

## SURROGATE BASED SENSITIVITY ANALYSIS OF PROCESS EQUIPMENT

**D.W. STEPHENS<sup>1\*</sup>, D. GORISSEN<sup>2</sup> AND T. DHAENE<sup>2</sup>**

<sup>1</sup>Parker Centre (CSIRO Minerals), Clayton, Victoria 3169, AUSTRALIA

<sup>2</sup>Ghent University – IBBT, Department of Information Technology (INTEC), Gaston Crommenlaan 8, 9050 Ghent, BELGIUM.

\*Darrin.Stephens@csiro.au

### ABSTRACT

The computational cost associated with the use of high-fidelity Computational Fluid Dynamics (CFD) models poses a serious impediment to the successful application of formal sensitivity analysis in engineering design. Even though advances in computing hardware and parallel processing have reduced costs by orders of magnitude over the last few decades, the fidelity with which engineers desire to model engineering systems has also increased considerably. Evaluation of such high-fidelity models may take significant computational time for complex geometries.

In many engineering design problems, thousands of function evaluations may be required to undertake a sensitivity analysis. As a result, CFD models are often impractical to use for design sensitivity analyses. In contrast, surrogate models are compact and cheap to evaluate (order of seconds or less) and can therefore be easily used for such tasks.

This paper discusses and demonstrates the application of several common surrogate modelling techniques to a CFD model of flocculant adsorption in an industrial thickener. Results from conducting sensitivity analyses on the surrogates are also presented.

### NOMENCLATURE

$A$	radial angle between flocculant sparge and feed pipe
$A_j$	Fourier coefficients
$b$	LS-SVM model parameter
$b'_j$	neuron bias
$B$	vertical location of flocculant sparge
$B_j$	Fourier coefficients
$C$	distance from feedwell wall to flocculant sparge
$D$	variance
$e_k$	acceptable error
$E(y)$	expectation of $y$
$F$	ANN activation function
$G_i$	transformation functions
$K$	RBF kernel function
$N$	number of sample points or number of neurons in ANN
$p$	integer number
$s$	real number
$S$	main effect
$ST$	total effect

$u$	input to neuron
$V(y)$	variance of $y$
$w$	LS-SVM model parameter
$w_{ij}$	ANN weights
$x$	input sample points
$y$	true model output
$\bar{y}$	mean true model output
$\hat{y}$	predicted model output
$\alpha_k$	LS-SVM model parameters
$\beta$	RBF model parameters
$\varphi$	input to output mapping function
$\gamma$	LS-SVM regularization constant
$\omega$	Frequency

### Subscripts

$i, j, k$  indices

### Superscript

$l$  neuron layer

### INTRODUCTION

For many industrial fluid dynamics problems, it is impractical to perform experiments on the physical world directly. Instead, complex, physics-based simulation codes are used to run experiments on computer hardware. Accurate, high-fidelity Computational Fluid Dynamics (CFD) models are typically time consuming and computationally expensive, this poses a serious impediment to the successful application of formal sensitivity analysis in engineering design. While advances in High Performance Computing and multi-core architectures have helped, routine tasks such as visualisation, design space exploration, sensitivity analysis and optimisation quickly become impractical (Simpson et al., 2008; Forrester et al., 2008). As a result, researchers have turned to various methods to mimic the behaviour of the simulation model as closely as possible, while being computationally cheaper to evaluate. This work concentrates on the use of data-driven, global approximations using compact surrogate models in the context of computer experiments. The objective is to construct a surrogate model that is as accurate as possible over the complete design space of interest using as few simulation points as possible. Once constructed, the global surrogate model is reused in other stages of the computational engineering pipeline, such as sensitivity analysis.

Sensitivity analysis of model output aims to quantify how a model depends on its input factors. Global sensitivity determines the effect on model output of all the input parameters acting simultaneously over their ranges. Most global sensitivity techniques are variance-based methods and determine the fractional contribution of each input factor to the variance of a model output. The main difficulty with global sensitivity analysis is that the number of model evaluations required is often large. As a result, CFD models are often impractical to use for design sensitivity analyses.

The objective of this paper is to discuss and demonstrate the application of several common surrogate modelling techniques to a case study of a CFD model of flocculant adsorption in an industrial thickener. A sensitivity analysis is then conducted on the produced surrogate models.

## SURROGATE MODELLING

### Radial Basis Functions

Radial basis function (RBF) models use linear combinations of radially symmetric functions to interpolate samples data points. The simplest form of these models is

$$\tilde{y}(x) = \beta_0 + \sum_{i=1}^N \beta_i \|x - x_i\| \quad (1)$$

Where  $\|\cdot\|$  is the Euclidean distance,  $N$  is the total number of sample points,  $x_i$  is the  $i^{\text{th}}$  sample point, and the  $\beta$ 's are model parameters found solving a linear system of  $N$  equations. RBF models are shown to produce a good fit for arbitrary contours (Powell, 1987).

### Artificial Neural Networks

Artificial neural networks (ANN) are massively parallel highly interconnected simple processors (neurons). A typical ANN structure consists of an input layer into which the independent variables are fed, the output layer that produces the dependent variables and one or more hidden layers. The hidden layers link the input and output layers together and allow for complex, nonlinear mapping from the input to the output. The mapping is not specified but is learned. An artificial neural network can be described in terms of the individual neurons, the network connectivity, the weights associated with the interconnections between neurons, and the activation function of each neuron.

A typical architecture, known as the multilayer feed-forward network, is shown in Figure 1 and is used in the simulations in this work. In this figure the lines represent the unidirectional feed-forward communication links between the neurons. A weight associated with each of these connections controls the output passing through a connection. The output of a neuron in the feed-forward network shown in Figure 1 may be represented as (Mahajan, 1999):

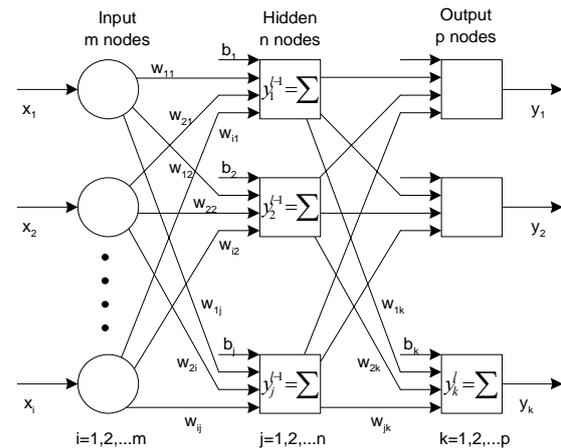
$$\tilde{y}_j^l = F \left( \sum_{i=1}^{N_{l-1}} w_{ij}^l y_i^{l-1} + b_j^l \right) \quad (2)$$

where  $\tilde{y}_j^l$  is the output of the  $j^{\text{th}}$  neuron in the  $l^{\text{th}}$  layer,  $w_{ij}^l$  is the weight on the connection from the  $i^{\text{th}}$  neuron in the  $(l-1)^{\text{th}}$  layer to the  $j^{\text{th}}$  neuron in the  $l^{\text{th}}$  layer,  $b_j^l$  is the bias connected to the  $j^{\text{th}}$  neuron in the  $l^{\text{th}}$  layer and  $N_{l-1}$  is the number of neurons in the  $(l-1)$  layer. The activation function  $F$  is typically given by the sigmoid function:

$$F(u) = \frac{1}{(1 + e^{-u})} \quad (3)$$

It serves to bound the output from any neuron in the network and allows the network to handle both small and large inputs.

The network is trained by minimising the mean-squared error between the output of the output layer and the target output for all the input patterns. An artificial neural network starts out with random weights, and the weights are adjusted until the required degree of accuracy is obtained, this process is called learning. A number of learning schemes are available to train an ANN (Møller 1993). These schemes govern how the weights are to be varied to minimize the error at the output nodes. In this study the training is done using Levenberg Marquard back propagation with Bayesian regularization (MacKay, 1992; Foresee and Hagan, 1997).



**Figure 1:** Multilayer feed-forward Artificial Neural Network (ANN).

### Least Squares-Support Vector Machines

The theoretical background of support vector machines (SVM) is mainly inspired from statistical learning theory (Vapnik, 1999). Major advantages of SVM over other machine learning models (such as neural networks) are that there is no local minima during learning and the generalization error does not depend on the dimension of the space. The core attraction of support vector methods is the promise of a global solution. Only the solution of linear equations is needed in the optimisation process, which not only simplifies the process, but also avoids the problem of local minima in SVM. The LS-SVM model (Sundaram et al., 2005) is defined in its primal weight space by:

$$\tilde{y}(x) = w^T \varphi(x) + b \quad (4)$$

where  $\varphi(x)$  is a function which maps the input space into a higher dimensional feature space,  $x$  is the  $M$ -dimensional vector of inputs  $x_j$ , and  $w$  and  $b$  the parameters of the model. Given  $N$  input-output training pairs  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ , LS-SVM for function estimation formulate the following optimisation:

$$\min_{w, b, e} J_p(w, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^N e_k^2 \quad (5)$$

subject to  $\tilde{y}_k = w^T \varphi(x_k) + b + e_k$  for  $k = 1, \dots, N$

The parameter set consists of vector  $w$  and scalar  $b$ . Solving this optimisation problem in dual space leads to finding the  $\alpha_k$  and  $b$  coefficients in the following LS-SVM regression model:

$$\tilde{y}(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b \quad (6)$$

The function  $K(x, x_k)$  is the dot product between the  $\varphi(x)^T$  and  $\varphi(x)$  function mappings.  $K$  is called the kernel function in the SVM formulation and there are a number of possible choices for kernels, each with its own advantages and disadvantages. The most common kernels used are RBF, Sigmoid and Linear. We have used the RBF kernel in our simulations.

## MODEL ACCURACY ASSESSMENT AND CROSS-VALIDATION

It is essential to assess the accuracy of a surrogate for prediction before it can be used for sensitivity analysis studies. Model accuracy is usually assessed by comparing some response data produced by the analysis code (CFD output) with corresponding response data predicted by the model. An error measure is calculated based on these two sets of values to quantify the degree of accuracy.

After a model is built, a straightforward way to assess its accuracy is to run the analysis code at a large additional set of sample points and compare the output with those predicted by the model (Wang and Lowther, 2006). However, this comes at a high cost, as CFD simulations may take significant computational time. Cross-validation is a model assessment method that can estimate the accuracy of a model without requiring any additional sample points. In general, the data is divided into  $k$  subsets ( $k$ -fold cross-validation) of approximately equal size. A surrogate model is constructed  $k$  times, each time leaving out one of the subsets from training, and using the omitted subset to compute the error measure of interest. The average of the error measure is calculated from the result in each of the  $k$  iterations. The average is then the cross-validation's estimate of the model accuracy.

Cross-validation works by estimating a model accuracy measure with only a limited number of samples points. The computational cost of cross-validation (i.e. generating  $k$  surrogates) is justified by the fact that model fitting and assessment usually takes only seconds to minutes, while a single run of the CFD code could last hours to days, if not longer.

While cross-validation almost always performs very well (Bengio and Chapados, 2003), Gorissen et al. (2009a) showed that it is not always efficient at preventing

unwanted ripples or bumps in the final model response. A new measure called Linear Reference Model (LRM) has been developed by Gorissen et al. (2009b) that helps reduce this problem. Any visible behavioural complexity (bumps, ripples, etc) should be explainable by data points at (or in the vicinity of) those locations. Thus, intuitively, if there is no such evidence nearby, the bump should be regarded as an artefact of the model. The LRM metric is based upon the core assumption that if nothing else is known, the model behaviour between two neighbouring points should be linear. This is achieved by penalising a model proportional to how much it deviates from a linear fit. For a more detailed description of the LRM algorithm see Gorissen et al. (2009b).

## SENSITIVITY ANALYSIS

Sensitivity, in this context, is a measure of the contribution of an independent variable to the total variance of the dependent data. Sensitivity analysis allows addressing questions such as (Queipo et al., 2005):

- Can we safely fix one or more of the input variables without significantly affecting the output variability?
- How can we rank a set of input variables according to their contribution to the output variability?
- If and which parameters interact with each other?
- Does the model reproduce well known behaviour of the process of interest?

There are alternative approaches for sensitivity analysis, differing, for example, in scope (local versus global), nature (qualitative versus quantitative), and in whether they assume a particular model. In this paper we discuss the application of the Fourier amplitude sensitivity test (FAST), a variance-based non-parametric approach for analysis applications.

### Fourier Amplitude Sensitivity Test (FAST)

FAST is a sensitivity analysis method which works irrespective of the degree of linearity or additivity of the model (Cukier et al., 1973; Schaibly et al., 1973; Cukier et al., 1975; Cukier et al., 1978; Koda et al., 1979a; Koda et al., 1979b; Pierce et al., 1981; McRae et al., 1982). This procedure provides a way to estimate the expected value and variance of the output variable and contribution of individual input parameters to this variance. An advantage of FAST is that the evaluation of sensitivity estimates can be carried out independently for each parameter using just one simulation, because all the terms in a Fourier expansion are mutually orthogonal.

The main idea of the FAST method is to convert the  $n$ -dimensional integral required for the sensitivity calculation into one-dimensional integral in  $s$  by using the transformation  $X_i = G_i \sin(\omega_i s)$  for  $i = 1, \dots, n$ . For properly chosen  $\omega_i$  and  $G_i$ , the expectation of  $y$  can be approximated by:

$$E(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s) ds \quad (7)$$

where  $f(s) = f(G_1 \sin(\omega_1 s), \dots, G_n \sin(\omega_n s))$ .

Using the properties of Fourier series (Saltelli et al., 1998), an approximation to the variance of  $y$  is given by:

$$V(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2(s) ds - E(y)^2 \quad (8)$$

$$\approx 2 \sum_{j=1}^{\infty} (A_j^2 + B_j^2)$$

Where  $A_j$  and  $B_j$  are the Fourier coefficients. The expressions in (7) and (8) provided a means to estimate the expected value and variance associated with  $y$ .

Saltelli et al. (2000) proposed Extended FAST to not only calculate the main effects but also total effects to allow full quantification of the importance of each variable. Consider the frequencies that do not belong to the set  $p\omega_i$ . These frequencies contain information about interactions among factors at all orders not accounted for by the main effect indices. To calculate the total effects, we assign a high frequency  $\omega_i$  for the  $i^{\text{th}}$  variable and different low frequency  $\omega_{-i}$  to the remaining variables. By evaluating the spectrum at  $\omega_{-i}$  frequency and its harmonic, we can calculate the partial variance  $D_{-i}$ . This is the total effect of any order that does not include the  $i^{\text{th}}$  variable and is complementary to the variance of the  $i^{\text{th}}$  variable. Similar to Sobol' indices, the total variance due to the  $i^{\text{th}}$  variable is  $D_{T_i} = D - D_{-i}$  and the total effect is:

$$ST_i = S_i + S_{(i,-1)} \quad (9)$$

Where  $S_{-i}$  is the sum of all  $S_{i_1 \dots i_k}$  terms which do not include the term  $i$ .

The indices from Extended FAST are used within this paper and were calculated using the MATLAB version of SimLab distributed freely by the Joint Research Centre (<http://simlab.jrc.cec.eu.int>).

## CASE STUDY

### Problem description

CFD modelling of process unit operations is a tool that is being used increasingly within the minerals processing industry to reduce operating and capital costs and increase throughputs, one such unit operation where CFD has been applied is gravity thickening (Kahane et al., 1997; Johnston et al., 1998; Fawell et al., 2009; Kahane et al., 2002; White et al., 2003; Owen et al., 2009). The principle is simple – adding flocculant to aggregate the fine particles and allowing them to settle to produce clear overflow liquor and concentrated underflow slurry.

Understanding the flows within a feedwell is of critical importance to the overall thickener performance, as this is where most flocculation is induced, with the degree of turbulence a major factor in determining the structure and size of aggregates formed (Owen et al., 2009). Predictions of the solids distribution, liquid velocity vectors, shear rate and flocculant adsorption throughout a feedwell may be obtained from detailed knowledge of its geometry and the physical properties of the incoming feeds (Kahane et al., 2002). Attempts to optimise the performance of thickeners or clarifiers have usually depended on feedwell modification and positioning of the flocculant sparge being done on a trial and error basis. Owen et al. (2009) presented details of a multi-phase CFD model developed specifically for the analysis of the flow within feedwells

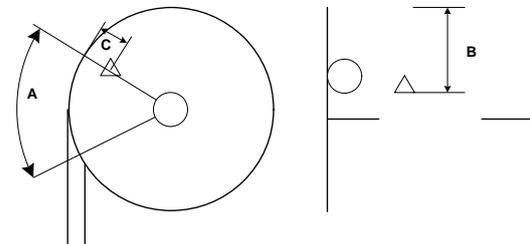
of industrial thickeners including the prediction of the adsorption of flocculant onto the feed solids. Such a model can be used to investigate the effect of flocculant sparge location on the flocculant particle coverage.

The aim of this study was to build a surrogate utilising the outputs from the CFD model detailed in Owen et al. (2009). Once built, the surrogate can be used to investigate the sensitivity of the model output to the various model inputs.

The feedwell geometry investigated in this study is a “generic” open feedwell with an attached shelf and a single feed inlet system entering the feedwell tangentially clockwise. The feedwell is 4.0 m in diameter with 3 m side-wall below the liquor surface and the feed inlet is 0.5 m in diameter, centred at 1 m below the liquor surface. The shelf is located 0.1 m below the base of the feed inlet. The model thickener has a diameter of 20 m, side-wall height of 5 m and is feed at  $1000 \text{ m}^3 \text{ h}^{-1}$ . The solids concentration is set 10 w/w%, the solids density to  $2710 \text{ kg m}^{-3}$  and the liquid density to  $1000 \text{ kg m}^{-3}$ . For modelling purposes, no specific flocculant chemistry is required, but a flocculant dosage (in this case  $20 \text{ g t}^{-1}$ ) needs to be specified.

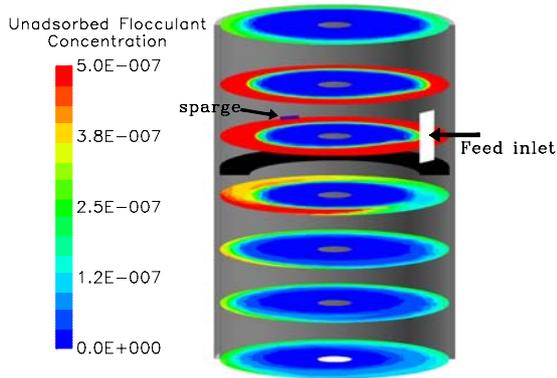
The parameters relating to the flocculant sparge location (Figure 2) evaluated for their influence on feedwell flocculant mixing and adsorption were as follows:

- Radial angle between flocculant sparge and feed pipe (A),
- Vertical location of flocculant sparge (B),
- Distance from feedwell wall to flocculant sparge (C).



**Figure 2:** Flocculant sparge location parameters used in surrogate building.

CFD computations were performed with the CFD software package PHOENICS (CHAM, London, UK) using the model described in Owen et al. (2009) with a mesh of about 155,000 grid cells. The information produced by the CFD thickener model is presented in terms of CFD images and calculated parameters. An example of the graphical output from the model is shown in Figure 3, where a series of horizontal slice planes through the feedwell are presented displaying the unadsorbed flocculant. In this figure, the feedwell is not plotted to scale, but stretched in the vertical direction to adequately display all slices.

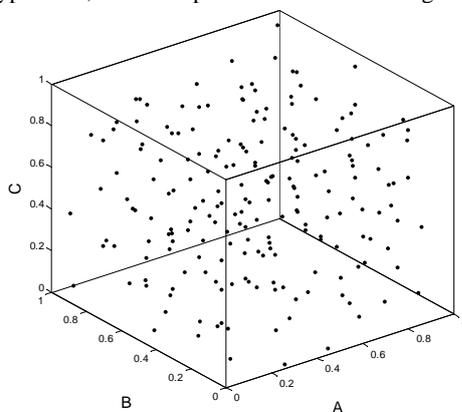


**Figure 3:** Plan view of unadsorbed flocculant concentration in the feedwell for one flocculant sparge position.

The approach detailed in Owen et al. (2009) was used for the simulations, where the time-dependent flocculant adsorption simulations were started from a converged steady-state flow solution and each took approximately 30 minutes. The flocculant solution was introduced into the feedwell with no actual velocity and hence the same steady-state flow solution could be used for each of the transient flocculant adsorption simulations. The output quantity from the CFD used for the surrogate building was the ‘floculant loss’. The flocculant loss is the steady-state value of the amount of unadsorbed flocculant leaving the feedwell divided by the incoming amount of flocculant through the sparge. Therefore, a value of zero represents complete adsorption of the flocculant onto the solid phase within the feedwell and a value of one represents all flocculant having left the feedwell without any adsorption onto the solids.

#### Sampling strategy

Each of the input variables, A, B, and C were scaled in the range 0-1 to simplify the sampling strategy. To select the set of N sample points with which to build a surrogate model, we use Latin hypercube sampling (Santner et al., 2003). This strategy samples all regions of the design space equally, making it especially suitable for modelling with computer analysis code (Santner et al., 2003). For this case study we used 200 sample points in the Latin hypercube, these samples are illustrated in Figure 4.



**Figure 4:** Samples of the input parameters used to build the surrogate models.

## RESULTS

### Comparison of model types

All of the surrogate models were built using the SUrogate MOdeling (SUMO) MATLAB toolbox (Gorissen et al., 2009c), a plug-in based adaptive tool that automatically generates a surrogate model within the predefined accuracy and time limits set by the user. Different plug-ins are supported: model types (ANN, LS-SVM, RBF, Kriging, etc), model parameter optimization algorithms (BFGS, EGO, Genetic Algorithm, etc), sample selection (random, error based, etc), and sample evaluation methods (data file, local, cluster, etc). For the LS-SVM models SUMO uses the implementation from the LS-SVMlab toolbox (<http://www.esat.kuleuven.be/sista/lssvmlab>) from Katholieke Universiteit Leuven (Suykens et al., 2002)

For the work described in this paper the toolbox was used with the following control flow: a set of samples were chosen as described in the Sampling strategy section. The simulator (CFD model) was run for each of the chosen sample points to generate the model output for each of these input points. Based upon this set, one or more surrogate models of the chosen type (e.g. ANN, LS-SVM and RBF) are constructed and their parameters optimized using a Genetic Algorithm (GA). Models are assigned a score based on an equal weighting of two measures (e.g. cross-validation and LRM). The score for each model is used as the fitness to drive the GA to a good optimum in the model parameter optimization landscape. The optimization continues until one of the following three conditions is satisfied: (1) no further improvement is possible, (2) the maximum allowed time has been reached, or (3) the user required accuracy has been met. The maximum time allowed for the GA was 1 hour. In all cases this was longer than necessary since the best solution was found quite fast.

The GA as implemented in the MATLAB GADS toolbox was used to search the parameter space of possible models. Each model type has a set of hyper-parameters, the parameter space searched by the GA for each of the model types was:

- RBF – optimal basis function combination and the parameters of each selected basis function.
- ANN – the initial weights (from which to start the back propagation training) and the network topology (number of neurons per hidden layer and number of hidden layers).
- LS-SVM – the regularization constant ( $\gamma$ ) and the spread of the RBF kernel.

The population size of each model type is set to 10 as is the maximum number of generations. The cross over fraction was set to 0.7.

The error function used to measure the fitness of each model is defined as:

$$Error(y, \tilde{y}) = \frac{\sum_{i=1}^N |y_i - \tilde{y}_i|}{\sum_{i=1}^N |y_i - \bar{y}|} \quad (10)$$

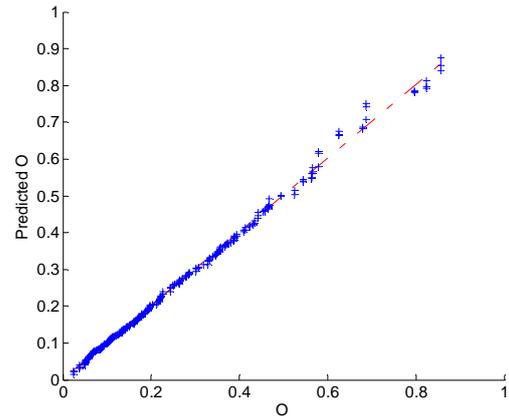
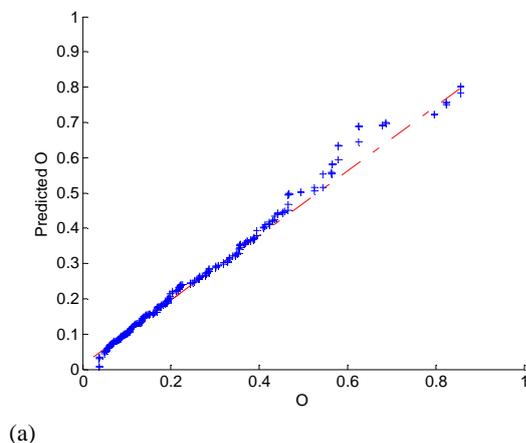
where  $y_i$ ,  $\tilde{y}_i$ ,  $\bar{y}$  are the true, predicted and mean true response values respectively. For the RBF and LS-SVM

models the metric used to drive the hyper-parameter optimisation was an equally weighted combination of the error function in (10) on a 5-fold cross-validation and the LRM measure. Cross-validation is very expensive for Neural Networks as it requires the model to be re-trained for each fold. Therefore, the ANN used an equally weighted combination of the error function in (10) calculated on the provided training samples and the LRM measure to drive the hyper-parameter optimisation.

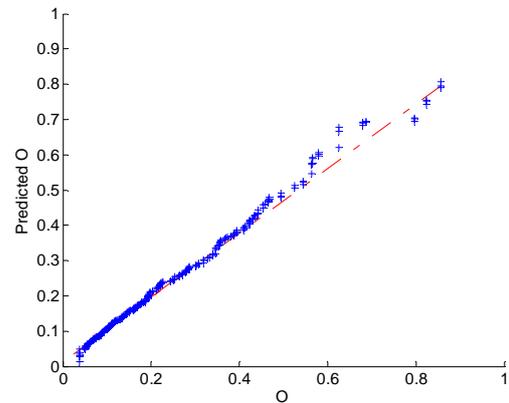
The plots in Figure 5 show that the predictions tend to be poorest when the flocculant loss is largest. This is possibly due to insufficient sample points with a high value of flocculant loss to build the models. This is a drawback of using a fixed number of sample points to build the surrogate models. Future work will look at using a small number of initial sample points followed by an adaptive sampling strategy.

From Figure 5 it can also be seen that the RBF and LS-SVM models have a similar level of accuracy at predicting the flocculant loss when the value of flocculant loss is high. Compared to the RBF and LS-SVM models the ANN model does the best job at predicting the flocculant loss for high values.

Once the surrogate models have been created they can be used for many purposes, including visualisation of the relationship of models inputs to output. One such method of visualisation is the generation of 2D contour slice plots. In these plots the contour variable is the model output and the axes represent different model inputs. Figure 6 shows a comparison of the contour slice plots for each of the surrogate model types for the input parameters A and C. The third input parameter B is held at a constant value for each of the slices shown in the figure. At low values of B (i.e. high in the feedwell) all three models show output. As the slice plane moves closer to the feedwell outlet the response (magnitude) of the ANN model deviates from the other two models. The contours of the flocculant loss remain similar just the magnitude of the ANN output becomes higher. Expert knowledge of the underlying application allows us to judge that the ANN model is giving a more accurate response at high B values than the RBF and LS-SVM models. It is assumed that the deviation between the three models at higher values of input variable B result from a low number of sample points in this region.



(b)



(c)

**Figure 5:** Cross-validated predictions versus actual values for (a) RBF, (b) ANN and (c) LS-SVM models.

### Global sensitivity analysis

While it is well established that variations in flocculant sparge location can have a significant impact on feedwell flocculant mixing and adsorption (Kahane et al., 2002; Owen et al, 2009), these sensitivities are seldom well quantified. Once a surrogate is available, these sensitivities can be computed using Extended FAST. The computation of the terms in (8) for each input variable needs the evaluation of the surrogate model at a number of points (sample size). Three sample sizes (2500, 5000 and 12500) were used in Extended FAST for the evaluation of the indices. Evaluation time for the large sample size took approximately 10 minutes for each of the model types. For each model type, very little variation in the calculated indices was found, with the mean values being reported in Table 1. The sensitivity indices for each of the model types are very similar in value for both the main effect and total effect for each of the parameters. This indicates that the global response for each of the models is very similar and therefore from a sensitivity analysis point of view any of the three surrogate models can be used.

Let us take into account the results of Extended FAST. We can see that the most important factor is C, which has a main effect of approximately 35% and a total effect of 82%. The next most important parameter is B, with a main effect of 18% and a total effect of 59%. The main effect of A is negligible at 0.39%, indicating A has only

interaction effects of the flocculant loss. The total effect for A is also small at 5% indicating the interactions are not that important.

Adding up the three sensitivity indices we can see their value is much less than 1. This means the model is non-additive with significant interactions between variables. The majority of interaction is occurring between parameters B and C.

**Table 1:** Main effect and total sensitivity indices for each of the input variables and surrogate model types.

	RBF	ANN	LS-SVM
S <sub>A</sub>	0.0015	0.0082	0.0020
S <sub>B</sub>	0.1821	0.1730	0.1898
S <sub>C</sub>	0.3767	0.3575	0.3295
ST <sub>A</sub>	0.0375	0.0721	0.0528
ST <sub>B</sub>	0.5883	0.5739	0.6244
ST <sub>C</sub>	0.8241	0.8149	0.8048

## CONCLUSION

Three types of surrogate models (RBF, ANN and LS-SVM) have been described along with the approach to optimising their parameters. The output from a CFD model of thickener feedwells that incorporates flocculant adsorption has been used to produce surrogate models of the adsorption process within thickener feedwells. The produced surrogate models have been used to investigate the sensitivity of the model output (flocculant loss) to the various models inputs (A, B and C). The calculated sensitivity indices for each of the model types were very similar in value for both the main and total effect for each of the parameters. This indicates that the global response for each of the models is very similar despite their underlying architecture being very different.

The sum of the main effects indicated the models are non-additive with significant interactions between variables. The most important parameter is C (distance from feedwell wall to flocculant sparge), the next most important parameter is B (vertical location of flocculant sparge), with the effect of parameter A (radial angle between flocculant sparge and feed pipe) being negligible. The total sensitivity indices showed strong interaction effect for C followed by B and again minor interaction for A. Therefore most of the model output can be explained by C and B with strong interaction between these two parameters.

## ACKNOWLEDGEMENTS

The support of the Parker CRC for Integrated Hydrometallurgy Solutions (established and supported under the Australian Government's Cooperative Research Centres Program) is gratefully acknowledged.

## REFERENCES

BENGIO, Y. and CHAPADOS, N. (2003). "Extensions to metric based model selection", *Journal of Machine Learning Research*, **3**, 1209-1227.

CUKIER, R.I., FORTUIN, C.M. SHULER, K.E., PETSCHKE, A.G. and SCHAIBLY, J.H. (1973). "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory", *Journal of Chemical Physics*, **59**, 3873-3878.

CUKIER, R.I., SCHAIBLY, J.H. and SCHULER, K.E. (1975). "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III Analysis of the approximations", *Journal of Chemical Physics*, **63**, 1140-1149.

CUKIER, R.I., SCHAIBLY, J.H. and SCHULER, K.E. (1978). "Nonlinear sensitivity analysis of multi-parameter model systems", *Journal of Computational Physics*, **26**, 1-42.

FAWELL, P.D., FARROW, J.B., HEATH, A.R., NGUYEN, T.V., OWEN, A.T., PATERSON, D., RUDMAN, M., SCALES, P.J., SIMIC, K., STEPHENS, D.W., SWIFT, J.D. and USHER, S.P. (2009). "20 years of AMIRA P266 'Improving Thickeners Technology' - how has it changed the understanding of thickener performance?", Proceedings of the 12th International Seminar on Paste and Thickened Tailings, 21-24 April 2009, Vina del Mar, Chile. JEWEL, R., FOURIE, A., BARRERA, S. and WIERTZ, J. (Editors), Australian Centre for Geomechanics: Nedlands, Australia, pages 59-68.

FORESEE, F.D. and HAGAN, M.T. (1997). "Gauss-Newton Approximation to Bayesian Learning". Proceedings of the International Joint Conference on Neural Networks: June 1997.

FORRESTER, A., SOBESTER, A. and KEANE, A. (2008). *Engineering Design Via Surrogate Modelling: A Practical Guide*. Wiley.

GORISSEN, D., TOMMASI, L.D., CROMBECQ, K. and DHAENE, T. (2009a). "Sequential modelling of a low noise amplifier with neural networks and active learning," *Neural Computing and Applications*, **18**, 485-494.

GORISSEN, D., COUCKUYT, I., DHAENE, T. and De TURCK, F. (2009b). "A Generic Metric for Measuring Response Nonlinearity in Regression", *Submitted to Journal of Machine Learning Research*.

GORISSEN, D., DHAENE, T. and LAERMANS, E. (2009c). "Adaptive Response Modeling with Global Surrogates", *Submitted to Advances in Engineering Software*.

JOHNSTON, R.R.M., SCHWARTZ, P., SIMIC, K., NEWMAN, M., FARROW, J.B. and SWIFT, J.D. (1998). "Improving thickener performance at the Pasminco Hobart smelter", in: DUTRIZAC, J.E., GONZALEZ, J.A., BOLTON, G.I. and HANCOCK, P. (Editors), *Zinc and Lead Processing*, Met. Soc. CIM, 697-713.

KAHANE, R.B., SCHWARZ, M.P. and JOHNSTON, R.R.M. (1997). *International Conference on CFD in Minerals and Metal Processing and Power Generation*.

KAHANE, R., NGUYEN, T. and SCHWARZ, M.P. (2002). "CFD modelling of thickeners at Worsley Alumina Pty Ltd", *Applied Mathematical Modelling*, **26**, 281-296.

KODA, M., McRAE, G.J. and SEINFELD, J.H. (1979a). "Automatic sensitivity analysis of kinetic mechanisms", *International Journal of Chemical Kinetics*, **11**, 427-444.

KODA, M., DOGRU, A.H. and SEINFELD, J.H. (1979b). "Sensitivity analysis of partial differential equations with application to reaction and diffusion processes", *Journal of Computational Physics*, **30**, 250-258.

MacKAY, D.J.C. (1992). "Bayesian Interpolation", *Neural Computation*, **4**, 415-447.

MAHAJAN, R.L. (1999). "Neural nets for modeling, optimization and control in semiconductor

manufacturing”, Proceedings of SPIE annual meeting, July 1999, Denver, Colorado.

McRAE, G.J., TILDEN, J.W. and SEINFELD, J.H. (1982). “Global sensitivity analysis – a computational implementation of the Fourier Sensitivity Test (FAST)”, *Computers and Chemical Engineering*, **6**, 15-25.

MØLLER, M.F. (1993). “A Scaled Conjugate Gradient algorithm for Fast Supervised Learning”, *Neural Networks*, **6**, 525-533.

OWEN, A.T., NGUYEN, T.V. and FAWELL, P.D. (2009). “The effect of flocculant solution transport and addition conditions on feedwell performance in gravity thickeners”, *International Journal of Mineral Processing*, **93**, 115-127.

PIERCE, T.H. and CUKIER, R.I. (1981). “Global nonlinear sensitivity analysis using walsh functions”, *Journal of Computational Physics*, **41**, 427-443.

POWELL, M.J.D. (1987). “Radial basis functions for multivariable interpolation: A review, in Algorithms for Approximation”. Oxford, U.K.: Oxford Univ. Press, 105-210.

QUEIPO, N.V., HAFTKA, R.T., SHYY, W., GOEL, T., VAIDYANATHAN, R. and TUCKER, P.K. (2005). “Surrogate-based analysis and optimization”, *Progress in Aerospace Sciences*, **41**, 1-28.

SALTELLI, A., and BOLADO, R. (1998). “An alternative way to compute Fourier amplitude sensitivity test (FAST)”, *Computational Statistics and Data Analysis*, **26**, 445-460.

SALTELLI, A., CHAN, K. and SCOTT, E. M. (2000). *Sensitivity Analysis*, New York: Wiley.

SANTNER, T. J., WILLIAMS, B. J. and NOTZ, W.I. (2003). *The Design and Analysis of Computer Experiments*, New York: Springer-Verlag.

SCHAIBLY, J.H. and SHULER, K.E. (1973). “Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II Applications”, *Journal of Chemical Physics*, **59**, 3879-3888.

SIMPSON, T.W., TOROPOV, V., BALABANOV, V. and VIANA, F.A.C. (2008). “Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come or not”, Proceedings of the 12<sup>th</sup> AIAA/ISSMO Multidisciplinary Analysis and Optimisation Conference, 10-12 September, Victoria, British Columbia, Canada.

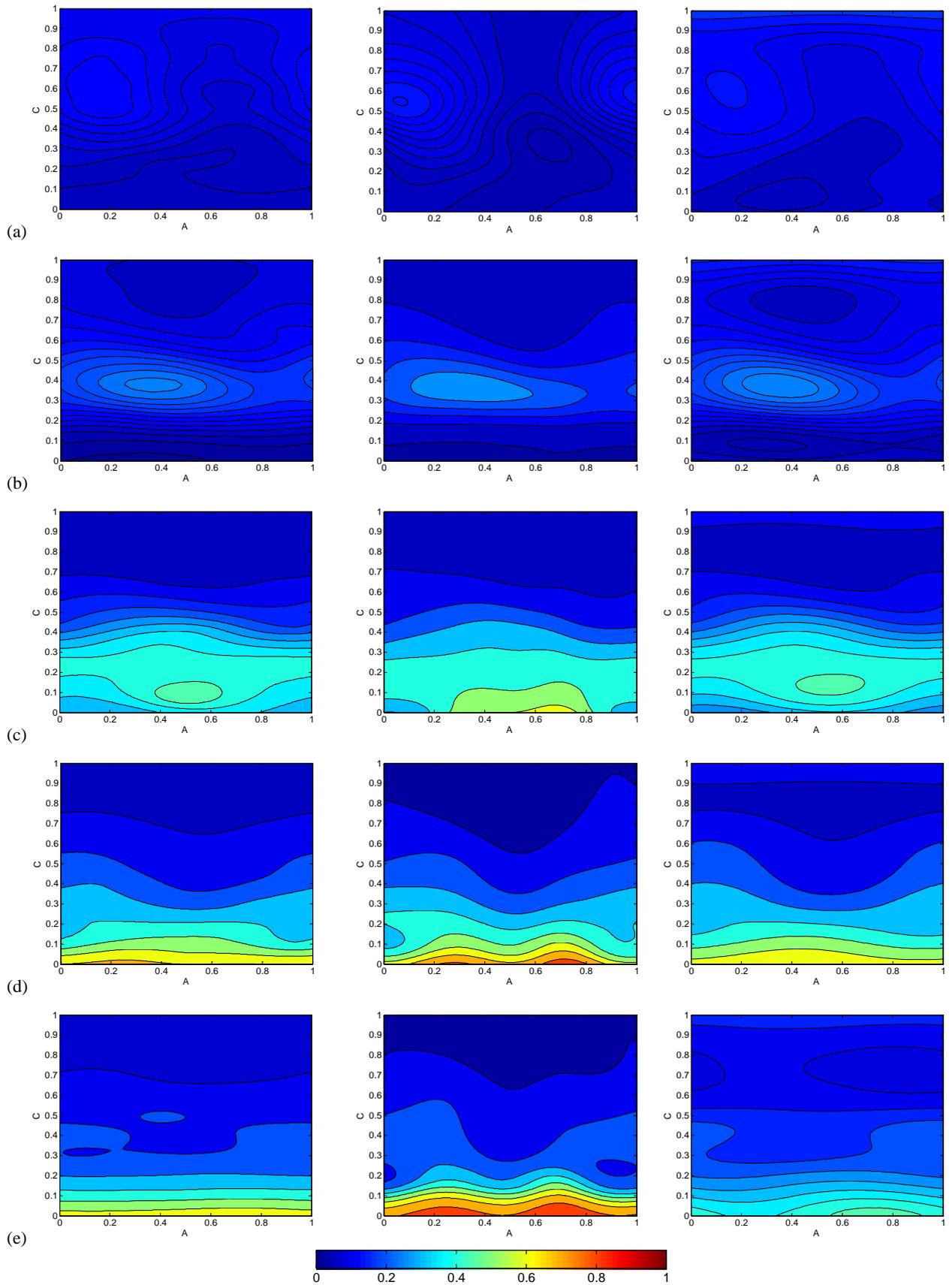
SUNDARAM, M., PALANISWAMI, S.R. and SINICKAS, M. (2005). “Radar Localization with multiple Unmanned Aerial Vehicles using Support Vector Regression”, Proceedings of the 2005 3<sup>rd</sup> International Conference on Intelligent Sensing and Information Processing, 14-17 December 2005, pages 232-237.

SUYKENS, J.A.K., Van GESTEL, T., De BRABANTER, J., De MORR, B. and VANDEWALLE, J. (2002) *Least Squares Support Vector Machines*, Singapore, World Scientific.

VAPNIK, V. (1999). *The Nature of Statistical Learning Theory*, New York: Springer-Verlag.

WANG, L. and LOWTHER, D.A. (2006). “Selection of Approximation Models for Electromagnetic Device Optimization”, *IEEE Transactions on Magnetics*, **42**, 1227-1230.

WHITE, R.B., SUTALO, I.D. and NGUYEN, T. (2003). “Fluid flow in thickener feedwell models”, *Minerals Engineering*, **16**, 145-150.



**Figure 6:** Comparison of surrogate model output (floculant loss) for the three model types tested for B values (a) 0, (b) 0.25, (c) 0.5, (d) 0.75 and (e) 1. Left images are for RBF, centre for ANN and right for LS-SVM models.