

## APPLICATION OF METAMODELLING TECHNIQUES FOR MECHANIZED TUNNEL SIMULATION\*

KAVAN KHALEDI, TOM SCHANZ

*Chair of Foundation Engineering, Soil and Rock Mechanics,  
Ruhr Universität Bochum, Germany,  
e-mails: kavan.khaledi@rub.de, tom.schanz@rub.de*

SHORASH MIRO

*Chair of Computing in Engineering,  
Ruhr Universität Bochum, Germany,  
e-mail: shorash.miro@rub.de*

[Received 19 November 2013. Accepted 16 December 2013]

**ABSTRACT.** Complex engineering problems require using computationally expensive simulations which take relatively long time. In such cases, routine tasks such as design optimization, parameter identification, or sensitivity analysis become impractical since they require thousands or even millions of simulations. A common practice for engineers to solve this problem is to use metamodels in place of actual simulation models. In this paper, we investigate the performance of four metamodeling approaches, namely, Response Surface Methodology, Moving Least Squares, POD-RBF, and Neighborhood Approximation considering the effect of sample size and sampling methods. Our main goal in this work is to find a reliable and robust metamodel technique in order to construct an approximated function for mechanized tunnel simulation. For this reason, a numerical study is carried out on a 3D tunnel modeled in Plaxis3D and the accuracy and robustness of the aforementioned metamodeling techniques are discussed.

**KEY WORDS:** metamodeling, radial basis functions, proper orthogonal decomposition, sampling.

### 1. Introduction

Most of the engineering problems involve a very complex mathematical model which cannot be solved by analytical approaches, therefore there is no way but using numerical simulations such as Finite Element Method (FEM) or

---

\*Corresponding author e-mail: kavan.khaledi@rub.de

Computational Fluid Dynamics (CFD). During the past decades, along with the significant progresses in computer science, the numerical simulations have been established as powerful tools used practically in all fields of engineering and science. In a number of situations, the engineering problems require using computationally expensive simulations. Therefore, routine tasks such as design optimization, parameter identification, inverse problem or sensitivity analysis become impossible since they require thousands or even millions of simulations. A common practice for engineers to solve this problem is to develop simplified models to approximate the original model with high level of accuracy. The approximated model which can capture the behaviour of the original model is called *metamodel* or *surrogate model*. In a mathematical language, “metamodel techniques involve the construction of multidimensional parametric surfaces by using expensive codes at strategically placed design locations, These computationally cheap metamodels can then be used in the place of original analysis for rapid evaluation”[8].

Generally, researches about metamodelling can be categorized into three groups: (1) papers in which a new metamodel approach has been introduced; see, e.g. [2, 6, 9], (2) papers which perform comparative study between existing metamodelling approaches; see, e.g. [4, 5, 8, 11] and papers which apply the metamodelling concept for engineering problems; see, e.g. [1, 8]. In this paper, first, a brief introduction to the RSM, POD-RBF, MLS and NA approaches is presented. Second, a comparative study is performed, that is, the selected metamodelling methods are applied to a real-world engineering problem to demonstrate their effectiveness and robustness.

## 2. Metamodelling approaches

Assume that we look for an unknown function  $u$  which describes the behaviour of an engineering problem. The only information that we have are the input and output data in the form of some scattered sample points like  $(\mathbf{x}, u(\mathbf{x}))$  obtained from physical or computational experiments. Depending on the structure of input parameters  $(\mathbf{x})$  and output values  $(u(\mathbf{x}))$ , several techniques for approximating  $u$  may be applicable. In this research, four metamodelling approaches, including *response surface methodology*, *POD with RBF method*, *moving least squares method* and *neighborhood approximation*, are selected for our comparative study.

### 2.1. Response surface methodology

Response surface methodology (RSM) is a technique of producing a metamodel using low order polynomials (linear, quadratic or cubic) in a rela-

tively small region of parameter space. This method is also known as **Multivariate polynomial method**. Using higher order polynomials (more than three) for big parameter spaces may improve the quality of approximation but it may increase instability of equation system and it would be too difficult to take enough sample points to determine all the coefficients in the polynomial equation [4]. In this research, we use second order polynomials for interpolation of input data. First, the approximation function is assumed to be a quadratic polynomial function, as below[8]:

$$(1) \quad \tilde{u}(\mathbf{x}) = \beta_0 + \sum_{i=1}^s \beta_i x_i + \sum_{i=1}^s \beta_{ii} x_i^2 + \sum_{i=1}^s \sum_{j>i}^s \beta_{ij} x_i x_j = \mathbf{X}^T(\mathbf{x})\boldsymbol{\beta},$$

$$(2) \quad \mathbf{X}^T(\mathbf{x}) = [1, x_1, x_2, \dots, x_s, x_1^2, x_2^2, \dots, x_s^2, x_1 x_2, x_1 x_3, \dots, x_{s-1} x_s],$$

where  $x_i$  is  $i$ -th variable of input point  $\mathbf{x}$  (vector of design variables) and  $s$  is the number of design variables or dimension of the problem. Then, the unknown coefficients  $\boldsymbol{\beta}$  are calculated using least square regression. For this reason, the following error function should be minimized:

$$(3) \quad \min_{\boldsymbol{\beta}} J(\boldsymbol{\beta}) = \sum_{i=1}^{n_p} \|\tilde{u}(\mathbf{x}^i) - u(\mathbf{x}^i)\|^2,$$

where  $\mathbf{x}^i$  is  $i$ -th sample point and  $n_p$  is the number of sample points (or supporting points). Substituting  $\boldsymbol{\beta}$  in the equation 1, the RSM metamodel is created. This approach is considered to be the simplest and most popular approach to metamodelling.

**2.2. POD with RBF** Radial basis functions are used for scattered multidimensional interpolation. In this method, linear combinations of a radially symmetric function are used to approximate response function. Assume that we have  $n_p$  sample points in the  $s$  dimensional space, then the function value in an arbitrary point  $\mathbf{x}$  is equal to:

$$(4) \quad u(\mathbf{x}) = a_1 g_1(\mathbf{x}) + \dots + a_{n_p} g_{n_p}(\mathbf{x}) = \sum_{i=1}^{n_p} a_i g_i(\mathbf{x}),$$

where  $a_i$  are unknown coefficients and  $g_i(\mathbf{x})$  is the value of radial function  $g$  with center point  $\mathbf{x}^i$  at point  $\mathbf{x}$ . There are various types of radial functions

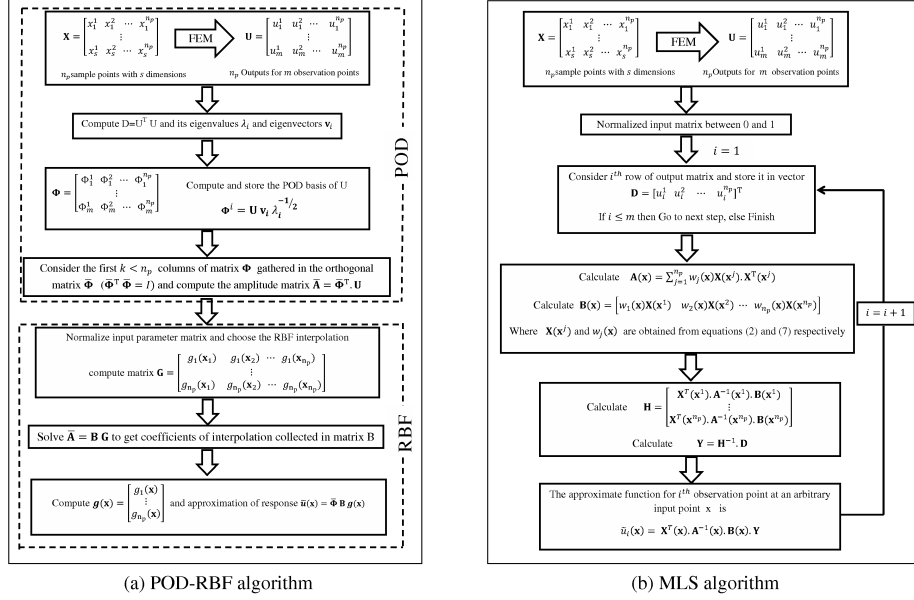


Fig. 1. Schematic flow-chart of the POD-RBF and MLS approaches

introduced in literatures (see [5, 1]). In this paper, *Inverse Multiquadratic* function is applied, which has the form:

$$(5) \quad g_i(\mathbf{x}) = \left( \|\mathbf{x} - \mathbf{x}^i\|^2 + c^2 \right)^{-0.5},$$

where  $\|\mathbf{x} - \mathbf{x}^i\|$  is euclidean distance between the points  $\mathbf{x}$  and  $\mathbf{x}^i$ . Moreover, the smoothing coefficient  $c$  has been assumed herein equal to 1. It is computationally of advantage to select this value within the 0–1 range [2]. Substituting all the sample points and their corresponding function values in equation 4 and organizing the unknowns in a linear equation system, the unknown coefficients  $a_i$  can be obtained. Proper Orthogonal Decomposition (POD) combined with Radial Basis Functions(RBF) is a recently developed method by Vladimir Buljak [1, 2]. He used the concept of orthogonality in vector spaces to reduce the dimension of problem and to simplify the approximation procedure of discrete function values. In other word, POD-RBF finds the projection of the system response into a reduced space and then the approximation is carried out by using radial basis functions. The Schematic flow-chart of POD-RBF method has been depicted in Fig. 1(a).

### 2.3. Moving least squares

The moving least squares method (MLS) uses the same logic as response surface methodology but the coefficients  $\beta$  are not constant any more, that means, they are functions of the spatial coordinates. These coefficients can be determined if the number of sample points used for interpolation is equal to the number of coefficients. In MLS method, the unknown coefficients  $\beta$  are determined by minimizing the weighted least squares error  $J(\mathbf{x})$  at the  $n_p$  sampling points[5]:

$$(6) \quad J(\mathbf{x}) = \sum_{i=1}^{n_p} w_i(\mathbf{x}) \|\mathbf{X}^T(\mathbf{x}^i) \cdot \beta(\mathbf{x}) - u(\mathbf{x}^i)\|^2,$$

where  $w_i(\mathbf{x})$  is the weight function associated with  $i^{th}$  sample point. Different types of weighting functions can be found in literatures (see [5, 6]). One of the most common types is the *Gaussian* weighting function of exponential type, which is given as:

$$(7) \quad w_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}^i\|^2}{h^2}\right),$$

where  $h$  is a spacing parameter, which can affect the accuracy and efficiency of approximation. It should be mentioned here that the MLS method satisfies a condition in which the real function,  $u$ , and the approximation,  $\tilde{u}$ , are equal at all the  $n_p$  sample points. The Schematic flow-chart of MLS method has been depicted in Fig. 1(b).

### 2.4. Neighborhood approximation

Having a set of  $n_p$  sample points in model space for which the function values have been determined, we use the geometrical construct as the *Voronoi* diagram. This is a unique way of dividing the  $s$ -dimensional model space into  $n_p$  regions (convex polyhedra) which we call Voronoi cells(see [9]). Each cell is the nearest neighbor region about one of the sample points as measured by a particular distance measure ( in our case we used euclidean distance). The Voronoi cells can be generated by '*Voronoi*' function in MATLAB software. Since the function value is known at all sample points, the neighbourhood approximation is generated by simply setting the function value to a constant inside each cell. Therefore to obtain an approximate function value at any new point we need only to find which of the samples it is closest to.

In this research, we used MATLAB 2010b to make NA metamodels in  $s$  dimensional parameter space where the sample points are irregularly distributed.

### 3. Sampling strategies

The accuracy of a metamodel strongly depends on the number of sample points and their distribution inside the input parameter space. Due to the time and cost limitations, we have to generate a limited number of samples. Because of this reason, *sampling strategies* are used to find the best sample points inside the design domain. The general idea of a sampling strategy is to generate a series of points that can be uniformly scattered in the input parameter space[11]. In this work, we applied the following strategies in order to study their effect on the accuracy of approximation methods.

- *Hammersley Sequence Sampling* (HSS) : This method is based on the representation of a decimal number in the inverse radix format [8]. It is a non-random method that generates high-uniform samples in all dimensions. The algorithm of the method can be found in ref. [3].
- *Latin Hypercube Sampling* (LHS) : This method has a random nature and the generated samples are uniform if each dimension is viewed separately. The algorithm of the method can be found in ref. [3, 8].
- *Pseudo-Random Sampling* (RND): MATLAB software has been used to generate Pseudo-Random points.

### 4. Numerical example

In July 2010 a new collaborative research center (SFB 837) started at Ruhr-Universität Bochum Germany, entitled “Interaction Models in Mechanized Tunneling”. The center consists of 14 sub-projects and it is funded by the German Research Foundation (DFG) [7]. This paper is related to the work conducted in the subproject C2 “Methods of System Identification for the Adaptation of Numerical Simulation Models”.

The goal in this section is to make a metamodel for mechanized tunneling simulation in order to apply it later for system identification purposes. To achieve this goal, a three dimensional finite element model for the tunnel excavation has been made, using the FE-code PLAXIS 3D, version 2010. Since the geometry, the material properties, the initial and excavation conditions are in total symmetric with respect to a vertical plane parallel to the tunnel axis (X-axis), only one-half of the model needs to be analyzed (see Fig. 2). More details about the model and the applied material properties can be found in [7, 10]. There are four parameters related to the soil properties which are considered as input variables of the model (i.e.,  $\phi$ ,  $c$ ,  $E_{50}^{ref}$  and  $E_{ur}^{ref}$ ) and two observation points i.e.,  $O_{12}$  and  $S_{12}$  have been set for recording the output values (see Fig. 2). Here, a comparative study is performed to evaluate the performance

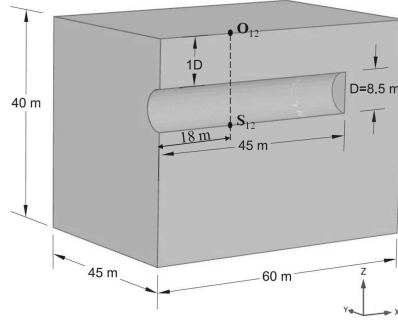


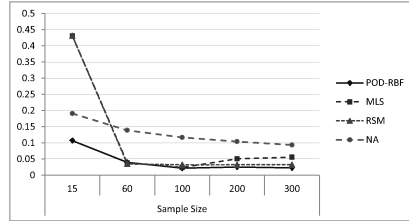
Fig. 2. The geometry of the three dimensional tunnel modeled in Plaxis 3D

of each metamodelling approach. The six step procedure for performing the comparison is as follows:

1. Generate  $n_p$  sample points inside the input parameter domain and store them in a matrix called *input parameter* matrix.
2. Run Plaxis3D for each sample point and store the vertical displacement of the observation points in a matrix called *snapshot* matrix.
3. Normalize input parameter matrix between 0 and 1 in order to avoid potential scaling errors due to disparate magnitudes of input parameters. Then, construct metamodels  $\tilde{u}$  using RSM, POD-RBF, MLS and NA.
4. Perform all the above-mentioned steps for  $n_p = 15, 60, 100, 200, 300$  points and sampling methods HSS, LHS and RND. Thus,  $4(\text{methods}) \times 3(\text{sampling techniques}) \times 5(\text{sample sizes}) = 60$  metamodels are constructed.
5. Construct a test sample of 300 points randomly in the design domain. Evaluate the exact output value  $u$  and the metamodel  $\tilde{u}$  at each test point.
6. Compute Normalized Root Mean Square Error (NRMSE):

$$\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n_p} (u(\mathbf{x}^i) - \tilde{u}(\mathbf{x}^i))^2}{\sum_{i=1}^{n_p} (u(\mathbf{x}^i))^2}}.$$

All the aforementioned steps have been performed and the obtained results have been reported in Figs 3–5.

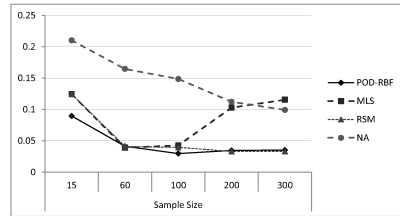


(a)

Sample Size	Method			
	POD-RBF	MLS	RSM	NA
15	0.106804	0.430685	0.430685	0.190719
60	0.039572	0.037966	0.034974	0.138657
100	0.021826	0.023705	0.032178	0.116844
200	0.025195	0.05067	0.03232	0.103806
300	0.022712	0.055796	0.03222	0.093443

(b)

Fig. 3. NRMSE error comparison for different sample sizes using HSS sampling technique

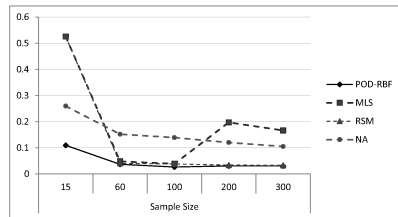


(a)

Sample Size	Method			
	POD-RBF	MLS	RSM	NA
15	0.089584	0.124431	0.124431	0.210201
60	0.04131	0.039012	0.040845	0.164587
100	0.029589	0.042731	0.039537	0.1486
200	0.034583	0.10277	0.033093	0.112219
300	0.03526	0.115479	0.033056	0.09917

(b)

Fig. 4. NRMSE error comparison for different sample sizes using LHS sampling technique



(a)

Sample Size	Method			
	POD-RBF	MLS	RSM	NA
15	0.109269	0.525172	0.525172	0.259321
60	0.036579	0.047742	0.040103	0.151724
100	0.026292	0.038479	0.037902	0.138771
200	0.030413	0.197078	0.033711	0.119802
300	0.0299	0.166175	0.032878	0.104774

(b)

Fig. 5. NRMSE error comparison for different sample sizes using RND sampling technique

## 5. Results and discussion

In this section, we discuss the results obtained after constructing meta-models using the comparison procedure described before. Regarding the POD-RBF and RSM methods, the obtained results are very close to each other for the sample size more than 60 samples. When the sample size exceeds 100 samples, the error values remain constant, that means, we do not get benefit from increasing the sample size. For samples less than 60, the POD-RBF has significantly better performances compared to the other methods. The POD-RBF, RSM and NA do not show a serious sensitivity against sampling methods where the MLS method shows a different behaviour. Regarding MLS approach, its performance is very close to the RSM method when the HSS technique is used, but, due to its sensitivity against sampling methods, we obtain inaccurate results for LH and RND methods specially for the high sample sizes. As for NA method, it is not a good choice for metamodelling and the calculated results are not as accurate as what we obtained from other approaches. Finally, the POD-RBF approach can be selected as the most efficient and robust method, because it generates very accurate metamodels even for low sample sizes and its performance is not influenced by sampling strategies. Moreover, the generated metamodel gives practically the same result in about 9000 times shorter computing time compared with the original model.

## 6. Conclusion

In this paper, application of four metamodelling approaches were examined for a 3D tunnel model simulated in Plaxis3D. A comparative study was performed considering different sample sizes and sampling strategies. Finally, the POD-RBF method was selected as the most efficient and robust approach due to its accuracy and reliability.

## REFERENCES

- [1] BOLZON, G., V. BULJAK. An Effective Computational Tool for Parametric Studies and Identification Problems in Materials Mechanics. *Comput. Mech.*, **48** (2011), 657–687.
- [2] BULJAK, V. Proper Orthogonal Decomposition and Radial Basis Functions Algorithm for Diagnostic Procedure Based on Inverse Analysis. *FME Transactions*, **38** (2010), 129–136.

- [3] GIUNTA, A., S. WOJTKIEWICZ, M. ELDRED. Overview of Modern Design of Experiments Methods for Computational Simulations, 41st Aerospace Sciences Meeting and Exhibit, 6–9 January, Nevada, Reno.
- [4] JIN, R., W. CHEN, T. W. SIMPSON. Comparative Studies of Metamodelling Techniques under Multiple Modelling Criteria, *Struct Multidisc Optim* **23**, 1–13.
- [5] KRISHNAMURTHY, T. Comparison of Response Surface Construction Methods for Derivative Estimation Using Moving Least Squares, Kriging and Radial Basis Functions, 46th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 18–21 April 2005, Texas, Austin.
- [6] LANCASTER, P., K. SALKAUSKAS. Surfaces Generated by Moving Least Squares Methods. *Mathematics of Computation*, **37** (1981), No. 155, 141–158.
- [7] MIRO, SH., V. ZAREV, D. HARTMANN, T. SCHANZ. Scenario-Driven System Identification for the Specification of Ground Models in Mechanized Tunneling, Ruhr-Universität Bochum, DFG Collaborative Research Center SFB 837 “Interaction Modelling in Mechanized Tunneling”, Germany.
- [8] MULLER, A., A. MESSAC. Metamodelling using Extended Radial Basis Functions: a Comparative Approach. *Engineering with Computers*, **21** (2006), 203–217.
- [9] SAMBRIDGE, M. Geophysical Inversion with a Neighbourhood Algorithm-I. Searching a Parameter Space. *Geophys. J. Int.*, **138** (1999), 479–494.
- [10] ZAREV, V., T. SCHANZ, M. DATCHEVA. Three-Dimensional Finite-Element Analysis of Mechanized Shield Tunneling in Urban Areas, In: Proceedings of the Thirteenth International Conference on Civil, Structural and Environmental Engineering Computing, (Eds B. H. V. Topping, Y. Tsompanakis), UK, Stirlingshire, Civil-Comp Press, Paper 223, 2011, doi:10.4203/ccp.96.223.
- [11] ZHAO, D., D. XUE. A Comparative Study of Metamodelling Methods Considering Sample Quality Merits. *Struct. Multidisc. Optim.*, **42** (2010), 923–938.