Interdisciplinary Description of Complex Systems 21(3), 316-323, 2023

REVIEW OF DACE-KRIGING METAMODEL

Muzaffer Balaban*

Turkish Statistical Institute
Ankara, Turkey

DOI: 10.7906/indecs.21.3.8
Regular article

Received: 12 June 2023.
Accepted: 28 June 2023.

ABSTRACT

This paper presents a conceptual review of the kriging metamodel that is introduced for the design and analysis of computer experiments (DACE). Kriging is a statistical interpolation method to build an approximation model from a set of evaluations of the function at a finite set of points. The method originally developed for geostatistics, and it is now widely used in the domains of spatial data analysis and computer experiments analysis. The main difference between these domains the dimensionality of the problems. Geostatistics and spatial data are mainly deal with the coordinates. Computer experiments, simulation outputs and other engineering problems have multidimensional input variables. With this study, it is aimed to examine the limitations of the prediction performance of the DACE-kriging metamodel. The result of the study shows that the regression part of the DACE-kriging metamodel is the most important part to develop an approximation, and if there is a spatial relationship of the residuals, kriging part will also contribute to the improvement of the prediction performance. Otherwise, kriging will have no contribution to the DACE-kriging metamodel, and even worsen the prediction performance. If the regression part perfectly fit to the observations, the residual will have poor spatial relationship and the kriging part will be meaningless anymore.

KEY WORDS

DACE-kriging, regression, basic kriging, correlogram

CLASSIFICATION

JEL: C15

*Corresponding author, η: balabanmuzaffer@gmail.com; +90 535 507 4252
Turkish Statistical Institute, 06100, Ankara, Turkey

INTRODUCTION

Kriging term covers several spatial interpolation models. Kriging theory was originally developed as a geostatistical interpolation method [1]. The kriging model makes predictions at unobserved locations using a linearly weighted combination of observations. Each observation influences the kriging prediction is based on geographical proximity to the unobserved location, the spatial spreads and the pattern of spatial correlation of the observations. Kriging models are meaningful only if the observations are spatially correlated. The kriging weights are recalculated using the appropriate variogram or correlogram model for each prediction point. There are many kinds of kriging in the literature such as simple kriging, ordinary kriging, universal kriging, cokriging, median polish kriging etc. [2].

Sacks et al. [3] presented a modified kriging approach as a metamodel to deterministic computer experiments. The use of kriging metamodels has been remarkably effective for global metamodeling in the design and analysis of computer experiments (DACE) community when the simulation models are complex and/or very expensive to run [4]. Their approach is a hybrid method that combines a regression between the output variable and input variables with the simple kriging (SK) of the regression residuals. Firstly, a polynomial regression model is applied to the outputs and then basic kriging applied to the residuals. Their main contribution to the kriging literature is expanding of the problem dimensions from two-dimensional coordinate to the high dimensional computer experiments. Additionally, they have used high dimensional correlogram models instead of variogram mostly used in geostatistics to find the kriging weights. Prediction at each new point is performed by summing the predicted trend and residual. The parameter set used for the regression part are estimated once for the whole search space, and for the kriging part the weights are re-estimated at each new point. In the literature it is known that Regression models are local and kriging models are global. Kriging models are flexible because of the diversity of the correlogram model obtained from the experiments. Therefore, it reveals the importance of the kriging part to develop a global metamodel for the whole search space.

There are several names of this method in the DACE literature. Some of them are as follows. Kriging [3, 5-16], spatial correlation metamodels [8, 17, 18], Gaussian process models [4, 19, 20], Gaussian stochastic process models [21, 22], Gaussian kriging [21] are used as the name of the method in the related references. In the geostatistical literature this method is called regression kriging [23, 24]. Some authors variously call this method as regression with residual simple kriging [25], detrended kriging [26, 27] and residual kriging [28, 29]. I prefer to use “DACE-Kriging” in the metamodeling process as the same meaning with the “regression kriging” in Geostatistics as the name for this method to prevent some misunderstanding because the kriging term refer to a general class of geostatistical interpolation methods.

The aim of this study is to examine the limitations of the prediction performance of the DACE-kriging metamodel. The results of this study show that the regression part of the DACE-Kriging model is the most important part to develop an approximation, and if there is a spatial relationship of the residuals, kriging part will also contribute to the improvement of the prediction performance. Otherwise, kriging will have no contribution to the DACE-Kriging model, and even worsen the prediction performance. If the regression part perfectly fit to the observations, the residual will have poor spatial relationship and the kriging part will be meaningless anymore.

Remaining of this article as follows. Model formulation of DACE-Kriging metamodel is presented in Section 2, numerical examples are given in Section 3, and Section 4 presents conclusions.

MODEL FORMULATION OF DACE-KRIGING METAMODEL

DACE-Kriging metamodel is a mixed estimation method that is a combination of multiple regression methods and simple kriging. It can be defined as the estimation of residual values obtained from the difference between the estimation values made by methods such as regression and the observation values by kriging method [3]. Simply, prediction at each new point is performed by summing the predicted trend and residual. Predicted trend is obtained by linear or quadratic regression (or higher order) and predicted residual is obtained by simple kriging applied to regression residuals.

Model assumptions of $\mathrm{Y}(\mathbf{x})$ are given in the followgin equations:

$$\mathrm{Y}(\mathbf{x}) = \mathrm{M}(\mathbf{x}) + \mathrm{Z}(\mathbf{x}), \tag{1}$$

$$\mathrm{M}(\mathbf{x}) = \sum_{j=0}^{b} \beta_j f_j(\mathbf{x}), \tag{2}$$

$$\mathrm{Z}(\mathbf{x}) = \mathrm{Y}(\mathbf{x}) - \mathrm{M}(\mathbf{x}), \tag{3}$$

$$\mathrm{Z}(\mathbf{x}) = \mathrm{Y}(\mathbf{x}) - \sum_{j=0}^{b} \beta_j f_j(\mathbf{x}). \tag{4}$$

DACE-Kriging predictor is given in the following two equations:

$$\hat{y}(\mathbf{x_0}) = \sum_{j=0}^{b} \beta_j f_j(\mathbf{x_0}) + \sum_{i}^{n} \lambda_i\, z(\mathbf{x_i}), \tag{5}$$

$$\hat{y}(\mathbf{x_0}) = \sum_{j=0}^{b} \beta_j f_j(\mathbf{x_0}) + \sum_{i}^{n} \lambda_i \left(y(\mathbf{x_i}) - \sum_{j=0}^{b} \beta_j f_j(\mathbf{x_i}) \right). \tag{6}$$

The first part of the model in (5) and (6) shows the regression model and the second part shows the simple kriging model. Where, $\mathbf{x_0}$ is a new point vector for prediction, $\mathbf{x_i}$ is the observed point vector, β_j is the j. coefficients of the regression model, f_j indicates j. regression design unit and f_0 is equal to 1. DACE-Kriging predictor can be rewritten as a vector form in (7) and (8).

$$\hat{y}(\mathbf{x_0}) = \mathbf{f_0'}\hat{\boldsymbol{\beta}} + \boldsymbol{\lambda}'\mathbf{Z}, \tag{7}$$

$$\hat{y}(\mathbf{x_0}) = \mathbf{f_0'}\hat{\boldsymbol{\beta}} + \boldsymbol{\lambda}'(\mathbf{y} - \mathbf{f}'\hat{\boldsymbol{\beta}}). \tag{8}$$

Where, $\mathbf{Z}$ is residual vector, $\mathbf{f_0}$ is design vector of input variables at $\mathbf{x_0}$, $\hat{\boldsymbol{\beta}}$ is regression model parameters vector and $\boldsymbol{\lambda}$ is the kriging weights vector. Considering the spatial correlation of the residuals, the model coefficients are solved with the following generalized least squares estimator [2].

$$\hat{\boldsymbol{\beta}} = (\mathbf{f}'\mathrm{R}^{-1}\mathbf{f})^{-\mathbf{1}}\mathbf{f}'\mathrm{R}^{-1}\mathbf{y}, \tag{9}$$

$$\boldsymbol{\lambda} = \mathrm{R}^{-1}\mathbf{r}. \tag{10}$$

Where, **f** is the input variables design matrix at the observation point, **Y** is the observation vector, **r** is the correlogram vector of new point, and R is the nxn-dimensional correlogram matrix of the residuals. R and **r** are obtained from correlogram model.

$$\mathrm{R} = \begin{vmatrix} r(x_1 - x_1) & \ldots.. & r(x_1 - x_n) \\ \ldots. & \ldots.. & \ldots. \\ r(x_n - x_1) & \ldots.. & r(x_n - x_n) \end{vmatrix}$$

The covariogram is estimated with the following equation:

$$\hat{c}(\mathbf{h}) = \frac{\mathbf{1}}{\mathbf{n(h)}} \sum_{\mathbf{i=1}}^{\mathbf{n(h)}} (\mathbf{Z}(\mathbf{x_i}) - \boldsymbol{\mu})(\mathbf{Z}(\mathbf{x_i} + \mathbf{h}) - \boldsymbol{\mu}), \tag{11}$$

where $\hat{c}(h)$ is the covariogram estimator, μ is the mean of the stochastic process and $\hat{c}(0) = \sigma^2$ is the variance of the stochastic process, $n(h)$ is the number of experiment pairs. The correlogram is estimated as in (12).

$$\hat{r}(h) = \hat{c}(h)/\sigma^2, \tag{12}$$

where $\hat{r}(h)$ is the correlogram estimator. A theoretical correlogram is used to calculate the kriging weights for each point. Theoretical correlogram model must fit to the experimental correlogram. General theoretical correlogram model is given in (13). Where, θ_i is the correlogram model parameter, h_i is the univariate distance and p_i is the power of the model valued one or two [3, 30-32]. Estimation of parameters in (13) is realized by maximum likelihood estimation method (MLE) or least squares estimator. Some mostly used theoretical correlogram models regarded Euclidean norm are given in Table 1.

$$r(h) = \prod_{i=1}^{k} \exp\left(-(\theta_i h_i)^{p_i}\right). \tag{13}$$

Table 1. Theoretical correlogram models.

Model name	**Model**
Gaussian	$r(h) = \exp\left(-\left(h/\theta\right)^2\right)$
Exponential	$r(h) = \exp\left(-h/\theta\right)$
Linear	$r(h) = \max\ (1-\theta h, 0)$
Mathern 1	$r(h) = \exp\left(1-\theta h\right)\left(1+\theta h+\theta^2 h^2/3\right)$
Mathern 2	$r(h) = \exp\left(-\theta h\right)\left(1+\theta h\right)$

NUMERICAL EXAMPLES

Numerical examples are performed on the four test problems which are Six-hump camel back function, Perm function, Stablinski–Tang function and Quintic function given in the Table 2. LHD is the one of the popular experimental design methods for computer experiments since developed by Mc Kay et al. [33]. It is convenient both for kriging and regression because of the space filling property. The levels of each factor are included in the design once. All factors have the same number of levels. Experiments are designed as many as the number of levels. In this design, the permutation of the levels is determined randomly. Training and validation datasets are generated with LHD. Training datasets consisting of 20, 27, 45 and 72 experiment were generated for the test problems (Six-hump camel back function, Perm function, Stablinski–Tang function, Quintic function). Since the kriging models are the best unbiased linear estimators, the validation data set consisting of 500 experiments was generated for each test problem. The validation is assessed by standard accuracy measures. The measures used in this study are Root Mean Squared Error (RMSE) and R^2 given as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y(\mathbf{x_i}) - \hat{y}(\mathbf{x_i})\right)^2}, \tag{14}$$

$$R^2 = 1 - \sum_{i=1}^{n}\left(y(\mathbf{x_i}) - \hat{y}(\mathbf{x_i})\right)^2 / \sum_{i=1}^{n}\left(y(\mathbf{x_i}) - \bar{y}\right)^2. \tag{15}$$

Table 2. Test functions.

No.	Function Name	Test Functions
1	**Six-hump camel back**	$$f(x) = \left(4 - 2.1x_1^2 + \frac{1}{3}x_1^4\right)x_1^2 + x_1x_2 + (-4 + 4x_2^2)^2x_2^2$$ $$-3 \leq x_1 \leq 3, -2 \leq x_2 \leq 2$$
2	**Perm**	$$f(x) = \sum_{i=1}^{3}(\sum_{j=1}^{3}(j + 10)(x_j^i - \frac{1}{j^i}))^2$$ $-3 \leq x_i \leq 3$
3	**Styblinski–Tang**	$$f(x) = 0.5\sum_{i=1}^{5}(x_i^4 - 16x_i^2 + 5x_i)$$ $-5 \leq x_i \leq 5$
4	**Quintic**	$$f(x) = \sum_{i=1}^{8}\left\lvert x_i^5 - 3x_i^4 + 4x_i^3 + 2x_i^2 - 10x_i - 4\right\rvert$$ $$-10 \leq x_i \leq 10$$

Linear regression and quadratic regression parameters were estimated with (9) for four numerical examples. Residual values were calculated with (4). By applying simple kriging to the residuals as given by (5), estimation errors for validation points were obtained.

Table 3. RMSE of the metamodels.

No.	Function Name	LR	QR	LR+SK	QR+SK
1	**Six-hump camel back**	31,5011	16,1573	19,94	13,8584
2	**Perm**	87761,27	44295,82	66390,76	44085,06
3	**Styblinski–Tang**	75,056	69,770	72,124	70,726
4	**Quintic**	79565,51	58847,17	59626,04	59437,34

LR in the Table 3 denotes linear regression and QR denotes quadratic regression. LR+SK shows simple kriging applied to linear regression residuals and QR+SK shows simple kriging applied to quadratic regression residuals. Considering the RMSE, it is seen that the residuals after LR have a spatial relationship for all test problems and the DACE-Kriging model produces more successful predictions when SK is applied to the residuals. According to RMSE, the DACE-Kriging model, as a result of the SK applied to the residuals after QR, produced more successful predictions at the validation points for the first and second test problems, and the prediction success at the validation points for the third and fourth test problems was worse due to the weak spatial relationship of the residuals.

Table 4. R^2 of the metamodels.

No.	Function Name	LR	QR	LR+SK	QR+SK
1	**Six-hump camel back**	–0,357	0,645	0,459	0,739
2	**Perm**	–0,068	0,728	0,389	0,731
3	**Styblinski–Tang**	–0,028	0,111	0,048	0,087
4	**Quintic**	0,014	0,458	0,460	0,447

According to the R^2 performance criterion given in Table 4, QR produces better predictions for all test problems than the LR and LR+SK models. QR+SK gives better prediction performance for the first and second test problems because of the spatial relationship of the residuals among all applied models, and worse for the third and fourth test problems due to the weak spatial relationship of the residuals at the validation points.

These four test problems show that the regression part of the DACE-Kriging model is the most important part of the model, and if there is a spatial relationship of the residuals, the SK model will also contribute to the improvement of the prediction performance. Otherwise, the SK will have no contribution to the DACE-Kriging model, and even worsen the prediction performance.

CONCLUSION

QR produces better predictions for all test problems than the LR and LR+SK models. QR+SK gives better prediction performance for the first and second test problems because of the spatial relationship of the residuals among all applied models, and worse for the third and fourth test problems due to the weak spatial relationship of the residuals at the validation points.

These four test problems show that the regression part of the DACE-Kriging model is the most important part of the model, and if there is a spatial relationship of the residuals, the SK model will also contribute to the improvement of the prediction performance. Otherwise, the SK will have no contribution to the DACE-Kriging model, and even worsen the prediction performance.

Future studies will focus on developing new kriging approaches to increase the prediction performance of the metamodel.

REFERENCES

[1] Matheron, G.: *Principles of geostatistics.*
Economic Geology **58**, 1246-1266, 1963,
https://dx.doi.org/10.2113/gsecongeo.58.8.1246,

[2] Cressie, N.A.C.: *Statistics for Spatial Data.*
A Wiley-Interscience publication, New York, 1993,

[3] Sacks, J., et al.: *Design and analysis of computer experiments.*
Statistical Science **4**, 409-435, 1989,
http://dx.doi.org/10.1214/ss/1177012413,

[4] Santner, T.J.; Williams B.J. and Notz, W.I.: *The Design and Analysis of Computer Experiments.* 2nd edition.
Springer, New York, 2019,

[5] Simpson, T., et al.: *On the use of statistics in design and the implications for deterministic computer experiments.*
Proceedings of DETC'97 ASME Design Engineering Technical Conferences. Sacramento, pp.1-14, 1997,

[6] Jin, R.; Chen, W. and Simpson, T.: *Comparative Studies of Metamodeling Techniques Under multiple modeling criteria.*
8th Symposium on Multidisciplinary Analysis and Optimization. AIAA, Long Beach, 2000, http://dx.doi.org/10.2514/6.2000-4801,

[7] Jones, D.R.: *A Taxonomy of Global Optimization Methods Based on Response Surfaces.*
Journal of Global Optimization **21**, 345-383, 2001,
http://dx.doi.org/10.1023/a:1012771025575,

[8] Meckesheimer, M., et al.: *Metamodeling of Combined Discrete/Continuous Responses.*
American Institute of Aeronautics and Astronautics **39**(10), 1950-1959, 2001,
http://dx.doi.org/10.1115/detc2000/dtm-14573,

[9] Srivastava, A., et al.: *A method for using legacy data for metamodel-based design of large-scale systems.*
Structural and Multidisciplinary Optimization **28**, 146-155, 2004,
http://dx.doi.org/10.1007/s00158-004-0438-4,

[10] Kim, B.S.; Lee, Y.B. and Choi, D.H.: *Comparison study on the accuracy of metamodeling technique for non-convex functions.*
Journal of Mechanical Science and Technology **23**, 1175-1181, 2009,
http://dx.doi.org/10.1007/s12206-008-1201-3,

[11] Martin, J.D.: *Using Kriging Models to Perform Sensitivity Analysis.*
American Institute of Aeronautics and Astronautics 2011-2156, 1-11, 2011,
http://dx.doi.org/10.2514/6.2011-2156,

[12] Nik, M.A., et al.: *A comparative study of metamodeling methods for the design optimization of variable stiffness composites.*
Composite Structures **107**, 494-501, 2014,
http://dx.doi.org/10.1016/j.compstruct.2013.08.023,

[13] Van Gelder, L., et al.: *Comparative study of metamodeling techniques in building energy simulation: Guidelines for practitioners.*
Simulation Modelling Practice and Theory **49**, 245-257, 2014,
http://dx.doi.org/10.1016/j.simpat.2014.10.004,

[14] Dong, H., et al.: *Kriging-assisted discrete global optimization (KDGO) for black-box problems with costly objective and constraints.*
Applied Soft Computing Journal **94**, 1-17, 2020,
http://dx.doi.org/10.1016/j.asoc.2020.106429,

[15] Zhou, C., et al.: *Reliability and sensitivity analysis of composite structures by an adaptive kriging- based approach.*
Composite Structures **278**, 1-12, 2021,
http://dx.doi.org/10.1016/j.compstruct.2021.114682,

[16] Wang, J.; Sun, Z. and Cao R.: *An efficient and robust Kriging-based method for system reliability analysis.*
Reliability Engineering and System Safety **216**, 1-19, 2021,
http://dx.doi.org/10.1016/j.ress.2021.107953,

[17] Barton, R.R.: *Metamodels for simulation input-output relation.*
In: Swain, J.J.; Goldsman, D.; Crain, R.C. and Wilson, J.R.: *Proceedings of the 24th conference on Winter simulation*. IEEE, Arlington, pp.289-299, 1992,
http://dx.doi.org/10.1145/167293.167352,

[18] Barton, R.R.: *Simulation metamodels.*
In: 1998 Winter Simulation Conference. IEEE, Washington, pp.167-174, 1998,
http://dx.doi.org/10.1109/WSC.1998.744912,

[19] Marrel, A., et al.: *Calculations Sobol indices for the Gaussian process metamodel.*
Reliability Engineering and System Safety **94**, 742-751, 2009,
http://dx.doi.org/10.1016/j.ress.2008.07.008,

[20] Schweidtmann, A.M., et al.: *Deterministic global optimization with Gaussian processes embedded.*
Mathematical Progamming Computation **13**, 553-581, 2021,
http://dx.doi.org/10.1007/s12532-021-00204-y,
[21] Shao, W., et al.: *Extended Gaussian Kriging for computer experiments in engineering design.*
Engineering with Computers **28**, 161-178, 2012,
http://dx.doi.org/10.1007/s00366-011-0229-7,
[22] Erickson, C.B.; Ankenman, B.E. and Sanchez, S.M.: *Comparison of Gaussian process modeling software.*
European Journal of Operational Research, Elsevier **266**(1), 179-192, 2018,
http://dx.doi.org/10.1016/j.ejor.2017.10.002,
[23] Odeh, I.; McBratney, A. and Chittleborough, D.: *Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging.*
Geoderma **67**(3-4), 215-226, 1995,
http://dx.doi.org/10.1016/0016-7061(95)00007-b,
[24] Hengl, T.; Heuvelink, G.B.M. and Rossiter, D.G.: *About regression-kriging: From equations to case studies.*
Computers and Geosciences **33**(10), 1301-1315, 2007,
http://dx.doi.org/10.1016/j.cageo.2007.05.001,
[25] Asli, M. and Marcotte, D.: *Comparison of approaches to spatial estimation in a bivariate context.* Mathematical Geology **27**, 641-658, 1995,
http://dx.doi.org/10.1007/bf02093905,
[26] Jef, H., et al.: *Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium.*
Journal of Environmental Monitoring **8**, 1129-1135, 2006,
http://dx.doi.org/10.1039/b612607n,
[27] Nalder, I.A. and Wein, R.W.: *Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest.*
Agricultural and Forest Meteorology **92**, 211-225, 1998,
http://dx.doi.org/10.1016/s0168-1923(98)00102-6,
[28] Mardikis, M.G.; Kalivas, D.P. and Kollias, V.J.: *Comparison of interpolation methods for the prediction of reference evapotranspiration - an application in Greece.*
Water Resources Management **19**, 251-278, 2005,
http://dx.doi.org/10.1007/s11269-005-3179-2,
[29] Reyes, A., Giraldo R. and Mateu, J.: *Residual Kriging for Functional Spatial Prediction of Salinity Curves.*
Communications in Statistics - Theory and Methods **44**(4), 798-809, 2015,
http://dx.doi.org/10.1080/03610926.2012.753087,
[30] Mitchell, T.J. and Morris, M.D.: *Bayesian design and analysis of computer experiments: two examples.*
Statistica Sinica **2**, 359-379, 1992,
[31] Kleijnen, J.P.C.: *Kriging metamodeling in simulation: a review.*
European Journal of Operational Research **192**, 707-716, 2009,
http://dx.doi.org/10.1016/j.ejor.2007.10.013,
[32] Kleijnen, J.P.C.: *Regression and kriging metamodels with their experimental designs in simulation: a review.*
European Journal of Operational Research **256**, 1-16, 2017,
http://dx.doi.org/10.1016/j.ejor.2016.06.041,
[33] Mc Kay, M.G.D.; Beckman, R.J. and Conover, W.J.: *A comparison of three methods for selecting values of input variables in the analysis of output from a computer code.*
Technometrics **21**, 239-245, 1979,
http://dx.doi.org/10.1080/00401706.1979.10489755.