes.

4.1 Participant characteristics
The survey respondents were mainly from the finance and banking (33%), ICT (23%) or government public sector and defence industries (9%). Although the majority of respondents were from these industries, the data confirms a presence of requirements practitioners across all industries as illustrated in Figure 3.

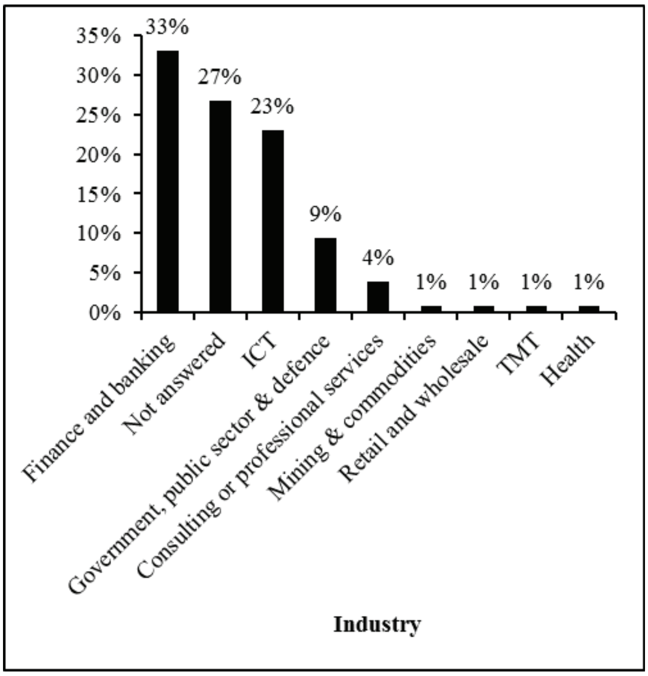

Figure 3: Practitioners' industry background

The requirements practitioners had very diverse job descriptions as summarised in Figure 4. There was, however, one specific job description that was commonly used, i.e. business analyst. This grouping also includes the job description of senior business analyst and constituted 47.2% of the respondents. The 'Other' group of job descriptions ranged from developers, project managers and specialists to programme managers.

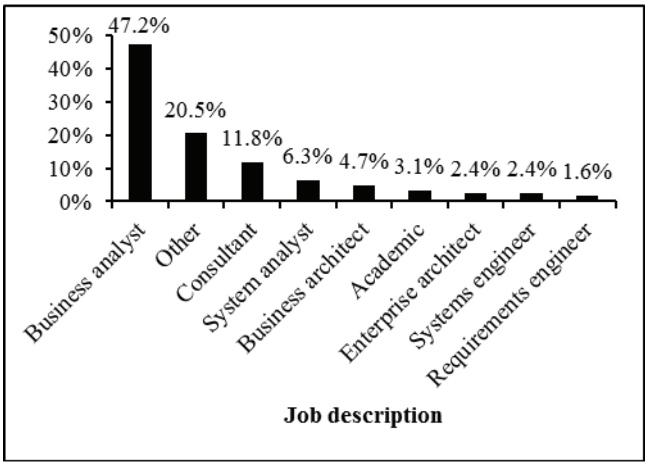

Figure 4: Respondents' job descriptions

The respondents were asked how many years' experience they had as a requirements practitioner. Only 3% of the respondents had less than one year of experience, 19% had between one and three years' experience and 22%

had between four and five years' experience. The majority of respondents were practitioners, with 56% of respondents with 6 years and more of experience.

Focused tertiary programmes and industry-specific certification for requirements engineering are available [53, 72, 73]. However, the practitioners surveyed held very diverse tertiary degrees with no uniform education amongst them as can be seen in Figure 5.

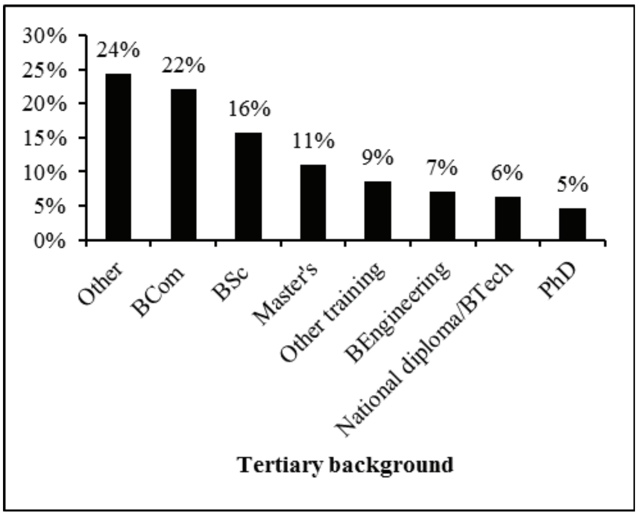

Figure 5: Practitioners' tertiary background

Respondents with BCom or BSc degrees had degrees within the information and communication technology (ICT) industry. Respondents with degrees outside these qualifications are listed as 'Other'. The industry certifications that supported the practitioners' qualifications were very diverse. One respondent was a certified system engineer. Six respondents were certified business analysis professionals. Additionally, five respondents held SAP certifications; one held a PRINCE certification and three respondents held Microsoft certified systems engineer or development certifications.

This concludes the analysis of the respondents. The following section focuses on how the practitioners in the requirements discipline execute the requirements process, including the usage of best practice tools and techniques.

4.2 Requirements engineering process

To describe how practitioners execute the requirements engineering process, data was collected to measure (i) whether any process models are adopted in practice, (ii) the involvement of practitioners during the activities of the requirements engineering process, including the tools and techniques adopted, and lastly, (iii) the quality delivered as an output of the requirements engineering process.

Requirements engineering process models: In 88% of the cases, a formal approach was used as guidance during project implementation versus the 12% of cases where no approach was followed. An analysis of the formal approaches indicates that incremental development was the most popular approach, with 33% of the respondents using it, followed by the prototyping (24%), agile (19%) and waterfall (13%) approaches as summarised in Figure 6.

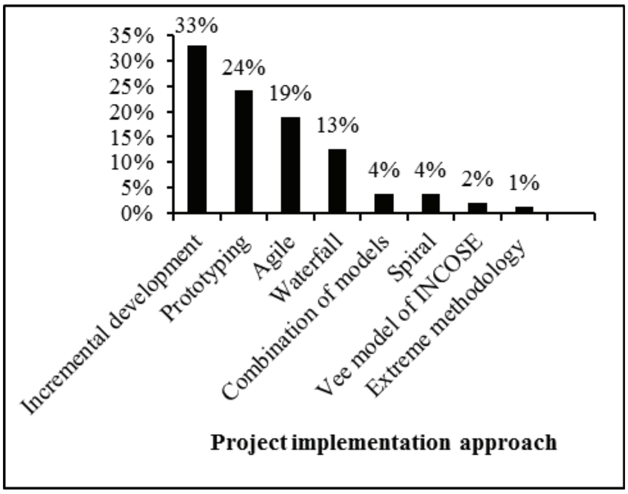

Figure 6: Project implementation approach

Requirements engineering process activities: A study in Australia investigated the barriers experienced by business analysts that prevent them from effective requirements analysis [27]. The original survey questions and data used by Wever and Maiden [27] were shared with the researchers. This was integrated into the questionnaire as a basis to determine the practitioners' involvement during the requirements activities.

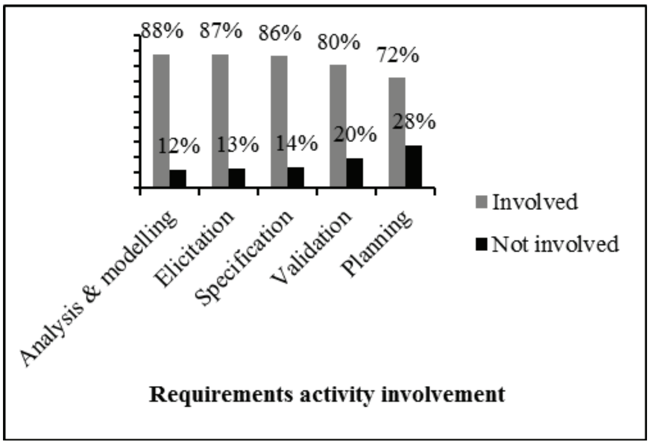

Figure 7: Practitioners' involvement in requirements activities

The practitioners appeared to be very consistently involved in the requirements activities. Most were involved in the analysis and modelling activity but were the least involved in the planning activity. If this is compared with the results of the Australian study by Wever and Maiden [27], no similarities can be identified. The business analysts in Australia reported to be engaged inconsistently across the requirements activities. In that study [27], the practitioners ranged from being the most involved in requirements elicitation (above 80%) to being the least involved in requirements validation (below 60%).

The potential reason for a more consistent involvement across requirements activities in this study compared with the Australian study could be attributed to the fact that the respondents were all experienced practitioners with six years or more experience. The Australian study, on the other hand, reported that the majority of respondents' experience was between one and three years [27].

In the cases where the respondents indicated that they were not involved in the specific requirements activity, an additional question captured the reasons for this non-involvement. Of the 127 respondents, 11 reported that planning was not part of their role. Two respondents mentioned a lack of resources and therefore planning was simply not done. One respondent mentioned that planning was not relevant to the project he was involved in. A further two respondents said: *within my organisation, there is no formal process in place to engage in this activity*, suggesting that planning was not done at all during project implementations. It was also mentioned that in some instances planning was done. However, this was inconsistent and without any continuity. This indicates that planning did not guide the implementation of the requirements during the project at all and was done on an ad hoc basis. In one instance, the planning of the project was done by external vendors. A few respondents also mentioned that they were not given the opportunity to take part in the planning or were not asked to do so.

The reasons provided for the respondents' non-involvement during requirements activities other than planning can be attributed either to the activities not being part of the respondents' role or not relevant to the project.

In addition to the respondents' involvement, the respondents were questioned about the deliverable of each activity and the tools or techniques used to produce the specific deliverable.

Requirements planning: During the planning activity, the requirements practitioner is responsible for creating a requirements management plan which is the key input to the overall project plan [74, 75]. From the results as illustrated in Figure 8, practitioners indicated that the project schedule was a deliverable of the requirements planning activity and not a requirements management plan as suggested by literature. The second important deliverable produced during the planning activity, according to the practitioners surveyed, was a scope of work or problem definition. Thirdly, the actual requirements specifications were mentioned by 10% of respondents as a deliverable, followed by a requirements management plan (9%).

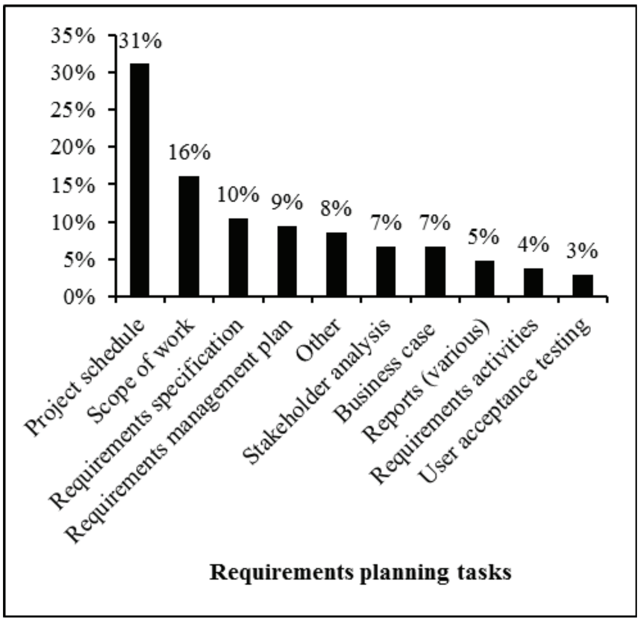

Figure 8: Summary of requirements planning task deliverables

Requirements elicitation: The elicitation activity is the discovery of the knowledge about the problem that should be solved. As time passes by during the problem-solving process, the sources change [76]. During the early stages conversations with colleagues and personal experience were used as illustrated in Figure 9. During the latter stages of the process, textbooks, codes and standards, industry newsletters and conversations with academics were used [77].

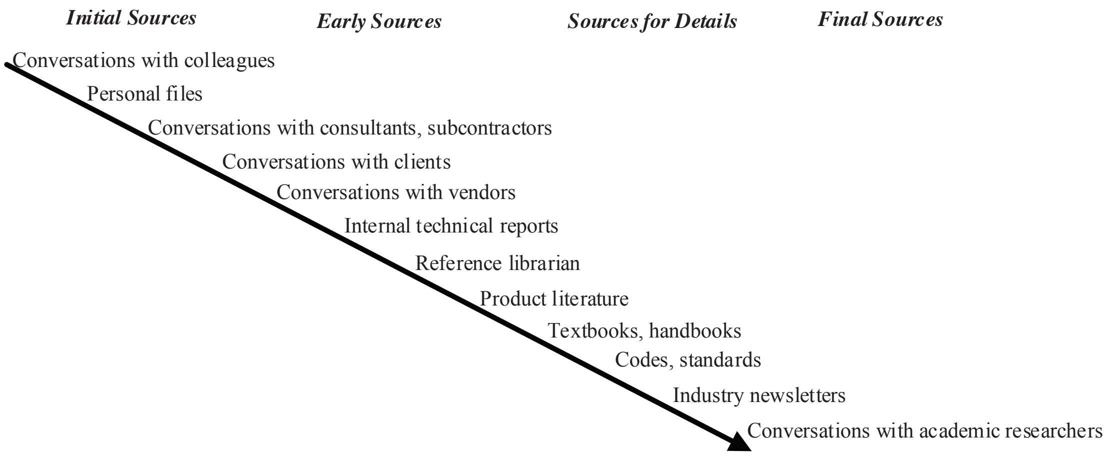

Figure 9: Information-seeking behaviour during the problem-solving process [77]

From the data collected as illustrated in Figure 10, academic, industry and textbook sources were used in a few cases (4%, 5%, and 7%, respectively). The preferred sources which were used by the practitioners were either personal experience or conversations with customers and colleagues, which are classified as initial sources as illustrated in Figure 9. The sources used by the practitioners indicate that the information-seeking behaviour did not follow a typical problem-solving process where the sources changed over time.

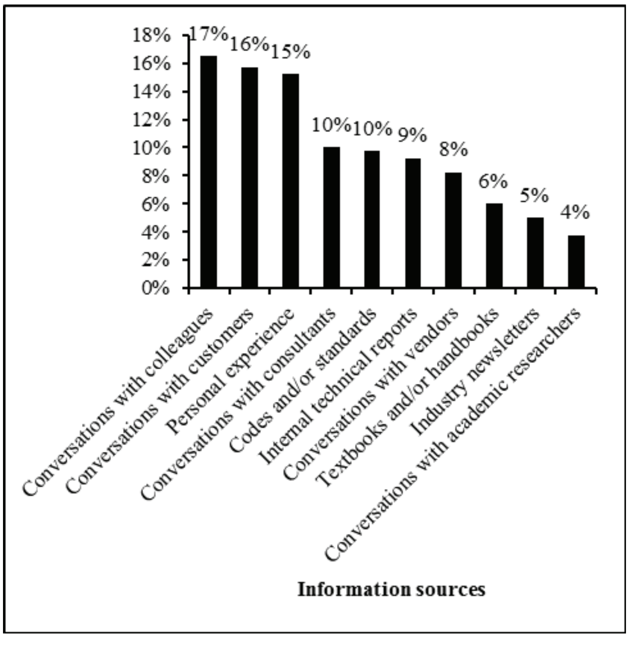

Figure 10: Practitioners' research patterns during elicitation

A list of techniques used to elicit information was derived from literature. The practitioners were requested to rate the usages of each technique on a scale of 'use', 'never use' and 'never heard of'. The results are displayed in Figure 11.

The practitioners utilised interviews, document analysis, groupwork and brainstorming extensively at 98%, 95%, 92% and 90%, respectively. Workshops (86%) and scenarios (81%) were also techniques that were used by the respondents. Prototyping and observation were utilised by 70% and 68%, respectively. Card sorting, ethnography, laddering and repertory grids were used by only a few practitioners.

The respondents preferred traditional techniques such as interviews and brainstorming during requirements elicitation. Literature suggests that the most appropriate combination of techniques must be considered to ensure that all types of knowledge are acquired [47, 78]. A framework for selecting the most effective techniques to access non-tacit, semi-tacit and tacit knowledge has been developed by Maiden and Rugg [47].

The results and comments by the respondents indicate that practitioners surveyed had a set of preferred techniques that were used. They selected techniques based on what technique was known, and not what the most appropriate combination of techniques was to ensure that all types of knowledge were acquired.

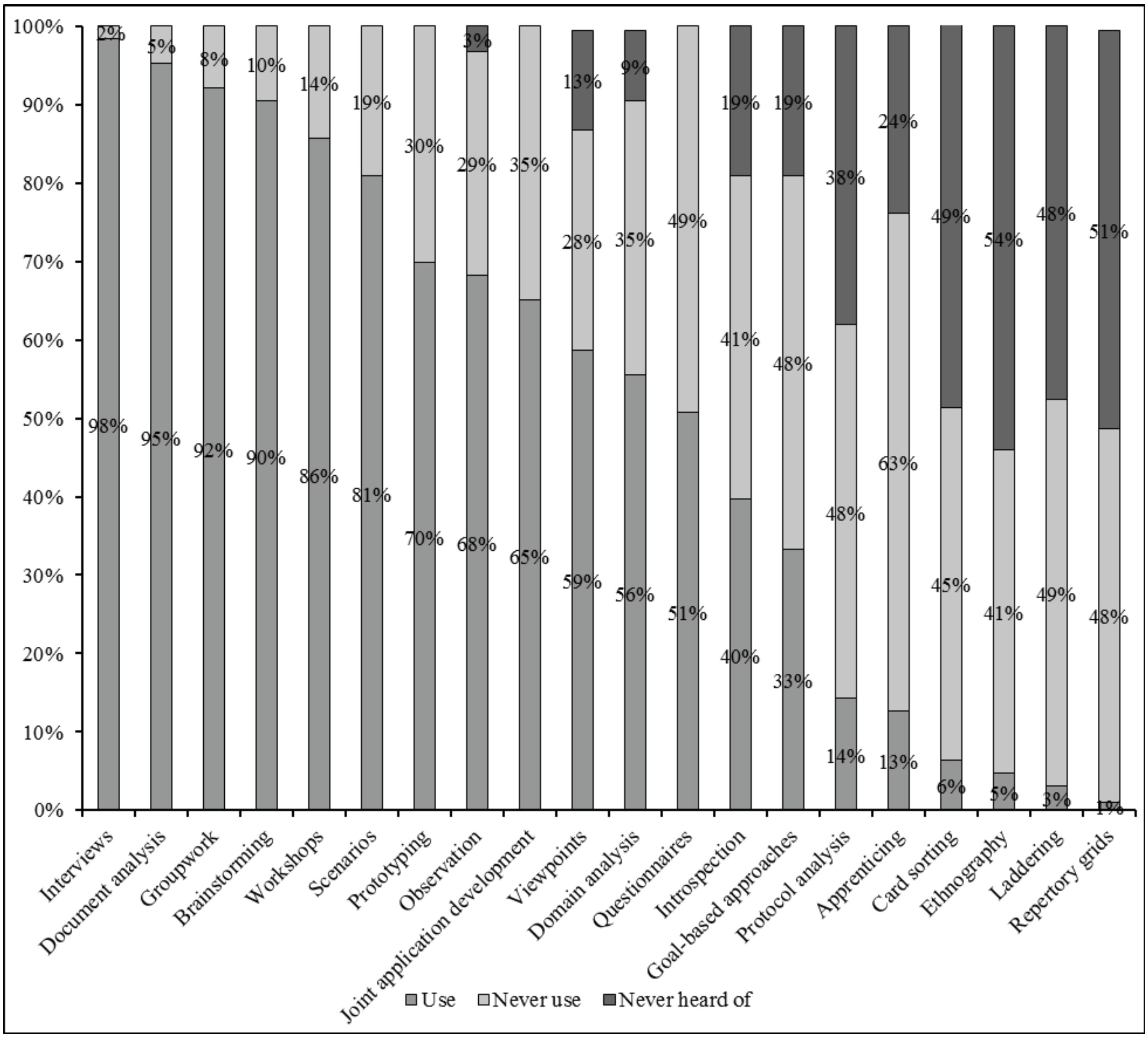

Figure 11: Elicitation techniques usage

Requirements analysis and modelling: The analysis activity analyses all the elicited information and generates a list of potential requirements [48]. The objective of the requirements analysis step is to increase understanding, identify problems and search for inconsistencies in the list of requirements produced [43, 79]. Models are generated to understand the requirements.

A list of techniques used during the analysis and modelling activity was derived from literature. The practitioners were requested to rate the usages of each technique on a scale 'use', 'never use' and 'never heard of'.

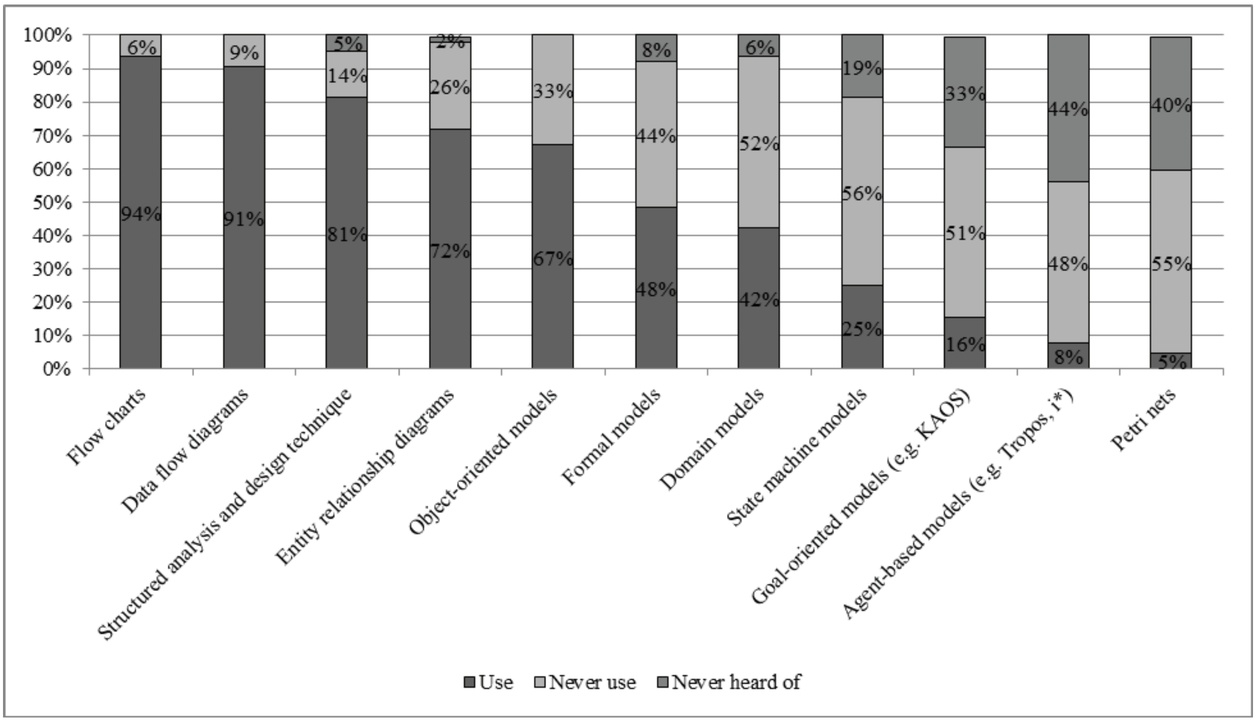

Figure 12: Analysis and modelling techniques

The techniques preferred were flow charts (94%), data flow diagrams (91%), structured analysis (81%) and entity relationship diagrams (72%). Petri nets, agent-based models, goal-oriented models and state machine models were either not in use or not known by most of the respondents.

Requirements specification: The specification is used to facilitate communication and should be complete to ensure that a fit-for-purpose working solution can be produced from the requirements [80]. Of the 127 respondents, 61 indicated how the specifications were generated. In 79% of the cases template guidelines were available to assist the practitioners in generating the specifications.

In 8% of the cases, practitioners used formal notation to present the requirements and 59% used semi-formal notation to present the requirements. The balance (33%) used informal presentation such as natural language to present the requirements. These results were compared with the preferred elicitation techniques of practitioners, which indicates a potential mismatch. The preferred techniques of respondents all deliver outputs in natural language except brainstorming, which will deliver a semi-formal format [46]. However, 59% of the respondents indicated that they produced a specification with a semi-formal presentation. The natural language percentage was expected to be higher, as the elicitation techniques used by practitioners deliver output in natural language. The practitioners stated that they presented their specifications in semi-formal format, but the elicitation techniques that they used did not produce semi-formal notation. It could be that the practitioners generated models from elicited information during the analysis and modelling activity in order to generate a semi-formal notation specification.

Requirements management: The respondents were questioned if a software tool was used to manage the delivered specifications. In total only 12.8% of the respondents used a software tool to develop or manage requirements. Three of the tools mentioned by the respondents used as requirements management tools can actually be classified as requirements management tools, i.e. Blueprint Requirements Center™ 2010, Enterprise Architect and IBM Rational RequisitePro. The other tools mentioned vary from Visio - a modelling tool; SharePoint - a document management tool; System Architecture - a tool used to model business operations and systems. From these results it can be concluded that requirements management tools were not generally used by the practitioners to present or manage requirements.

Requirements quality: The desired output of the requirements engineering process is a set of commonly agreed upon requirements by all stakeholders [39]. The literature prescribes that a requirement should consist of eight characteristics to ensure high quality [51, 81]. As these characteristics are subjective, an ordinal scale was used to measure the quality of the requirements. The scale elements used to determine the quality of the requirements were based on the eight characteristics of a quality specification as per Table 1.

Table 1: Quality scale

Characteristics	Question
Correct	Have all the requirements been validated by the source of the requirement, i.e. typically the stakeholder?
Unambiguous	Was there a single interpretation for each requirement to enable the common understanding by all stakeholders?
Complete	Were all the required requirements present in the specification ensuring a workable solution fit for purpose by the user?
Consistent	Did some requirements conflict with other requirements or with higher level system or business requirements?
Ranked for importance	Were all the requirements prioritised based on importance or in terms of expected changes associated with the requirement?
Verifiable	Was it possible to test each requirement to determine whether it has been properly implemented?
Modifiable	Was a history of changes made to each requirement kept?
Traceable	Was each requirement linked back to its source of origination?

To estimate whether the requirements possessed each quality characteristic, a 7-point frequency Likert-type scale was integrated into the scale with the eight characteristics. The respondents had to rate how frequently the requirements consisted of the eight quality characteristics. The scale used was 1 – Never; 2 - Rarely, in less than 10%; 3 - Occasionally, in about 30%; 4 - Sometimes, in about 50%; 5 - Frequently, in about 70%; 6 - Usually, in about 90%; 7 - Every time.

To simplify the analysis, exploratory factor analysis was used as a data reduction technique to validate whether the eight quality elements could be summarised. The quality of the requirements was summarised separately and displayed for practitioners involved in the requirements activity versus practitioners who were not involved in the activity to see whether this impacted the output of the requirements engineering process. The cross-tabulation in Table 2 indicates a higher percentage quality of requirements when the practitioners were involved. This is the case for each of the activities. The results show a clear pattern of dependency between the quality of requirements and the way the requirements engineering process is executed.

Table 2: Activity involvement by quality of requirements

Requirements activities	Quality				
	Rating 1	Rating 3	Rating 3 to 5	Rating 5 to 7	Total
Planning *Involved* (N=43)	0%	0%	23%	77%	100%
Planning *Not involved* (N=12)	8%	8%	42%	42%	100%
Elicitation *Involved* (N=50)	0%	2%	24%	74%	100%
Elicitation *Not involved* (N=5)	20%	0%	60%	20%	100%
Analysis & Modelling *Involved* (N=54)	2%	2%	26%	70%	100%
Analysis & Modelling *Not involved* (N=1)	0%	0%	100%	0%	100%
Specification *Involved* (N=55)	2%	2%	27%	69%	100%
Specification *Not involved* (N=0)	n/a	n/a	n/a	n/a	n/a
Validation *Involved* (N=49)	2%	2%	18%	78%	100%
Validation *Not involved* (N=6)	0%	0%	100%	0%	100%

1 (Never); 2 (Rarely, less than 10%); 3 (Occasionally, in about 30%); 4 (Sometimes in about 50%); 5 (Frequently in about 70%); 6 (Usually, in about 90%); 7 (Every time)

A Mann-Whitney test was done in each group (involved and not involved) to validate that these two groups did not have the same median. From the test results it was concluded in the case of planning, elicitation and validation that there were statistically significant differences between the groups where there was involvement in the activity: for the planning activity ($U = 124$, $p = 0.005$); for the elicitation activity ($U = 38.5$, $p = 0.009$); for the validation activity ($U = 51$, $p = 0.008$). In the analysis and modelling case the sample available for the not involved group was very small and could be the reason why there was no statistically significant difference between the groups.

These tests confirm that the more the practitioner is involved in the requirements activities, the higher the quality of requirements.

Customer satisfaction: The respondents were asked to indicate from their personal perspective whether the business stakeholders and end-users were satisfied with the implemented solution and to indicate whether the users actually used the system. The results showed an average of 82.46% for the business stakeholders' satisfaction and 81.16% for end-users' satisfaction. These satisfaction levels are high and it could merely be the perception of the practitioners. Correlations were calculated to determine whether there was a relationship between customer satisfaction and usage of the solution. A negative relationship was identified between end-user satisfaction and usage of the solution, with a correlation of -0.288* significant at the 0.05 level.

The following reasons were provided by respondents as to why customers do not use the solution:

- The users do not understand how the technology supports their business processes.
- The users are still using the old solution.
- The users do not use training and user manuals.
- Users are waiting for more requirements to be implemented.
- Development is still in progress.
- Users are forced to use the solution but don't like it (imposed on the industry).

The perception of the practitioners is that customer satisfaction levels are high; however, the relationship with the usage of the solution by stakeholders suggests that there is more information that must be explored. In future research, this relationship needs to be explored directly with the stakeholders as previously suggested.

5. CONCLUSIONS

Real data has been provided regarding the requirements engineering process in industry as a reference point. The researchers would, however, like to identify some important findings:

- Requirements planning: Practitioners do not follow a formal planning activity to consider how the requirements activities should be approached prior to executing the activities.
- Practitioners' information-seeking behaviour: The problem-solving process of the practitioners depends on sources of information which are either personal experience or conversations; their information-seeking behaviour does not change over time.
- Tools and techniques used during activities: The tools and techniques selected by practitioners to be used during activities are based on what is known, and not the most appropriate combination of techniques.
- Requirements management tools: Many requirements management tools are available but these tools are not generally used by practitioners to present or manage requirements.
- The results confirmed that in the cases where the practitioners are involved in the activities, high quality requirements are delivered. This implies a more efficient requirements engineering process as the output of the process delivers high quality.

The paper creates a South African context industry description of how practitioners execute the requirements engineering process. It confirms what is known by practitioners and how they use the knowledge of requirements practice. This knowledge provides adequate data on requirements practice within South Africa for future research. It also includes very specific focus areas for practitioners and managers on how to improve the requirements engineering process without adoption of any new tools or methodologies. The focus areas emerging from the results are practical with small changes in practitioners' behaviour that could have a major impact on the results of the requirements engineering process.
The second part of the survey explored how practitioners behave during the requirements engineering process.
This entails a description of how they gather information about the problem during the requirements process, use the information and share their resulting information. By discovering these interaction patterns, communication can be improved and made more effective.

Follow-up reporting will discuss the results from the second part of the survey. The knowledge of practitioners and how they use the knowledge of requirements practice can be applied to focus future research efforts. The knowledge on the behaviour of practitioner could form the basis of cross-disciplinary research.

6. REFERENCES

[1] S. Robertson and J. Robertson: *Mastering the Requirements Process*, Addison-Wesley, Boston, Second Edition. 2010.

[2] R. Gonzales: "Developing the requirements discipline: Software vs. systems", *IEEE Software,* Vol. 22 No. 2, pp. 59-61, 2005.

[3] K. El Emam and N. H. Madhavji: "A field study of requirements engineering practices in information systems development", *Proceedings of the Second IEEE International Symposium on Requirements Engineering*, pp. 68-80, 1995.

[4] A. Aurum and C. Wohlin: "Requirements Engineering: Setting the Context ", in *Engineering and Managing Software Requirements*, A. Aurum and C. Wohlin, Eds., Berlin Heidelberg: Springer, 2005, pp. 1-15.

[5] A. van Lamsweerde: "Requirements engineering in the year 00: A research perspective", *Proceedings of International Conference on Software Engineering*, New York, pp. 5-19, 2000.

[6] B. Berenbach, D. J. Paulish, J. Kazmeier, and A. Rudorfer: *Software & Systems Requirements Engineering: In Practice*, McGraw Hill, New York. 2009.

[7] S. Viller, J. Bowers, and T. Rodden: "Human factors in requirements engineering:: A survey of human sciences literature relevant to the improvement of dependable systems development processes", *Interacting with Computers,* Vol. 11 No. 6, pp. 665-698, 1999.

[8] F. P. Brooks, Jr.: "No silver bullet; essence and accidents of software engineering", *Computer,* Vol. 20 No. Copyright 1987, IEE, pp. 10-19, 1987.

[9] C. Schwaber: "*The Root of the Problem: Poor Requirements*", Forrester Research, Cambridge, September 2006.

[10] M. I. Kamata and T. Tamai: "How does requirements quality relate to project success or failure?", *15th IEEE International Requirements Engineering Conference*, pp. 69-78, 2007.

[11] K. Ellis: "*Business Analysis Benchmark: The Impact of Business Requirements on the Success of Technology Projects*", IAG Consulting Delaware, 2008.

[12] K. El Emam and A. G. Koru: "A replicated survey of IT software project failures", *IEEE Software,* Vol. 25 No. 5, pp. 84-90, 2008.

[13] R. L. Glass: *Facts and Fallacies of Software Engineering* Addison-Welsey, Boston. 2003.

[14] A. Rodrigues: "*Project Goals, Business Performance, and Risk*", Cutter Consortium, Arlington, 2001.

[15] H. F. Hofmann and F. Lehner: "Requirements engineering as a success factor in software projects", *IEEE Software,* Vol. 18 No. 4, pp. 58-66, 2001.

[16] D. Leffingwell. (1997, 12 May 2011). Calculating your Return on Investment from More Effective Requirements Management. Available: http://www-128.ibm.com/developerworks/rational/library/347.html.

[17] A. L. Lederer and J. Prasad: "Nine management guidelines for better cost estimating", *Communications of the ACM,* Vol. 35 No. 2, pp. 51-59, 1992.

[18] D. J. Reifer: "Profiles of level 5 CMMI organizations", *Crosstalk: The Journal of Defense Software Engineering*, pp. 24-28.

[19] I. F. Hooks and K. A. Farry: *Customer-centered Products Creating Successful Products through Smart Requirements Management*, AMACOM/American Management Association, New York. 2001.

[20] B. W. Boehm: "Verifying and validating software requirements and design specifications", *IEEE Software,* Vol. 1 No. 1, pp. 75-88, 1984.

[21] B. W. Boehm: "Software Engineering", *IEEE Transactions on Computers,* Vol. C-25 No. 12, pp. 1226-1241, 1976.

[22] B. W. Boehm and P. N. Papaccio: "Understanding and controlling software costs", *IEEE Transactions on Software Engineering,* Vol. 14 No. 10, pp. 1462-1477 1988.

[23] C. Jones: "Software change management", *Computer,* Vol. 29 No. 2, pp. 80-82, 1996.

[24] B. H. C. Cheng and J. M. Atlee: "Current and Future Research Directions in Requirements Engineering", in *Design Requirements Engineering: A Ten-Year Perspective*. Vol. 14, K. Lyytinen*, et al.*, Eds., Heidelberg: Springer-Verlag, 2009, pp. 11-43.

[25] D. Zowghi and C. Coulin "Requirements Elicitation: A Survey of Techniques, Approaches and Tools", in *Engineering and Managing Software Requirements*, A. Aurum and C. Wohlin, Eds., Berlin Heidelberg: Springer, 2005, p. 19.

[26] S. Hansen, N. Berente, and K. Lyytinen: "Requirements in the 21st Century: Current Practice and Emerging Trends", in *Design Requirements Engineering: A Ten-Year Perspective*. Vol. 14, K. Lyytinen*, et al.*, Eds., Heidelberg: Springer-Verlag, 2009, pp. 44-87.

[27] A. Wever and N. Maiden: "What are the day-to-day factors that are preventing business analysts from effective business analysis?", *19th IEEE International Requirements Engineering Conference*, pp. 293-298, 2011.

[28] C. J. Neill and P. A. Laplante: "Requirements engineering: The state of the practice", *IEEE Software,* Vol. 20 No. 6, pp. 40-45, 2003.

[29] M. Ciolkowski, O. Laitenberger, and S. Biffl: "Software reviews, the state of the practice", *IEEE Software,* Vol. 20 No. 6, pp. 46-51, 2003.

[30] T. E. Bell and T. A. Thayer: "Software requirements: Are they really a problem?," presented at the Proceedings of the 2nd International Conference on Software Engineering, San Francisco, California, pp. 61-68, 1976.

[31] D. Leffingwell: *Agile Software Requirements Lean Requirements Practices for Teams, Programs, and the Enterprise*, Addison-Wesley Boston. 2011.

[32] B. Boehm: "A view of 20th and 21st century software engineering", *Proceedings of the 28th International Conference on Software Engineering*, New York, pp. 12-29, 2006.

[33] I. Sommerville: *Software Engineering* Pearson Education, Harlow, Seventh Edition. 2004.

[34] I. Sommerville: *Software Engineering*, Addison-Wesley, Harlow, Six Edition. 2001.

[35] M. Dorfman: "Requirements Engineering", in *Software Requirements Engineering*, M. Dorfman and R. H. Thayer, Eds., Los Alamitos, California: IEEE Computer Society Press, 2000, p. 531.

[36] B. W. Boehm: "A spiral model of software development and enhancement", *International Workshop on the Software Process and Software Environments,* Vol. 11 No., pp. 22-42, 27-29 March 1986.

[37] A. van Lamsweerde: *Requirements Engineering from Systems Goals to UML Models to Software Specification*, John Wiley & Sons, Chichester. 2009.

[38] B. Boehm: "Get ready for agile methods, with care", *Computer,* Vol. 35 No. 1, pp. 64-69, 2002.

[39] K. Pohl: "The three dimensions of requirements engineering: A framework and its applications", *Information Systems,* Vol. 19 No. 3, pp. 243-258, 1994.

[40] P. Loucopoulos and V. Karakostas: *System Requirements Engineering*, McGraw-Hill, New York. 1995.

[41] K. Pohl: "Requirement engineering: An overview", in *Encyclopedia of Computer Science and Technology*. Vol. 36, A. Kent and J. G. Williams, Eds., 1996.

[42] K. E. Wiegers: *Software Requirements: Practical Techniques for Gathering and Managing Requirements throughout the Product Development Cycle*, Microsoft Press, Second Edition. 2003.

[43] G. Kotonya and I. Sommerville: *Requirements Engineering Processes and Techniques*, John Wiley & Sons, Chichester. 1998.

[44] L. Macaulay: "Requirements for requirements engineering techniques", *Proceedings of the Second International Conference on Requirements Engineering*, Los Alamitos, pp. 157-64, 1996.

[45] B. Nuseibeh and S. Easterbrook: "Requirements engineering: A roadmap," presented at the Proceedings of the Conference on The Future of Software Engineering, Limerick, pp. 35-46, 2000.

[46] S. Robertson: "Requirements trawling: Techniques for discovering requirements", *International Journal of Human-Computer Studies,* Vol. 55 No. 4, pp. 405-421, 2001.

[47] N. A. M. Maiden and G. Rugg: "ACRE: Selecting methods for requirements acquisition", *Software Engineering Journal,* Vol. 11 No. 3, pp. 183-192, 1996.

[48] A. M. Hickey and A. M. Davis: "A unified model of requirements elicitation", *Journal of Management Information Systems,* Vol. 20 No. 4, pp. 65-84, 2004.

[49] P. Hsia, A. M. Davis, and D. C. Kung: "Status report: Requirements engineering", *IEEE Software,* Vol. 10 No., pp. 75-9, 1993.

[50] S. Robertson and J. Robertson: "*Volere Requirements Specification Template*", Atlantic Systems Guild Limited, 2010.

[51] IEEE, "IEEE Guide for Developing System Requirements Specifications," *IEEE Std 1233, 1998 Edition*, p. i 1998.

[52] IEEE, "IEEE Recommended Practice for Software Requirements Specifications," *IEEE Std 830-1998*, p. i 1998.

[53] IIBA: "*A Guide to the Business Analysis Body of Knowledge (BABOK Guide)*, Toronto Vol. 2. 2009.

[54] P. Berander and A. Andrews: "Requirements Prioritization", in *Engineering and Managing Software Requirements*, A. Aurum and C. Wohlin, Eds., Berlin Heidelberg: Springer, 2005.

[55] IEEE Computer Society: *Guide to the Software Engineering Body of Knowledge (SWEBOK)*, IEEE Computer Society, Los Alamitos. 2004.

[56] The International Council on Systems Engineering: *System Engineering Handbook*, The International Council on Systems Engineering, San Diego Vol. 3.1. 2007.

[57] F. J. Fowler: *Survey Research Methods*, Sage, Thousand Oaks, Fourth Edition. 2009.

[58] L. M. Rea and R. A. Parker: *Designing and Conducting Survey Research - A Comprehensive Guide*, Jossey-Bass, San Francisco. 2005.

[59] D. R. Cooper and P. S. Schindler: *Business Research Methods*, Mc-Graw Hill, Boston, Tenth Edition. 2008.

[60] T. C. Wagenaar: "Survey Design", in *Encyclopedia of Social Measurement* Vol. 3 P-Y, K. Kempf-Leonard, Ed., Oxford: Elsevier Academic Press, 2005, p. 715.

[61] K. Kelly, B. Clark, V. Brown, and J. Sitzia: "Good practice in the conduct and reporting of survey research", *International Journal for Quality in Health Care,* Vol. 15 No. 3, pp. 261-266, 1 May 2003.

[62] K. A. Rasinski: "Surveys", in *Encyclopedia of Social Measurement* Vol. 3 P-Y, K. Kempf-Leonard, Ed., Oxford: Elsevier Academic Press, 2005, p. 733.

[63] R. Czaja and J. Blair. (2005). Designing Surveys *(*Second ed.). Available: http://dx.doi.org/null

[64] R. Sapsford: *Survey Research*, Sage, Thousand Oaks. 2007.

[65] M. Marpsat and N. Razafindratsima: "Survey methods for hard-to-reach populations: Introduction to the special issue", *Methodological Innovations Online,* Vol. 5 No. 2, 2010.

[66] F. Baltar and I. Brunet: "Social research 2.0: Virtual snowball sampling method using Facebook", *Internet Research,* Vol. 22 No. 1, pp. 57-74, 2012.

[67] A. Bryman and E. Bell: *Business Research Methods*, Oxford University Press, Oxford, Third Edition. 2011.

[68] P. Biernacki and D. Waldorf: "Snowball sampling: Problems and techniques of chain referral sampling", *Sociological Methods & Research,* Vol. 10 No. 2, pp. 141-163, 1 November 1981.

[69] B. Paech, T. Koenig, L. Borner, and A. Aurum: "An Analysis of Empirical Requirements Engineering Survey Data", in *Engineering and Managing Software Requirements*, A. Aurum

and C. Wohlin, Eds., Berlin Heidelberg: Springer, 2005, pp. 436-487.

[70] F. Houdek and K. Pohl: "Analyzing requirements engineering processes: A case study", *Proceedings of the 11th International Workshop on Database and Expert Systems Applications*, pp. 983-987, 2000.

[71] C. Quinlan: *Business Research Methods*, South-Western Cengage Learning, Hampshire. 2011.

[72] J. E. W. Holm, M. van der Walt, and P. W. Stoker: "Systems Engineering Training in Engineering Education", *Transforming Engineering Education: Creating Interdisciplinary Skills for Complex Global Environments, IEEE Conference*, pp. 1-14, 2010.

[73] IREB. (2006, 9 June 2012). *International Requirements Engineering Board*. Available: http://www.certified-re.de/en/

[74] B. Carkenord, C. Cartwright, R. Grace, L. Goldsmith, E. Larson, J. Skrabak, and C. Vonk: "Partnering for Project Success Project Manager and Business Analyst Collaboration", *Proceedings of the PMI Global Congress* Washington, DC, 2010.

[75] S. Weese and T. Wagner: *CBAP/CCBA Certified Business Analysis Study Guide*, Wiley. 2011.

[76] C. Tenopir and D. W. King: *Communication Patterns of Engineers*, Wiley, Hoboken. 2004.

[77] D. Veshosky: "Managing innovation information in engineering and construction firms", *Journal of Management in Engineering,* Vol. 14 No. 1, pp. 58-66, 1998.

[78] L. Upchurch, G. Rugg, and B. Kitchenham: "Using card sorts to elicit web page quality attributes", *IEEE Software,* Vol. 18 No. 4, pp. 84-89, 2001.

[79] A. M. Hickey and A. M. Davis: "Requirements elicitation and elicitation technique selection: Model for two knowledge-intensive software development processes", *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, 2003*, p. 10

[80] T. NATURE: "Defining visions in context: Models, processes and tools for requirements engineering", *Information Systems,* Vol. 21 No. 6, pp. 515-547, 1996.

[81] C. Denger and T. Olsson: "Quality Assurance in Requirements Engineering", in *Engineering and Managing Software Requirements*, A. Aurum and C. Wohlin, Eds., Berlin Heidelberg: Springer, 2005, pp. 178-200.

ROTATION HOUGH TRANSFORM

S. Du*

* *Dept. of Electrical and Mining Engineering, University of South Africa, Florida 1710 South Africa E-mail:dushengzhi@gmail.com*

Abstract: This article discusses how image space discretization errors affect the standard Hough transform (SHT). It is observed that the discretization errors are different on various columns, that is, they are related to the θ values. Furthermore, the columns corresponding to θ=0° and θ=90° are seen to have the smallest errors. Based on this observation, a new Hough transform (rotation Hough transform, RoHT) is proposed. The given image is rotated by Δθ and the SHT is employed, generating the HT data on the columns corresponding to θ=0° and θ=90° only. After the image is rotated 90°/Δθ times, the HT columns generated are registered to the HT matrix of the original image. In this way, only θ=0° and θ=90° are used to calculate the votes for each cell in HT space, and the errors are minimized. The experiments verified the performance of the proposed method.

Keywords: Hough transform, rounding errors, HT voting process, straight line detection.

1. INTRODUCTION

The Hough transform (HT) proposed by Hough C.V., 1962 [1] is a commonly used technique for detecting straight lines in digital images. Each feature point in the image space is mapped to a sinusoidal curve in the parameter space (also commonly called HT space) by means of the following equation:

$$\rho = x\cos\theta + y\sin\theta \qquad (1)$$

where (x, y) is a feature point in the image space, and (θ, ρ) are the normal parameters of a straight line. According to the predetermined resolutions on θ and ρ directions (denoted as $\Delta\theta$ and $\Delta\rho$ respectively), the HT space is quantized into a matrix of cells. Along the sinusoidal curve mapped from feature point (x, y), the votes for the cells are calculated from the feature point. Once all feature points are considered, votes for some cells in the HT space are calculated. The cell to receive the most votes is termed the peak, and represents the dominant straight line in the image. The discretization errors in both the image space and the parameter space affect the HT data. The voting process also includes rounding errors because the ρ value calculated from eq. (1) might be unequal to the ρ value of any cell and be rounded to the nearest one. All these errors definitely affect the HT performance, such as the peak height and detection errors.

The HT errors have been extensively discussed and various improvements proposed [2-5]. The symmetry [6] and self-similarity [7] were discovered and used to reduce the error. A method of HT error compensation by shifting the image is proposed in [8].

This article discusses how the image space discretization errors affect HT performance. A new method is proposed that entails rotating the given image and employing the standard Hough transform (SHT) generating the columns of $\theta = 0^{\circ}$ and $\theta = 90^{\circ}$ only. These columns are then registered to the HT space of the original image, and thus generate the full HT matrix. In this way, the effects of discretization errors in the image space are minimized.

The rest of the article is organized as follows: Section 2 contains observations on the effects of discretization errors and rounding errors on HT performance, and sets out a proposed new method for minimizing these effects; Section 3 discusses some experiments for evaluating the proposed method through comparison with the SHT; and Section 4 concludes the article.

2. OBSERVATIONS AND THE PROPOSED METHOD

2.1 Digital images do not really have straight lines except $\theta = 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$

In most digital images the pixels are square. Consequently, the straight lines with $\theta \neq 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$ are represented approximately by connected short segments with $\theta = 0^{\circ}, \pm 45^{\circ}$, and/or $\pm 90^{\circ}$ as shown in Fig.1.

This means that only the straight lines with $\theta = 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$ can be represented exactly. In another words, these straight lines have the lowest discretization errors. This characteristic gives rise to an interesting situation when the SHT is employed to detect straight lines in digital images. As shown in Fig. 2, the HT data (the resolutions in θ and ρ directions are $\Delta\theta = 1^{\circ}$ and $\Delta\rho = 1$ pixel respectively) of a digital image covers the range $\theta = -90^{\circ}$ to $\theta = 89^{\circ}$. Zooming in on the HT data around $\theta = 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$, may reveal these columns to be different from others. The contrast of the cells in these columns is higher than in surrounding columns. Considering the nature of the digital images mentioned above, these columns have a higher signal/noise ratio than others. This discovery prompted me to improve the HT performance by making full use of these columns.

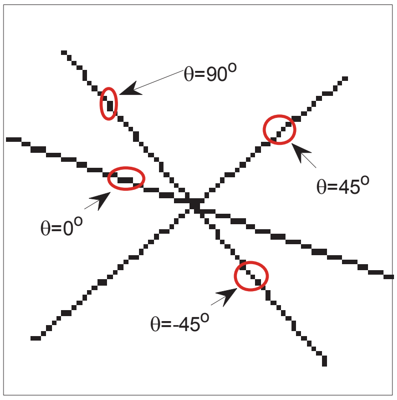

Fig. 1. Straight lines in digital images are actually composed of short segments with $\theta = 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$

2.2 Rotation HT

2.2.1 Rotating images

As discussed above, the straight lines with $\theta = 0^{\circ}, \pm 45^{\circ}, \pm 90^{\circ}$ have the lowest discretization errors and their corresponding columns in HT space have the highest signal/noise ratio. If an image is imagined as rotating by $\Delta\theta$ (the resolution in θ direction), then all straight lines with $\theta = \Delta\theta$ will become horizontal and the ones with $\theta = 90^{\circ} + \Delta\theta$ will become vertical. If the image is rotated again, then some other straight lines become either horizontal or vertical. Once the image has been rotated a sufficient number of times, all potential straight lines will have the opportunity to be approximately horizontal or vertical.

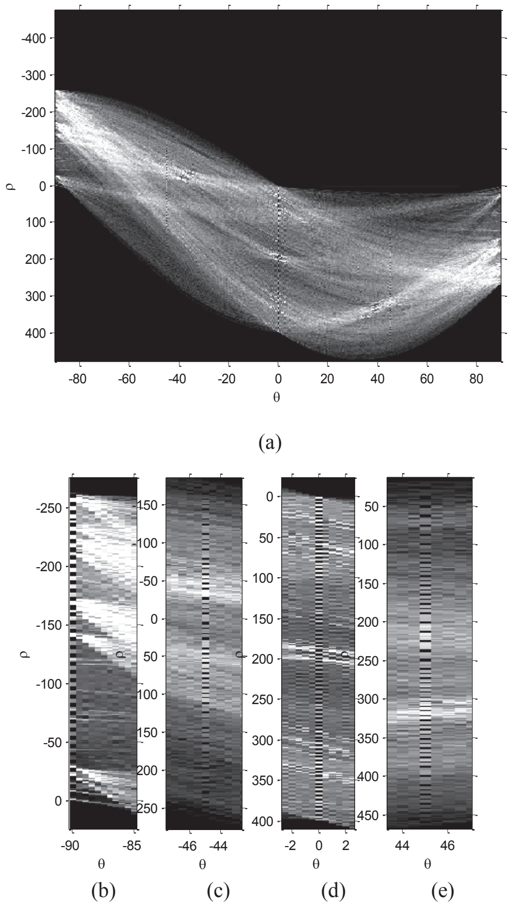

Fig. 2. (a) Global Hough data (b) Local HT data around $\theta = -90^{\circ}$ (c) Local HT data around $\theta = -45^{\circ}$ (d) Local HT data around $\theta = 0^{\circ}$ (e) Local HT data around $\theta = 45^{\circ}$

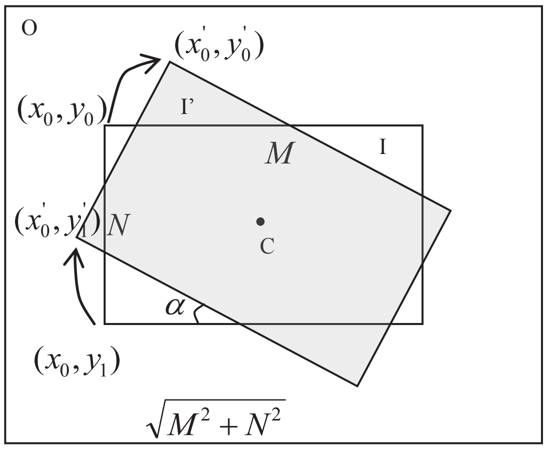

Fig. 3. An image is extended for rotation

Fig. 3 shows an image rotated around its center. For an $M \times N$ image (M, N being the width and height respectively), to avoid the feature point being rotated out of the rectangle area, the image is extended to a larger square area as shown in Fig. 3. The width of the square is $\sqrt{M^2 + N^2}$. The upper left-hand corner (O in Fig. 3) of the square area is the origin of the extended image.

2.2.2 Registering the HT data of the rotated image to the HT matrix of the original image

If the SHT is applied to each rotated image to generate the columns of $\theta = 0^o, \pm 45^o, \pm 90^o$ only, we obtain HT columns of highest signal/noise level.

Ideally, the peak height in the HT space of a straight line should be equal to its geometric length. However, the geometric length l of a straight line with $\theta = \pm 45^o$ is not equal to the number of its pixels (denoted as n). In fact

$$l = \sqrt{2}n. \tag{2}$$

To obtain a higher peak, only the columns of $\theta = 0^o, \pm 90^o$ are used. In fact, $\theta = \pm 90^o$ represents the same angle. Hence $\theta = 0^o, -90^o$ are used in the proposed method.

When a single column in HT data is considered, all cells have the same θ value, but discretized ρ values. Each cell corresponds to a belt in the image space with angle θ and width $\Delta\rho$. In a column, if we consider the top cell and the bottom cell, this is equivalent to scanning the image using a belt. In this article only the columns of 0^o and 90^o are considered, which means that only vertical and horizontal scans are used for each rotated image. From Fig. 3 it can be seen that the scans on the original image I with angle α and $\alpha + 90^o$ are equivalent to the horizontal and the vertical scans on the rotated image I' respectively. So, the columns of $\theta = \alpha$ and $\alpha + 90^o$ in the HT data of I correspond to the columns of $\theta = 0^o, -90^o$ of I'. It should be noted that the ρ value of the corner (x_0, y_1) changes after it is rotated to $(x_0^{'}, y_1^{'})$. From simple geometry analysis, one obtains

$$\rho = x_0 \cos(90^o - \alpha) + y_1 \sin(90^o - \alpha) \tag{3}$$

$$\rho' = x_0'. \tag{4}$$

Therefore one can simply copy specific cells generated by the horizontal scan on the rotated image to the $\theta = \alpha$ column in the HT data of I. The cells starting at $\rho' = x_0'$ are copied to the $\theta = \alpha$ column. The destination cells begin from $\rho = x_0 \cos(90^o - \alpha) + y_1 \sin(90^o - \alpha)$. For the vertical scan, similar results are obtained.

After the image is rotated $90^o / \Delta\theta$ times, the complete HT matrix of the original image will be filled.

3. EXPERIMENTS

To evaluate the proposed method, an image containing a straight line as shown in Fig. 4 is considered. The geometric length of the line is 80 pixels, and the normal parameters are θ and ρ respectively. To demonstrate the effects of θ values on the image space discretization error and then the HT performance, the θ values are changed from 1^o to 89^o but the length is kept. The SHT and the proposed method are then employed to detect the straight line.

Fig. 5 shows that the peaks obtained by the proposed rotation HT (RoHT) are closer to the "true" value than the SHT. Fig. 6 and Fig. 7 depict the comparison of detection errors. The proposed method may result in smaller errors than the SHT, which means that the proposed method detects straight lines with higher accuracy.

From Figs. 5, 6, and 7 one may find that the proposed method outperforms the SHT in terms of higher peaks and lower detection errors.

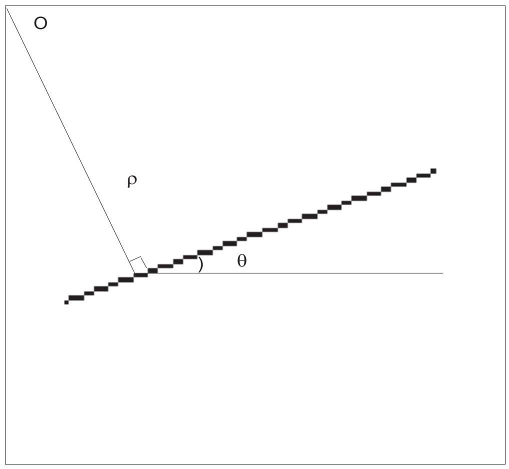

Fig. 4. Segment for experiments

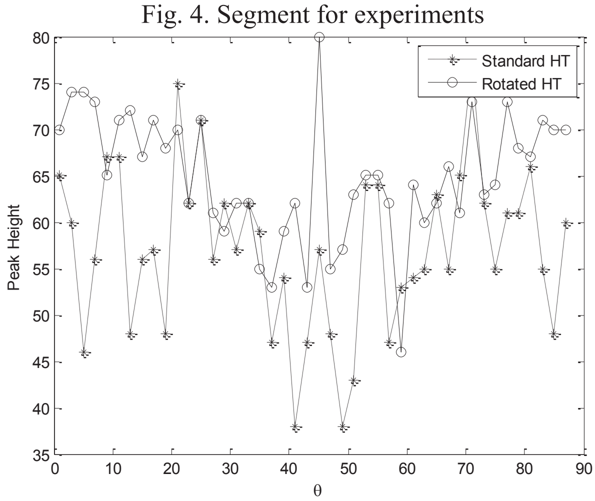

Fig. 5. Comparison of the peak height when the SHT and the proposed method are employed to detect straight lines

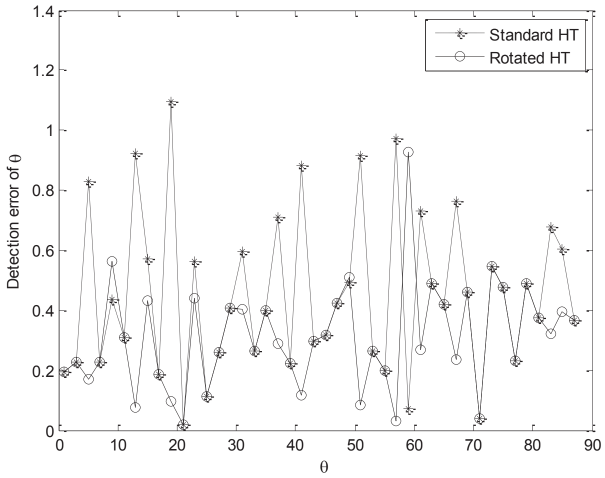

Fig. 6. Comparison of the θ detection error when the SHT and the proposed method are employed to detect straight lines

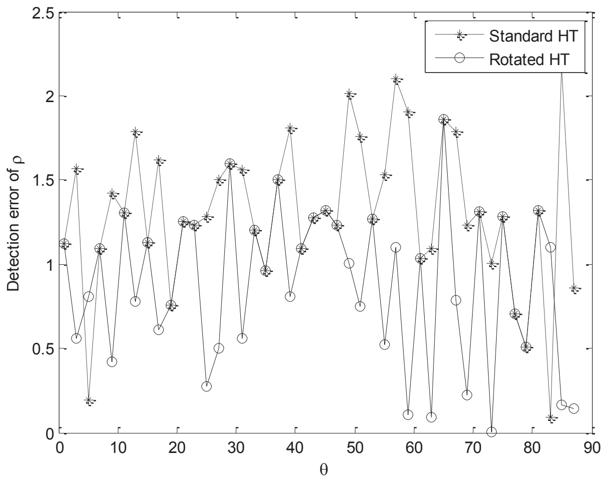

Fig. 7. Comparison of the ρ detection error when the SHT and the proposed method are employed to detect straight lines

4. CONCLUSION

This article analyzed the image space discretization error and its effects on the HT performance. It was discovered that these effects were minimized in the columns of $\theta = 0^{\text{o}}$ and $\theta = 90^{\text{o}}$, and on this basis the RoHT was proposed, which entailed rotating the given image and employing the SHT to generate these two columns only. The experiments showed the proposed method to be effective.

The computation complexity of rotating the given image is the main weakness of the proposed method, and this will constitute the focus of future work.

5. REFERENCES

[1] Hough PVC. A method and means for recognizing complex patterns. US Patent 3,069,654, 1962.

[2] Van Veen TM and Groen FCA. Discretization Errors in the Hough Transform. Pattern Recognition. Vol. 14, pp. 137–145, 1981.

[3] Yuen SY and Ma CH. An investigation of the nature of parameterization for the Hough transform. Pattern Recognition. Vol. 30, pp. 1009–1040, 1997.

[4] Guo LF and Chutatape O. Influence of discretization in image space on Hough transform. Pattern Recognition. Vol. 32, pp. 645–644, 1999.

[5] Zhang M. In the discretization of parameter domain in Hough transform, Proceedings of IEEE (ICPR '96), pp. 527–482, 1996.

[6] Du S, Tu C, and Sun M. High accuracy Hough transform based on butterfly symmetry. Electronics Letters. Vol. 48, No. 4, pp. 199–201, 2012.

[7] Tu C, Du S, van Wyk B, Djouani K, and Hamam Y. High resolution Hough transform based on butterfly self-similarity. Electronics Letters, Vol. 47, No. 25, pp. 1360–1361, 2011.

[8] Tu C, van Wyk B, Djouani K, Hamam Y, and Du S. Line positioning based on HT error compensation. IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1–4, 2011.

SAIEE AFRICA RESEARCH JOURNAL – NOTES FOR AUTHORS

This journal publishes research, survey and expository contributions in the field of electrical, electronics, computer, information and communications engineering. Articles may be of a theoretical or applied nature, must be novel and must not have been published elsewhere.

Nature of Articles

Two types of articles may be submitted:

- Papers: Presentation of significant research and development and/or novel applications in electrical, electronic, computer, information or communications engineering.
- Research and Development Notes: Brief technical contributions, technical comments on published papers or on electrical engineering topics.

All contributions are reviewed with the aid of appropriate reviewers. A slightly simplified review procedure is used in the case of Research and Development Notes, to minimize publication delays. No maximum length for a paper is prescribed. However, authors should keep in mind that a significant factor in the review of the manuscript will be its length relative to its content and clarity of writing. Membership of the SAIEE is not required.

Process for initial submission of manuscript

Preferred submission is by e-mail in electronic MS Word and PDF formats. PDF format files should be 'press optimised' and include all embedded fonts, diagrams etc. All diagrams to be in black and white (not colour). For printed submissions contact the Managing Editor. Submissions should be made to:

The Managing Editor, SAIEE Africa Research Journal,
PO Box 751253, Gardenview 2047, South Africa.
E-mail: researchjournal@saiee.org.za

These submissions will be used in the review process. Receipt will be acknowledged by the Editor-in-Chief and subsequently by the assigned Specialist Editor, who will further handle the paper and all correspondence pertaining to it. Once accepted for publication, you will be notified of acceptance and of any alterations necessary. You will then be requested to prepare and submit the final script. The initial paper should be structured as follows:

- TITLE in capitals, not underlined.
- Author name(s): First name(s) or initials, surname (without academic title or preposition 'by')
- Abstract, in single spacing, not exceeding 20 lines.
- List of references (references to published literature should be cited in the text using Arabic numerals in square brackets and arranged in numerical order in the List of References).
- Author(s) affiliation and postal address(es), and email address(es).
- Footnotes, if unavoidable, should be typed in single spacing.
- Authors must refer to the website: http: //www.saiee.org.za/arj where detailed guidelines, including templates, are provided.

Format of the final manuscript

The final manuscript will be produced in a 'direct to plate' process. The assigned Specialist Editor will provide you with instructions for preparation of the final manuscript and required format, to be submitted directly to:
The Managing Editor, SAIEE Africa Research Journal, PO Box 751253, Gardenview 2047, South Africa.
E-mail: researchjournal@saiee.org.za

Page charges

A page charge of R200 per page will be charged to offset some of the expenses incurred in publishing the work. Detailed instructions will be sent to you once your manuscript has been accepted for publication.

Additional copies

An additional copy of the issue in which articles appear, will be provided free of charge to authors.
If the page charge is honoured the authors will also receive 10 free reprints without covers.

Copyright

South African Institute for Electrical Engineers (SAIEE)
PO Box 751253, Gardenview, 2047, South Africa
Tel: 27 11 487 3003 | Fax: 27 11 487 3002
E-mail: researchjournal@saiee.org.za | Website: www.saiee.org.za