Evidence Based Model Calibration for Reliable Predictions

T.D. Wiles *Mine Modelling (Pty) Ltd, Australia*

Abstract

This paper discusses how to quantify the reliability of numerical model predictions, ways to improve the reliability, and finally how to apply this knowledge. By comparing model results with actual mine response we can quantify the reliability of numerical model predictions. This reliability can be improved by refining our model. This may require an improved representation of the geometry, better estimates of the pre-mining stress state, fine-tuning of the material properties and implementation of more complex plastic constitutive response. We can readily assess how important each of these refinements is simply by seeing how much improvement in reliability each of these refinements provides. This allows us to determine what type of model we require and assess whether the increased cost of a more complex model can be justified. Once we have quantified the reliability, we can use probability theory to make predictions of future mine behaviour. The use of probabilistic prediction allows for direct cost benefit comparisons of various alternates to be more readily considered.

1   Introduction

The reason we run numerical models is to try to determine why instabilities have happened in the past and to predict when these will occur in the future so that we can either avoid these problems or plan around them.

In one extreme, the predictions we make could be 100% reliable and our modelling exercise would be a fantastic success. Unfortunately, in the other extreme our predictions could be a dismal failure and our modelling exercise would not only be misleading but could compromise safety. Inevitably the reliability of our model predictions will lie somewhere between these two extremes being neither 100% reliable nor 0% reliable. We need to know the reliability of our predictions if they are to have any value.

Traditionally, the reliability has been assessed based on anecdotal evidence, with qualitative comparisons and trust in the opinions of the model operator make predictions more of a black art than a science. In order to reduce costs this too often allows overly simplified models to be used in situations where more complex models are required. But for the same reasons very expensive models are sometimes used where simplified models would provide adequate reliability for vastly reduced costs. Expensive lab testing, in situ monitoring and stress measurements are often made when they are not required, but neglected when necessary for model calibration.

There is a perception in the industry that lab testing and stress measurements constitute model calibration, that more complex constitutive models will naturally lead to better predictions regardless of how poorly calibrated the resulting model is. Too often a great deal of time, effort and money is wasted in addressing the wrong source of error. This leads to a situation where the true problems are not solved simply because the allocated budget is exhausted before they can be properly addressed.

This state of affairs exists and will continue to persist because without a quantitative measure of reliability there is no way to clearly state whether one prediction is in fact any better than another. This does not allow one to determine when the cost and effort of additional engineering is required and justified.

The tried and true scientific method requires that modelling should be approached as an evidence based objective exercise. Comparison with actual observed ground behaviour is the only way to confirm the efficacy of any model. Using this method, reliability can be readily quantified into a measure that allows direct comparison of the effectiveness of alternative models, as well as to make predictions with quantitative reliability. This allows us to determine whether the added expense of additional lab testing, monitoring, stress measurements and complex plasticity modelling actually quantitatively improve the reliability of our predictions or not, and hence whether these extra costs can be justified.

The thesis of this paper is that when we use modelling predictions we need to quantify the reliability so we will know to depend on them when they are good and disregard them when they are not.

2 Quantifying the reliability of model predictions

If we are to quantify the reliability of model predictions then we need to be very specific about exactly which parameters we are going to quantify. Fortunately this task is relatively easy since models predict very few parameters which include only stress, strain and displacement. Models do not directly predict ground stability problems such as pillar failure, over-break or ground support requirements. Expected mine behaviour can only be indirectly indicated by correlation with levels of stress or strain given by the model. By measuring the quality of this correlation we can quantify the reliability.

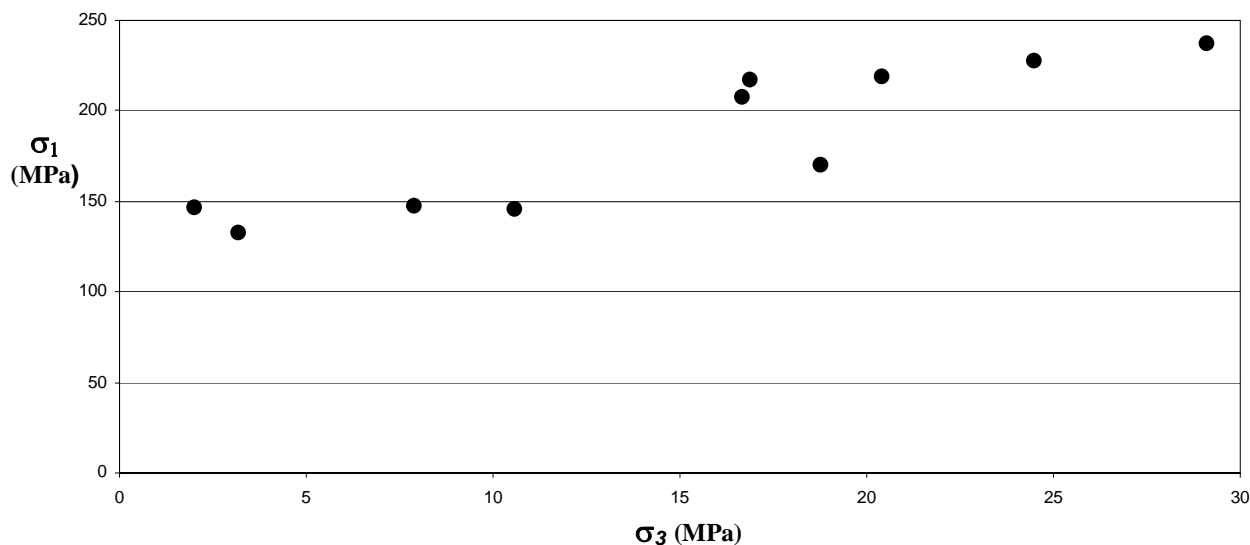### 2.1 Back-analysis of pillar failures

Let's now consider some back-analysis of pillar failures to illustrate this concept (Wiles, 2005).
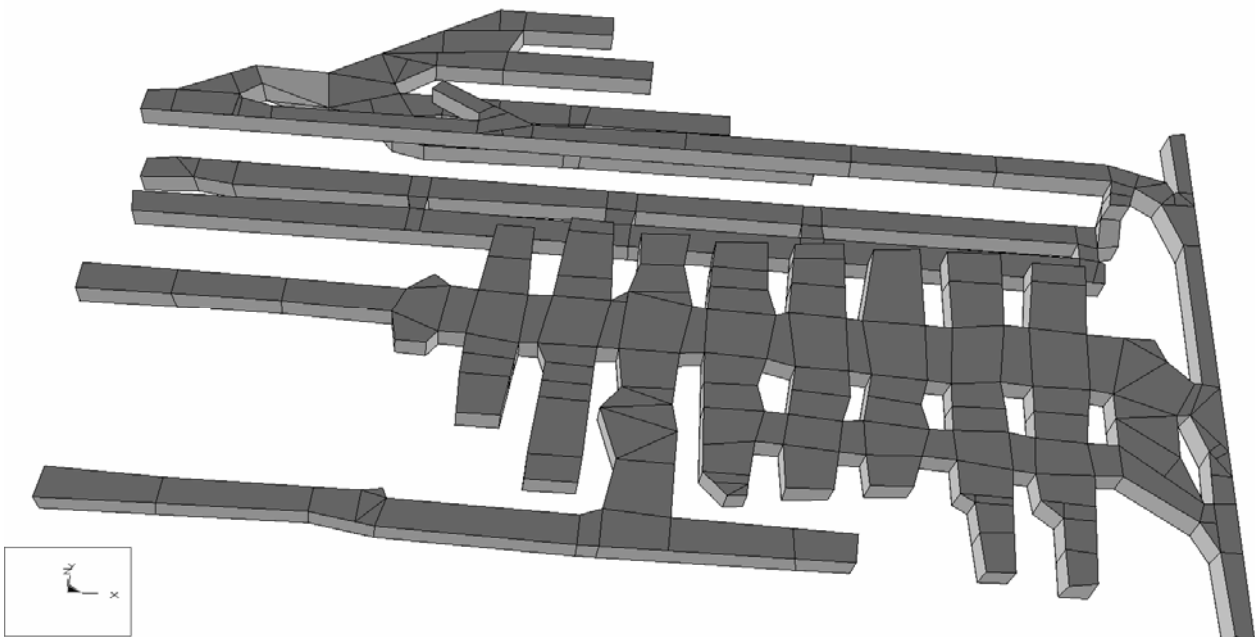
Figure 1   Pillar failures at Creighton mine

In Figure 1, the mining excavations are shown as solids and the pillars that failed are located between the cross-cuts. This is one of three different locations about 2 km deep in the mine where back-analyses were conducted to obtain the results shown in Figure 2. The maximum in situ stress of 100 MPa is oriented approximately along East-West (x-direction) and nearly horizontal. The σ_{c50} (50 mm intact unconfined core strength) is around 235 MPa.

In Figure 2, the elastic stress at the time failure for a series of historical pillar failures has been back-calculated using Map3D Fault-Slip (Wiles, 2007). Here, each point represents the principal stress determined for a different pillar from a separate back-analysis. A standard deviation of 18 MPa is readily calculated around the best fit line using linear regression. This in actual fact quantifies the reliability of this model into a single parameter that measures how well the model is able to match the observed mine behaviour. Equivalently, this could be expressed as a coefficient of variation (i.e. normalised standard deviation) simply by dividing by the mean stress giving a value of approximately ±10%.

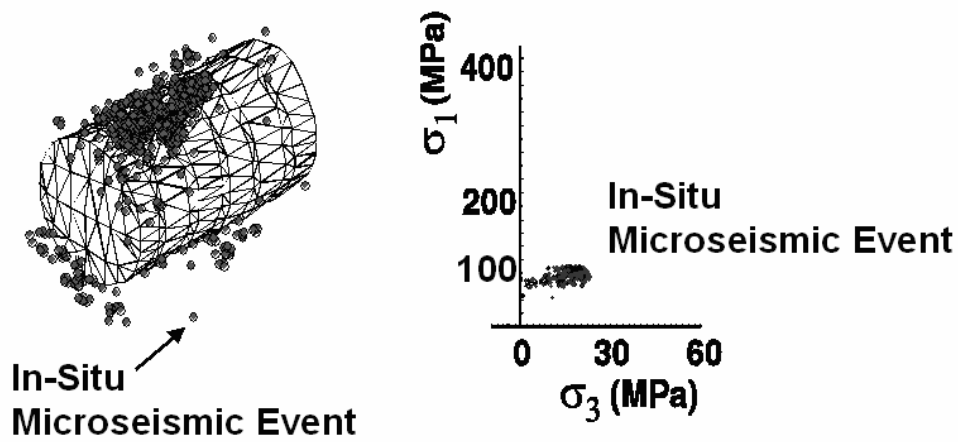**Figure 2   Elastic back-analysis of pillar failures at Creighton mine**

Here we can see a clear trend with low scatter which indicates that the model is able to predict pillars failures in the past quite accurately. Large scatter would indicate a very low reliability, conversely small scatter would indicate very high reliability. We anticipate that in future, the same model will be able to continue to predict pillar failures at the same location in the mine, with the same level of reliability.

Similar logic can be applied to any repeatedly observable mine behaviour such as ground support type, over-break depth, blast-hole collapse, stand-up time, etc. The limitation here is not in what or how this can be applied but rather in having a consistent series of observations available for back-analysis. This kind of data set is required so that the reliability can be quantified.

2.2   Back-analysis of seismic events

Now consider another example where stresses are calculated at seismic event locations (Martin, 1997) at AECL's Underground Research Laboratory. The mine-by tunnel illustrated in Figure 3 is 3.5 m diameter and is located at 420 m depth. The maximum in situ stress of approximately 60 MPa is oriented diametrically and nearly horizontal. The σ_{c50} (50 mm intact unconfined core strength) is around 210 MPa.

AECL's Mine-by Experiment

Figure 3   Back-analysis of seismic events at AECL's underground research laboratory

In Figure 3, the elastic stress at the location of seismic events has been back-calculated for several face advances. Here, each point represents the principal stress determined for a different seismic event. A standard deviation of 5 MPa (coefficient of variation of ±8%) is readily calculated around the best fit line using linear regression. We can see a clear trend with low scatter which indicates that the model is able to predict seismic events in the past quite accurately.

2.3   Back-analysis of structurally controlled over-break

In this example stresses are calculated at the over-break depth as illustrated in Figure 4 (Jamieson, 2007).
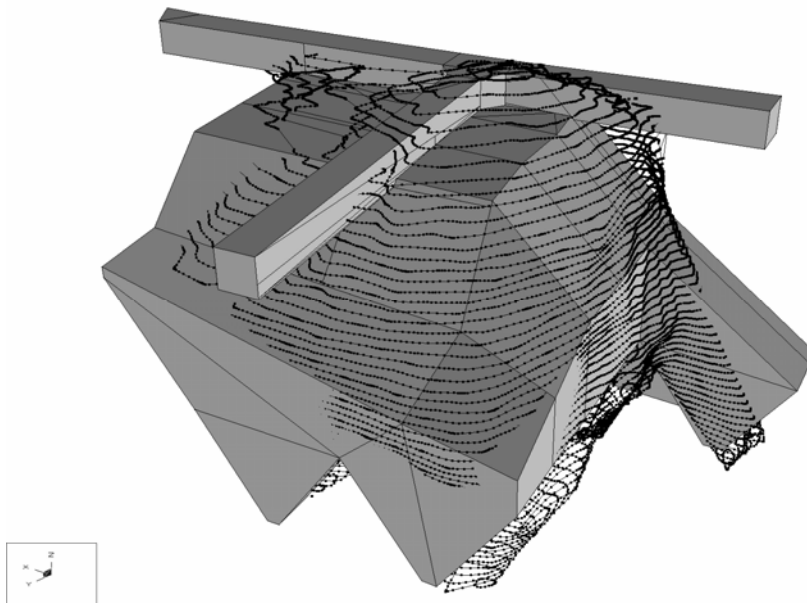
Figure 4   Stope over-break at Xstrata Copper, Mount Isa Copper Operations

This stope is located at 900 m depth. The maximum in situ stress of approximately 40 MPa is oriented nearly East-West (x-direction is east) dipping slightly to the East. The σ_{c50} (50 mm intact unconfined core strength) is around 150 MPa.
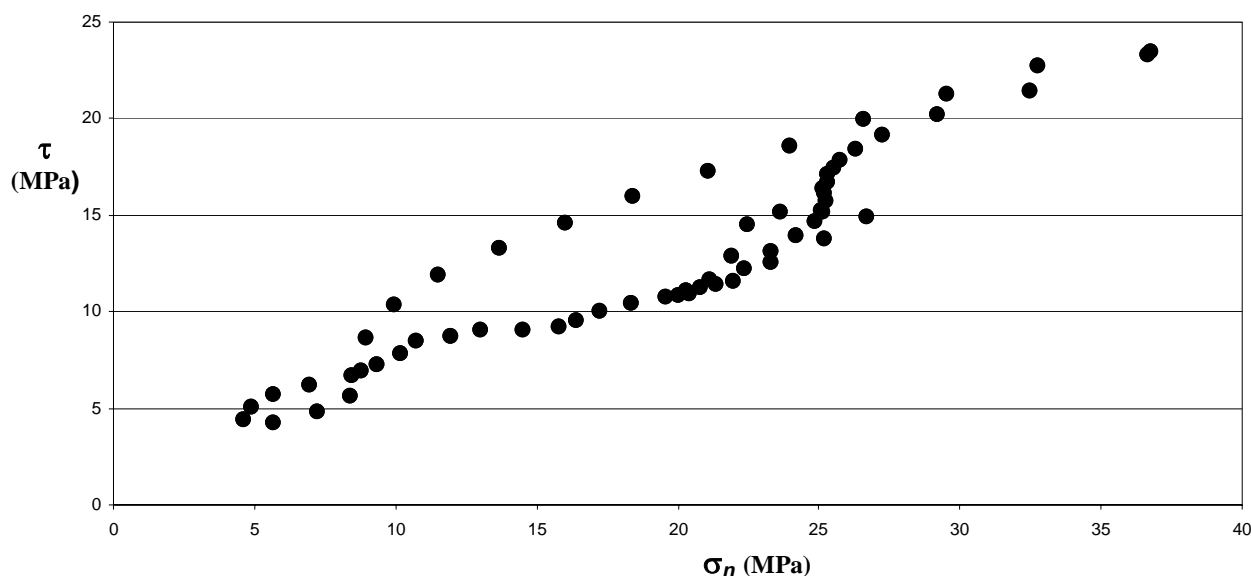
Figure 5   **Back-analysis of stope over-break at Xstrata Copper, Mount Isa Copper Operations**

In Figure 5, the elastic stress at points along the CMS over-break contour has been back-calculated using Map3D Fault-Slip (Wiles, 2007). Here, the stresses acting at the orientation of the local bedding plane (dip of 60° dip direction of 270°) are plotted. A standard deviation is 1.85 MPa (coefficient of variation of ±12%) is readily calculated around the best fit line using linear regression. In this case, we can see a clear trend with low scatter which indicates that the model is able to predict over-break in the past quite accurately.
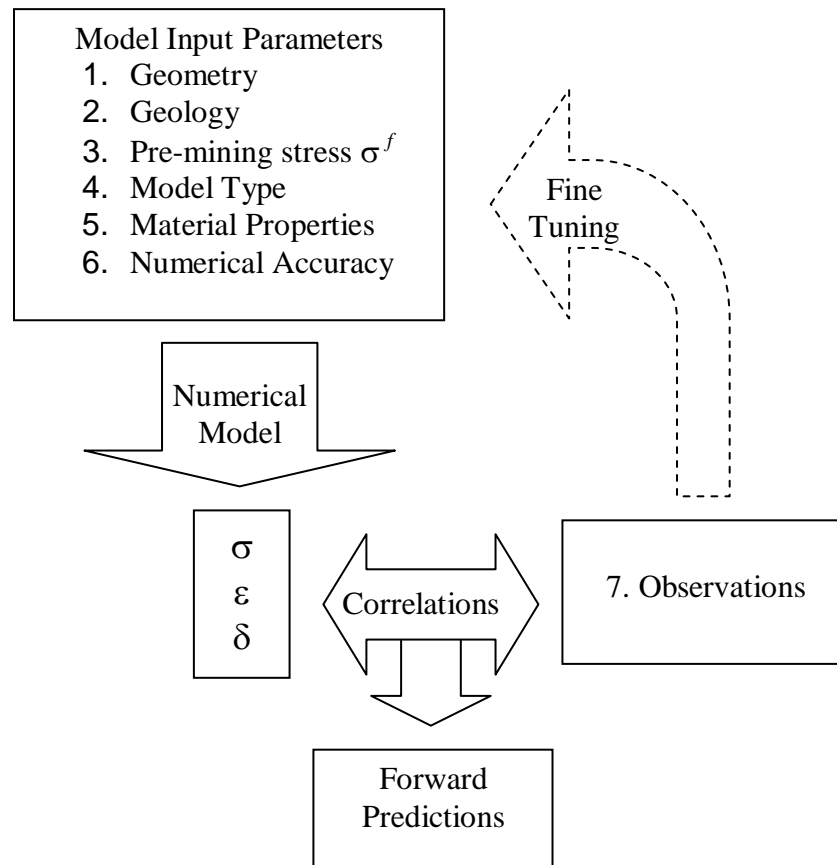
2.4   Discussion of back-analysis results

In each of these examples we have quantified the reliability of the model predictions by considering how well the model was able to match observed mine behaviour in the past. We anticipate that the same model will be able to continue to predict with the same level of reliability in future.

Since our modelling objective is to be able to make predictions with as much accuracy as possible, we are now in a position to re-run the same series of back-analyses with other models to see if they can improve the reliability (i.e. reduce the standard deviation) and by how much. Here we could consider various simplifications or refinements to the geometry. We could trial a series of alternative pre-mining stress state assumptions or we could incorporate additional geology in the form of faults or alternate material zones. We could re-run the same geometry with more complex constitutive behaviour with yielding faults or pillars. The model that gives the most reliability is the best choice since it best satisfies our modelling objective.

For each alternative model we can readily re-plot the back-analysis results as in the above figures then re-calculate the standard deviation to see if in fact we obtained any real improvement in the model reliability. We can of course also determine how much improvement was obtained and hence justify the cost of implementing the various alternatives we have trialled. Before considering this in detail let's first more clearly define the factors that affect the accuracy of model predictions.

3   Factors affecting the accuracy of model predictions

Since a numerical model is a mathematical calculation, the only things that can affect the results are the input parameters and the equations that are solved. A full list of model inputs is given in Figure 6.

Model Input Parameters
1. Geometry
2. Geology
3. Pre-mining stress σ^f
4. Model Type
5. Material Properties
6. Numerical Accuracy

Fine
Tuning

Numerical
Model

σ
ε
δ

Correlations

7. Observations

Forward
Predictions

Figure 6   Model inputs

The fundamental inputs which must be specified for a model are the geometry (2D, 3D, excavation shapes), geology (lithology), pre-mining stresses σ^f, model type (elastic, plastic), material properties and numerical accuracy (discretisation). From this information models calculate the stresses, strains and displacements. These can then be compared with observations (as in Figures 2, 3 and 5) to quantitatively determine the reliability.

3.1   Geometry

It is the geometry that controls how stresses are concentrated in the model. Narrow pillars will naturally tend to concentrate stresses while large span stopes will naturally tend to shadow nearby openings. It is therefore necessary that we construct our geometric approximation of the actual mine excavations to a consistent degree of accuracy. The most important consideration here is to get the pillar width/stope widths and position right. If this factor contributes significant errors to the model predictions the model will need to be re-built with more attention to detail.

Since analysis effort increases rather dramatically with the number of elements, there is always an overwhelming desire to use fewer elements and hence simplify the geometry. This is of course acceptable as long as one does not compromise the calculation accuracy.

3.2   Geology

We need to decide what components of the geology are to be included in the model. Stiff lithological units will tend to attract stress because of their increased resistance to deformation. Soft units trend to dissipate shear stresses. If these factors cause significant stress redistribution, such zones will need to be incorporated into the model. Unfortunately, each unit we add to our model adds additional unknowns that need to be calibrated.

3.3 Pre-mining stresses σ^f

Fundamentally modelling takes the pre-mining stress state and re-distributes this, hence prediction accuracy is directly dependant on what values are specified. It is therefore necessary that you have reasonable knowledge of the actual stress state in your mine. Pre-mining stresses can be determined through direct measurements such as AE or over-coring (Wiles et al., 2004). This can be supplemented with observations of mine response. Highly stressed corners can indicate the orientation of the maximum horizontal stress. High quality seismic monitoring data can in principal be used to determine both the stress ratios and the orientation. Back-analysis of instrumented mine response can also be used to determine both the stress ratios and the orientation (Wiles and Kaiser, 1994).

3.4 Model type

There are many model types to choose from including 2D, 3D, pseudo 3D tabular, elastic, plastic, creep, heat-flow, dynamic etc. There are also many modelling methods to choose from including Finite Element, Finite Difference and Boundary Element methods. Although it seems like there are an overwhelming number of options, these can be quickly sorted into a few categories with obvious advantages and disadvantages.

First of all, unless there are some unusual regularities in your mining geometry, 3D modelling will be required since 2D approximations will simply concentrate stresses incorrectly. Since all modelling methods give the same results for the same equations (i.e. elastic or plastic), it is the equations that are solved not the modelling method that matters. Clearly boundary element modelling has the distinct advantage of being relatively easy to set up and has short run times.

3D elastic modelling (Map3D, Examine3D, FLAC3D, 3DEC, ABACUS, ELFIN, etc.) allows one to accommodate complex 3D shapes thus overcoming the main problems with equilibrium and 2D modelling. Since elastic models solve the equations of equilibrium, continuity and elasticity only, very few input parameters need to be specified. The equations of elasticity require specification of the pre-mining stress state (σ^f), Young's modulus and Poisson's ratio only. Strength parameters are not used in these equations and hence do not affect the model predictions in any way. Unless you include multiple zones with contrasting stiffness, Young's Modulus has no effect on the stress predictions whatsoever and Poisson's ratio has only a very minor affect and this can generally be ignored. In most cases, there will only be 3 or 4 significant unknowns arising from σ^f, the maximum and minimum horizontal stresses and their orientation.

This small number of parameters presents a major advantage for elastic modelling since uncertainty in parameters is the primary cause of poor model predictions. With effectively only 3 or 4 parameters (σ^f) to calibrate and with relatively short run times, the uncertainty in these parameters can be quickly quantified and minimised through correlations with field observations (see Figure 7). The primary disadvantage of elastic modelling is that stresses are not re-distributed when the strength is exceeded. This applies to both pillar yielding as well as fault slip problems. This is of course only a problem if the amount of stress re-distribution causes a significant loss of accuracy. This remains to be demonstrated.

3D plastic modelling (Map3D, FLAC3D, 3DEC, ABACUS, ELFIN, etc.) allows one to incorporate stress transfer due to yielding pillars or abutments and slipping faults. Plastic models solve the same set of equations as elastic models, and in addition the equations of plasticity as well. Strength parameters are used as parameters in these later equations and hence directly affect the model predictions. These models give the same results as elastic models until the stresses exceed the strength.

The simplest possible plastic constitutive model is elastic-perfectly plastic (meaning the residual strength is set equal to the peak strength). This type of model at least doubles the number of model input parameters (over elastic models) since it requires the original 3 or 4 parameters from the pre-mining stress state plus additional parameters to describe the flow rule (strength criterion) and dilation rate. Also since strains are now used for calibrations, Young's modulus must also be specified. More complex plasticity models (such as strain softening) require specification of many more parameters (residual strength parameters and softening rates).

The advantages and disadvantages of plastic models turn out to be just the opposite of elastic models. The primary advantage of plastic modelling is the more realistic simulation capability. Stress transfer due to yielding pillars and faults is readily accommodated. Unfortunately, the large numbers of input parameters

presents a major disadvantage for plastic modelling since uncertainty in parameters is the primary cause of poor model predictions. In addition, significant yielding is always associated with long run times.

3.5   Material properties

Since elastic predictions do not depend on the strength parameters (and usually not on Young's modulus or Poisson's ratio) there are no material property input parameters to calibrate. Back-analyses do not have to be re-run in order to trial various strength parameters. Calibration can be done as post-processing.

In elasto-plastic analysis, the results depend directly on the material properties. Unlike elastic modelling, if you alter the strength parameters your model will give entirely different results. Calibration requires that for each input parameter trialled the back-analysis must be re-run.

3.6   Numerical accuracy

In all numerical models the calculation accuracy depends directly on the fineness of the discretisation. The numerical error is readily quantifiable and basically this boils down to ratio of the distance from the nearest excavation surface divided by the element width (i.e. size of the elements used to approximate the equations). To obtain accurate results near excavation surfaces you need to ensure that the element size is comparatively small. Each numerical model has its own error characteristic depending on what type of elements are used (constant, linear etc.) and the type of solution procedure (direct BEM, indirect BEM, FEM, FDM etc.).

Since analysis effort increases rather dramatically with the number of elements, there is always an overwhelming desire to use fewer elements and hence large element sizes. This is of course acceptable as long as one does not compromise the calculation accuracy. The important point here is that these errors are easily controlled to any desired magnitude simply by adjusting the discretisation fineness.

In 2D analyses it is common to see detailed excavation shapes with hundreds of elements around the perimeter of excavations and dozens of elements across the width of pillars. In 3D elastic analyses it is most common to see more simplified shapes with only dozens of elements around the perimeter of excavations and maybe a half-dozen elements across the width of pillars. In 3D plastic analyses it is common to see very blocky approximations for excavation shapes often with only one or two elements across the width of pillars.

Clearly numerical accuracy is routinely compromised simply to reduce run-times to more manageable levels. As soon as one compares results from different types of numerical models (using the same model type) it becomes very obvious that numerical errors can be significant as these predictions rarely agree very well.

3.7   Observations

While observations are not a model input, quality observations are required. If you are unsure of when pillars fail (Figure 2), the locations of the seismic events (Figure 3), or the actual over-break depth (Figure 5), this will introduce scatter into the back-analysis results, provide larger values for the standard deviation and hence result in less reliable model predictions.

4   Reliability of model predictions

Now that we have defined how to quantify the reliability of model predictions and identified the model input parameters, it is time to discuss how the various inputs interact to control prediction accuracy.

4.1   Combined interaction of input parameters

In the previous section a full list of input parameters were given. If we express the uncertainty in each parameter as a coefficient of variation (recall that coefficient of variation is just standard deviation normalised with respect to magnitude) then we can estimate the reliability of our model prediction (Wiles, 2006), C_{Output} as

$$\sqrt{C_{Geometry}^2 + C_{Geology}^2 + C_{\sigma^f}^2 + C_{Model}^2 + C_{Properties}^2 + C_{Numerical}^2} = C_{Output}$$

(1)

(here it has been assumed that each input parameter is independent and can be described by a normal distribution). C_{Output} represents the same coefficient of variation that was calculated from Figures 2, 3 and 5.

The most striking result of looking at modelling from this point of view is that it immediately becomes clear that the largest uncertainty dominates the prediction accuracy. For example, suppose that all coefficients in the above equation have very small values except for $C_{\sigma f}$ and C_{Model} which equal respectively ±20% and ±10%. Then it is easy to calculate that C_{Output} is ±22%. If we now improve the model type (for example going from an elastic model to a plastic model) to provide C_{Model} equal to ±5% then we see that C_{Output} reduces to only ±21%. In order to get a real improvement we must reduce $C_{\sigma f}$, not C_{Model}.

Another important point is that we must take care not to increase the magnitude of one parameter at the expense of reducing another. Most often when we reduce C_{Model} by improving the model type from 2D to 3D or from elastic to plastic, we unwittingly use larger element sizes in an effort to shorten run times and simplify the model construction process. This would have the affect of increasing $C_{Geometry}$ and/or $C_{Numerical}$ and it is easy to see that without realising it we could very well end up with a larger value for C_{output} and hence less reliability.

As each of the input parameters directly affects the model predictions, the most cost effective model is one where each parameter is specified with similar values of reliability. There is little to gain by specifying one parameter with high precision if another has large uncertainty.

We must consider the fact that there is natural variability in all rock masses and there will be some lower limit to the attainable value of C_{Output} regardless of how well we do our modelling. Small scale variations in the pre-mining stress state, material properties and geology can never be fully accommodated. For example, there is no point in setting a model to provide ±1% $C_{Numerical}$ since other contributions will always overwhelm any gains to be had from this extreme precision.

4.2 Model calibration - number of input parameters

The various model refinements that we can use (multiple materials zones, elastic, plastic, faults etc.) provide a potential improvement in accuracy of our model predictions. With increased constitutive capability, these models can simulate our rock mass response more realistically by taking into account full three-dimensional geometries, stiffness and pre-mining stress state varying across mine sites, stress re-distribution due to plastic yielding of pillars or faults, etc. However, this increase in constitutive capability comes at the cost of having to specify ever more material properties used as parameters in these equations.

Recall that in Figure 5 the standard deviation was determined as 1.85 MPa (coefficient of variation of ±12%). This was determined for the series of back-analyses that used the pre-mining stress state measured by over-coring using HI-cells (hollow inclusion stress meters) with the maximum principal stress oriented with a plunge of 30° and azimuth of 90°. In Figure 7, this same series of elastic back-analyses has been repeatedly re-run with a variety of different pre-mining stress states. In this case, only the orientation of the pre-mining stress state has been varied. A variety of stress ratios could be trialed as well. The standard deviation has been re-calculated (as was done in Figure 5) for each combination of azimuth and plunge. This same procedure can be used find the best value for any model input variable.

Clearly there is one pre-mining stress state (plunge of 20° and azimuth of 100°) that gives smallest value for standard deviation of 1.28 MPa (coefficient of variation of ±8.5%). It is this pre-mining stress state that allows our model to best match the historical observations of our mine response and hence this is the best choice of input parameters for our model. This begs the obvious question of whether this is a better estimate of the pre-mining stress state than what was obtained from the stress measurements. Alternatively, the bedding plane orientation could have been specified inaccurately. Although the results from only one cavity are illustrated here, it would obviously be desirable to superimpose the results for several cavities.
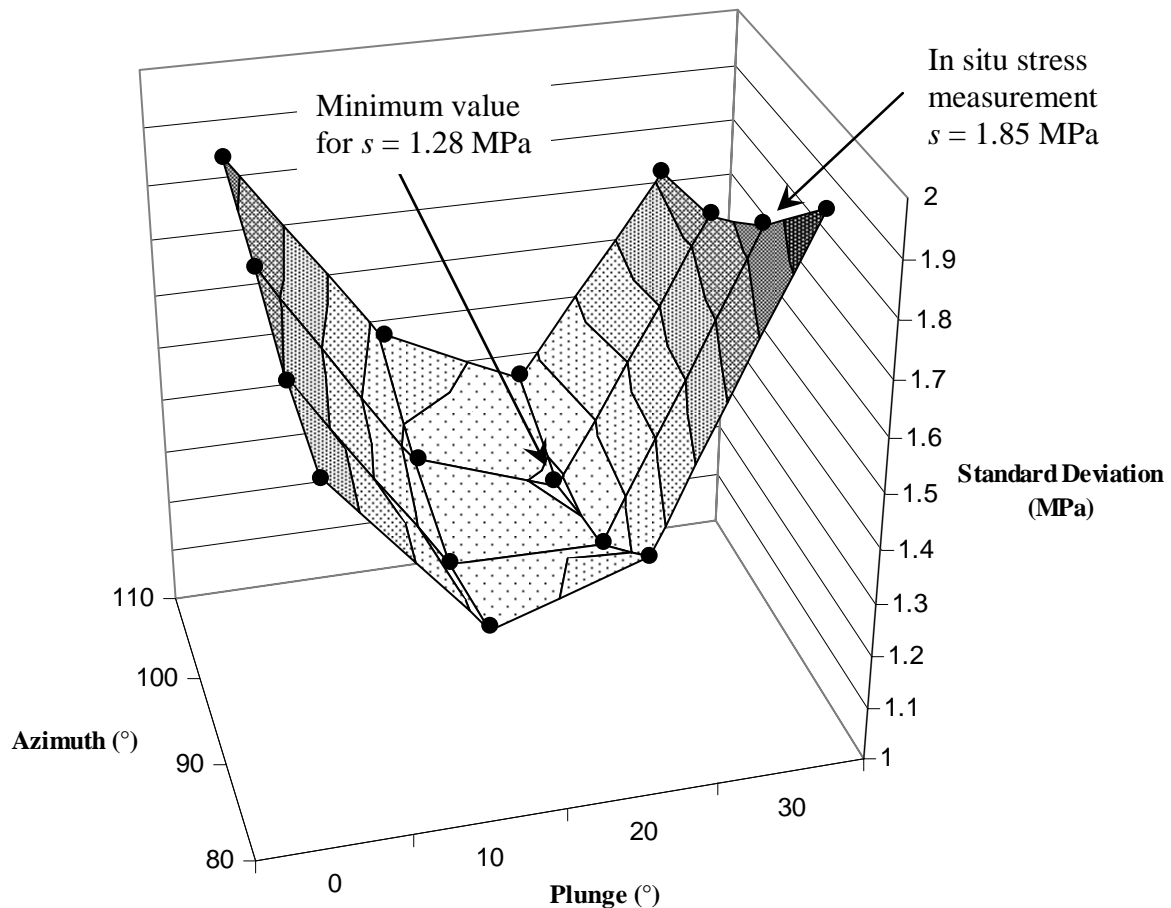
Figure 7   Error distribution for various pre-mining stress orientations

A close look at Figure 7 reveals that four values of plunge and four values of azimuth were trialled requiring that we re-ran our back-analysis series 16 times in order to define the best choice of input parameters. If we were to trial a 3rd parameter simultaneously then we would need to re-run the 16 back-analysis series four more times, or a total of 64 times. It is easy to see that the total number of back-analyses required depends heavily on the number of parameters we calibrate, and the total number of back-analysis series that need to be run is equal to 4^n where n represents the number of parameters we are calibrating.

Model calibration can be viewed as a curve fitting exercise. More complex models can be equated to higher order functional approximations with of course more parameters. With a large number of parameters, there are many ways of achieving the same model prediction hence there are also many more ways of getting the wrong prediction. When we simulate yielding material (pillars or faults) we are instructing the model how to behave. The strength parameters we specify directly control when, where and how much yielding occurs. This makes it important to ensure that each parameter chosen is tested to give accurate results under a wide range of conditions. Since these parameters are not inputs to simpler elastic models this concern does not arise.

While model input variables can be estimated by laboratory and in situ testing, back-analysis is the only way to verify that the estimated values are accurate and to systematically improve on those estimates.

For equilibrium models (tributary area) there is really only 1 parameter (hangingwall/footwall stress) and hence the effort for calibration is very small ($4^n = 4$). For elastic models, there are around 3 parameters hence the effort for calibration is larger ($4^n = 64$). For elasto-plastic models, there are at least 6 parameters

$(4^n = 4096)$ and usually 9 or more $(4^n \geq 262144)$. Clearly for more complex models the effort required for rigorous model calibration far exceeds practical limitations and hence this is never carried out.

5   Making model predictions with known reliability

In Figures 2, 3 and 5 it was shown how the best fit line and standard deviation could be calculated using linear regression. If we assume that a normal distribution (although any distribution would suffice) adequately represents the observed scatter it is easy to make model predictions with known reliability using probability theory. The procedure is to use the numerical model to predict the stresses for the new mining geometry then calculate the probability. Let's illustrate this procedure with a couple of examples (Wiles, 2006).

5.1   Ground support in highly stressed ground
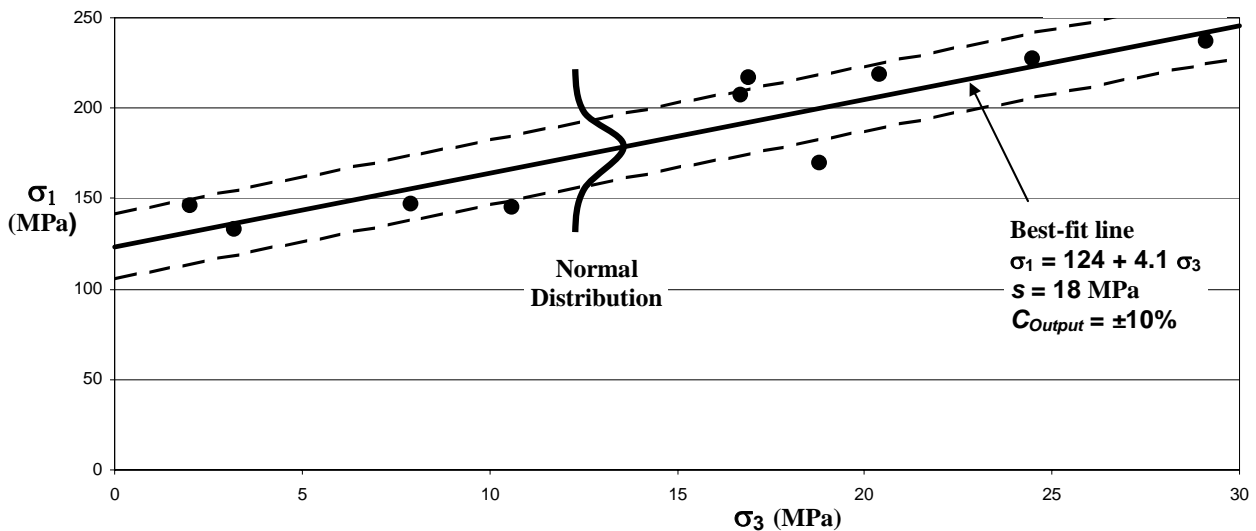
Figure 8   Back-analysis data with the best-fit line and normal distribution

Figure 8, is a reproduction of Figure 2 where linear regression has been used to determine the best-fit line (shown solid) and the lines representing plus and minus one standard deviation (shown dashed).
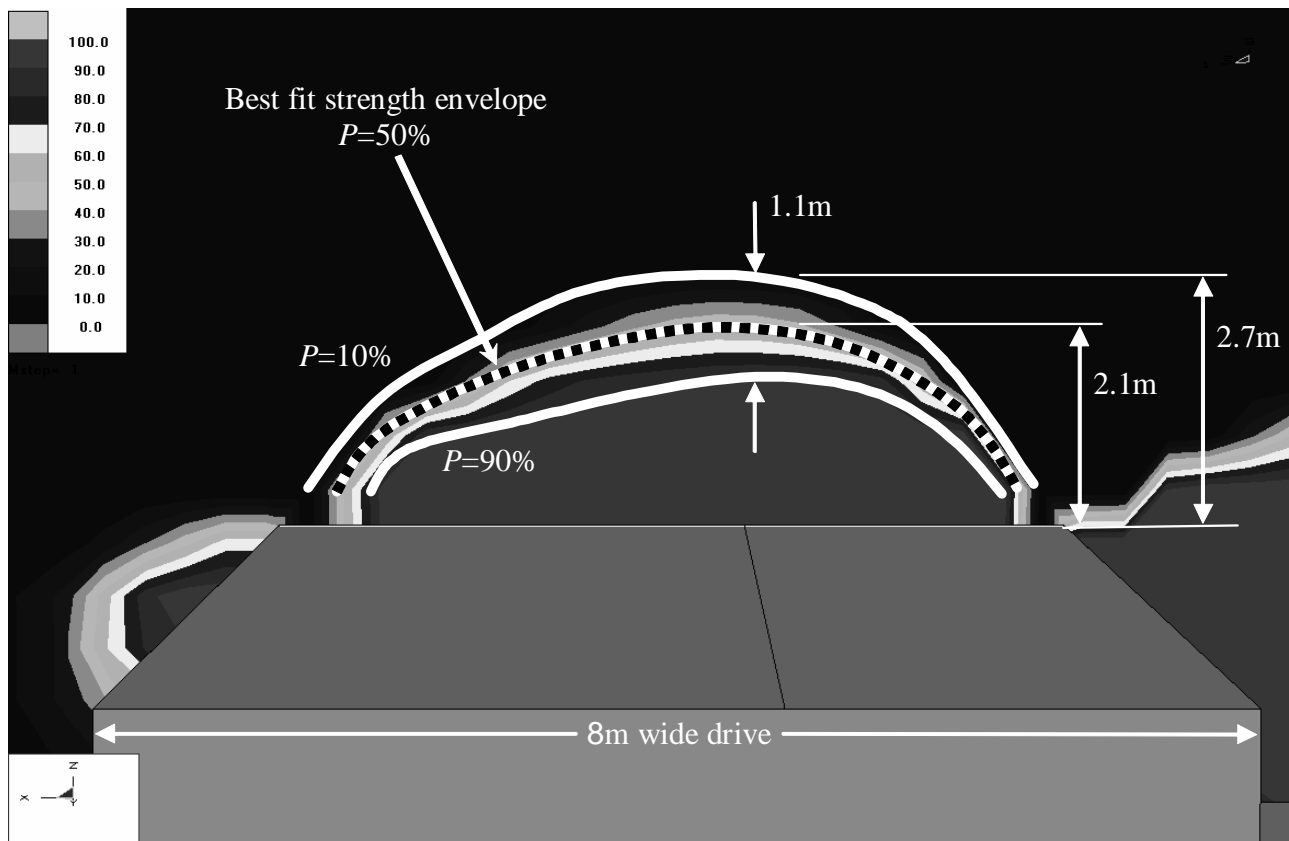
The probability of any stress prediction causing stress induced failure can be calculated using

$$P = N(\Delta\sigma_1 / s)$$

$$\Delta\sigma_1 = \sigma_1 - [UCS + q \sigma_3]$$

(2)

Where:

|		
N	=	normal distribution.
s	=	standard deviation.
UCS	=	the intercept of the best-fit line.
q	=	slope of the best-fit line.

Figure 9   Predicting depth of failure

This is done in Figure 9. The logic here is that based on historical observations we have calibrated a model and quantified how well the model was able to reproduce these observations. We now make a forward prediction by assuming that we expect the model to continue to be able to predict with the same degree of reliability into the future. The zone of uncertainty is the area between the 10% and 90% probability contours and is 1.1 m wide.

Over the back of a long drive we would expect to see the depth of failure varying around the mean depth of 2.1 m (*P* of 50%). We expect to see failure as deep as 2.7 m 10% of the time. Hence in order for our support design to be successful along 90% of the drift length we should design for 2.7 m of broken ground.

This method of interpretation can be used directly to conduct cost-benefit comparisons for various ground support designs. Here we could trade-off the cost of the heavier ground support required for fewer support failures and hence less reconditioning.

5.2   Stope dilution
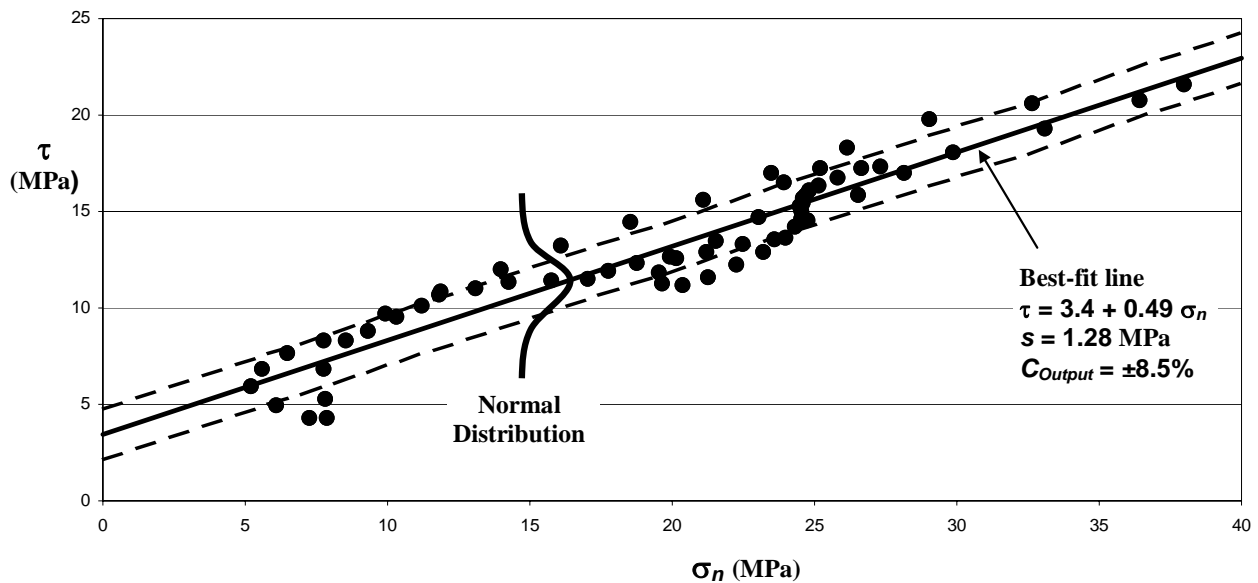
Figure 10   Back-analysis data with the best-fit line and normal distribution

The results in Figure 10 are similar to those in Figure 5, but here the pre-mining stress state that gives the smallest value for *s* has been used (as determined in Figure 7). Note the reduced scatter here. Linear regression has been used to determine the best-fit line (shown solid) and the lines representing plus and minus one standard deviation (shown dashed).

The probability of any stress prediction causing over-break can be calculated using

$$P = N(\Delta\tau / s)$$

$$\Delta\tau = \tau - [Coh + \sigma_n \ \tan(\varphi)] \tag{3}$$

Where:

Coh   =   the intercept of the best-fit line.

φ =   slope of the best-fit line.

This is done in Figure 11. As before, the logic here is that based on historical observations we have calibrated a model and quantified how well the model was able to reproduce these observations. We now make a forward prediction by assuming that we expect the model to continue to be able to predict with the same degree of reliability into the future. The zone of uncertainty is the area between the 10% and 90% probability contours and is 3.8 m wide.
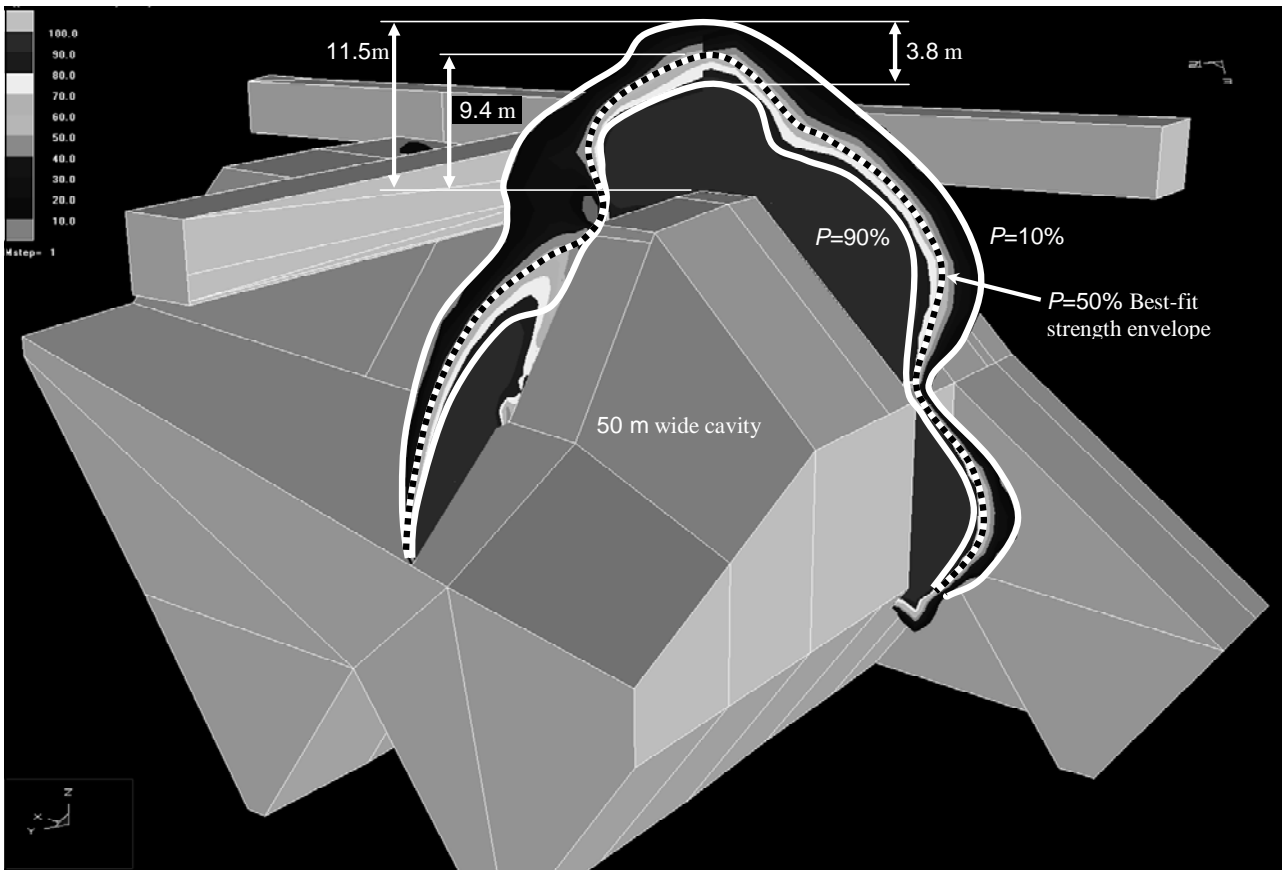
Figure 11   Predicting depth of failure

Over the back we would expect to see the depth of failure varying around the mean depth of 9.4 m with a 50% likelihood that the dilution volume will be 44000 m^3. Since we expect to see failure as deep as 11.5 m, 10% of the time, it follows that there is a 10% likelihood that the dilution volume could be as high as 76000 m^3. If we plan to depend on this model we must be prepared for the possibility that the dilution could be quite large.

5.3   Discussion on reliable model predictions

In both of the above examples there was a limited amount of accuracy with which the models were able to predict the mine response (i.e. the distance between the 10% and 90% probability contours). The uncertainty of the model predictions forces us to adopt a certain level of conservatism in our mine design. This leaves us with alternatives:

- Bear the cost of a more conservative mine design (larger mining costs).

- Attempt to improve the reliability of our model and hence narrow the zone of uncertainty (larger modelling costs).
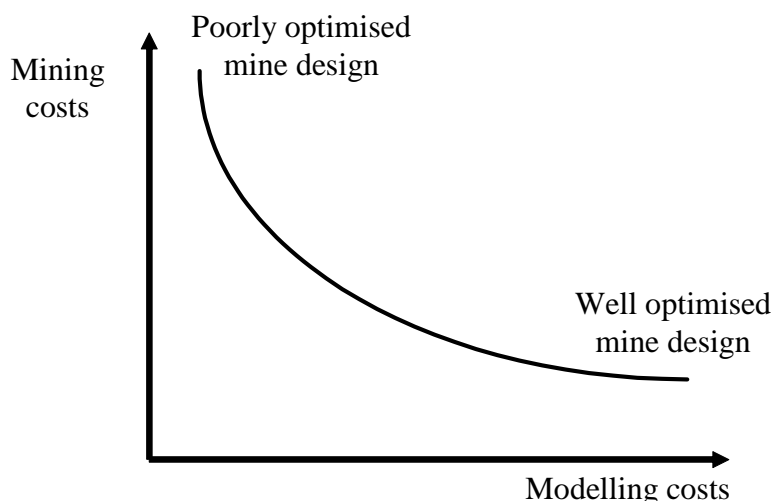
Figure 12   Design optimisation

These alternatives are illustrated in Figure 12. In order to improve the reliability of our model we would need to spend additional effort in calibration or perhaps utilise a more complex model, in an attempt to reduce the scatter in our back-analysis results (as was done in Figure 7). This would result in a smaller value for standard deviation and provide a narrower zone of uncertainty.

Improved reliability allows us to use smaller values of safety factor in our design. In fact it can be readily shown (Wiles, 2006) that safety factor can be directly related to probability as

$$SF \approx 1 \, / \, [\, 1 + C_{Output} \, N^{-1}(P) \,]$$

(4)

Where:

N^{1} =   the inverse normal distribution.

Clearly good fitting models (with small C_{Output}) allow for the use of small safety factors to achieve a low probability of failure. Unfortunately poorly fitting models require large safety factors to achieve the same probability of failure. Unless C_{Output} is determined the required safety factor is not known.

6   Discussion - selecting the best model

Clearly simple models with few parameters are desirable since they are the least costly to setup and run. Owing to the short run times and small number of input parameters these models can be rigorously calibrated at low cost. The argument is often made that simple models are unable to follow complex stress strain response and hence they can potentially give inaccurate predictions. While this is of course true, it remains to be seen how significant this effect is compared to other sources of error.

Complex 3D plastic models with a large number of parameters are costly to setup and run. These models are able to follow a more realistic stress strain response and hence have the potential to give more accurate predictions. This potential will only be realised if there is significant stress transfer resulting from the plastic yielding. The occurrence of plastic yielding is not by itself sufficient to cause a significant amount of load transfer. While this is likely true in soil mechanics problems (where the stresses routinely exceed the strength by a large margin), it remains to be demonstrated for hard-rock mining environments.

These predictions come at a high cost. Geometry is often over simplified and numerical accuracy is often sacrificed in order to keep the costs low (refer to Equation 1). Even though the level of accuracy can be readily determined by repeatedly conducting back-analyses with varying input parameters, this is rarely completed owing to the large costs involved. Rigorous calibration to find the best input parameters (as in Figure 7) is never completed because of the large number of input parameters.

Fortunately the level of accuracy can be readily measured by conducting back-analyses to determine exactly how well these models can match observed mine behaviour. If it is found that the level of accuracy is not satisfactory, there are many ways in which we could improve our model (as shown in Figure 6): *Geometry*, *Geology*, *Pre-mining stress* σ^f, *Model Type*, *Material Properties*, *Numerical Accuracy*, and *Quality Observations*.

We could check the uncertainty associated with each of these model inputs in turn to look for the sources of error that are causing the scattered back-analysis results. It is of course possible that we have simplified the geometry too much. Perhaps we have ignored some important geological feature such as stiff dyke. We may have used inappropriate values for the pre-mining stress state or material properties and could consider conducting an error minimisation study (Figure 7) to find better values. The discretisation may be too coarse or our observational data is in error. Finally we also need to consider use of a more complex plastic fault-slip or pillar yielding model.

If we now re-consider Equation 1 we can see that just because we add one refinement or another it does not necessarily follow that there will be any significant reduction in the model variability C_{Output}. Only those contributions that have large uncertainty will provide real improvements. By re-running the back-analyses we can readily compare the reliability obtained by adding each of the alternative refinements.

With this in mind let's consider the cost-benefit of alternative model types.
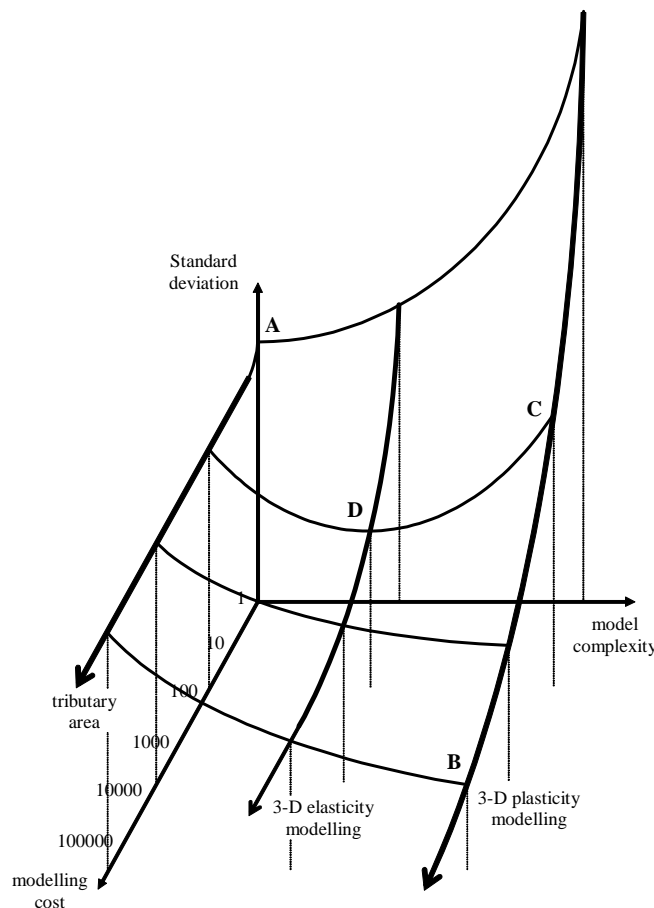
Figure 13   Trading off the cost-benefit of alternative model types

Figure 13 embodies Equation 1 pictorially. In this figure it can be observed that simple models such as tributary area are inexpensive to run but cannot provide much accuracy (point A). In spite of this limitation, with a very small budget you probably get the most benefit from these models. At the other extreme, 3D plasticity models are the most expensive to run but have the potential to provide the best accuracy (point B). Unfortunately, this potential for higher accuracy can only be realised with considerable calibration effort owing to the large number of input parameters that require specification. If you do not have a sufficiently

large budget for this calibration then you could conceivably arrive at lower accuracy predictions (point C) than you could have obtained by adopting a less expensive alternative (point D). Just because you have run a more complex plasticity model it does not mean that you will automatically get more useful answers.

In other words, for the same budget it is better to choose a well calibrated simple model than a poorly calibrated complex model. With a simpler model at relatively low cost (shorter setup and run times) you can systematically test and verify the effect of each variable. You can readily determine quantitatively what the reliability is. This allows you to model with a known level of confidence regardless of whether this turns out to be high or low. Even in cases where the reliability is low, at least you can quantify how uncertain the predictions can be.

In a complex model there will always be many untested input variables simply because a fixed budget will not allow for rigorous calibration (longer setup and run times). There are many ways of achieving the same prediction hence there are also many more ways of getting the wrong result. There will be uncertainty as to exactly what the reliability is. This provides models with an unknown level of confidence. With unknown reliability you cannot quantify how uncertain the predictions can be.

7   Conclusions

The reliability of numerical models can be readily quantified in terms of standard deviation (or coefficient of variation) by straight forward comparison with historical observations of mine behaviour (Figures 8 and 10).

By using a probabilistic approach to design, the relationship between model reliability and uncertainty in predictions can be clearly established (Figures 9 and 11).

With these two measures quantified, we are now in a position to directly test and justify the cost effectiveness of the many alternative options for improving model reliability. All we need to do is to re-run the same series of back-analyses with the various alternatives to see if they do improve the reliability (i.e. reduce the scatter in the back-analysis results) and see if they result in a sufficiently reduced zone of uncertainty (i.e. the distance between the 10% and 90% probability contours) to warrant the cost and effort.

The choice of the most cost effective model can be rationally made by careful consideration of the effort required to control the reliability of each input parameter (Equation 1 and Figure 6). With a limited budget the best reliability is likely to be obtained from simpler models that can be readily calibrated.

There is no doubt that more complex plasticity models are able to provide more realistic stress-strain response, but more accurate predictions can will only be realised if one has a sufficiently large budget to complete the verifications required. Keeping in mind that stress-strain response is just one of many sources of error, there is no guarantee that this will provide a significant improvement in reliability. It is entirely possible that other input variables dominate the total error (Equation 1) and it is these sources of error that need to be addressed.

Alternatives that can be considered include refinement of the geometry, geology, pre-mining stress-state, material properties, numerical errors and better observations. These can be refined through lab testing, in situ monitoring and stress measurements, but also by conducting a rigorous calibration by systematically trialling a series of alternative pre-mining stress states or other input parameters (as in Figure 7). In all cases we can determine how much improvement was obtained and hence justify the cost of the various alternatives we have trialled, keeping in mind that there is a limit to how accurate any model can be due to the natural variability of the rock mass.

In situations where yield induced stress transfer dominates the model response there is no doubt that plastic fault-slip or pillar yielding will need to be implemented. While there is no technical obstacle to this one needs to be prepared for the fact that calibration will be expensive and time consuming. This emphasises the necessity for clear thinking about exactly what verifications need to be made and how we are going to go about this. Due to the high costs involved (longer setup and run times) and the greater potential for error (due to the increased number of input parameters) there is little room for guesswork. The challenge here is to find a way to get the maximum value from these expensive exercises.

By making modelling an evidence-based objective exercise we can determine how reliable our modelling predictions are, so that we can depend on them when they are good and disregard them when they are not.

This allows us to evaluate the cost effectiveness of implementing the many alternatives we can try to reduce errors in our predictions. We need to consciously avoid situations where we exhaust our budget addressing the wrong sources of error before the problems at hand are properly solved.

In this paper I have laid out a systematic procedure for measuring the reliability of alternative models through repeated back-analysis on the same data set. Rather than speculate about which model will give a more favorable outcome from such comparisons, it is time to actually run the alternatives and compare the reliability to see if there is any real cost justification.

Acknowledgements

References

Jamieson, M. (2007) Personal Communication, Senior Rock Mechanics Engineer, Xstrata Copper – Mount Isa Copper Operations, Mt. Isa, Queensland, Australia.

Martin, C.D. (1997) The effect of cohesion loss and stress path on brittle rock strength. 17th Canadian Geotechnical Colloquium, Canadian Geotechnical Journal, Vol. 34(5), pp. 698-725.

Wiles, T.D. and Kaiser, P.K. (1994) In situ stress determination using the under-excavation technique – I. theory. Int. J. Rock. Mech. Min. Sci. Vol. 31, pp. 439-446.

Wiles, T.D. and Kaiser, P.K. (1994) In situ stress determination using the under-excavation technique – II. Applications. Int. J. Rock. Mech. Min. Sci., Vol. 31, pp. 447-456.

Wiles, T.D., Villaescusa, E. and Windsor, C. (2004) Rock reinforcement design for overstressed rock using three-dimensional numerical modelling. 5th International Symposium on Ground Support, Perth, Australia, 2004, pp. 483-489.

Wiles, T.D. (2005) Rockburst Prediction Using Numerical Modelling: Realistic Limits for Failure Prediction Accuracy. 6th International Symposium on Rockbursts and Seismicity in Mines (RaSiM 6), Perth, Australia, 2005, pp. 57-63.

Wiles, T.D. (2006) Reliability of Numerical Modelling Predictions, Int. J. Rock Mech. & Min. Sci. 43, pp. 454-472.

Wiles, T.D. (2007) Map3D User's Manual, available from http://www.map3d.com.