2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

An Exploratory Study of Discrete Time State-Space Models using Kriging

Andres F. Hernandez and Martha Grover Gallivan

Abstract— **An exploratory study of the application of kriging as an approximated model of complex dynamic systems is presented in this paper. This technique, which has its roots in the geology community, is gaining popularity for its simplicity and ability to track the error of the reduced model, making it attractive for engineering applications. Our proposed methodology, called dynamic mapping kriging, requires the generation of representative function evaluations, and then uses them in a recursive state-space formulation to represent the evolution of the process. By demonstrating on a highly simplified model, our intention is to present the relevant elements of this methodology, focusing on the design of the function evaluations as a crucial factor in the accurate representation of the system. We illustrate the advantages that the error prediction has as a tool to improve our technique and as a measurement to bound the error dynamics, and give some final recommendations about the applications of this methodology to build approximated models.**

I. INTRODUCTION

Dynamic models range in their complexity, from simple processes with linear expressions to nonlinear systems with a large number of parameters and assumptions. As computational capacity improves, the use of complex models and simulations has become an important part of engineering. Fields in which complex dynamic simulations are needed include combustion, turbulent flow, and nanoscale phenomena. Simulating these systems for practical purposes requires the reduction in the computational effort without losing accuracy in the prediction. Our group has shown interest in the creation of approximated models to represent dynamic systems in a fast and accurate mode [1], particularly in systems represented by discrete time state-space models:

$$\mathbf{w}(k+1) = f\left[\mathbf{w}(k), \mathbf{u}(k)\right] \qquad \mathbf{w} \in \mathbb{R}^m \tag{1}$$

Other examples of approximated models for complex systems are *in-situ* adaptive tabulation (ISAT) [2] and cell mapping [3]. Unfortunately, these approaches do not have a consistent framework to predict or bound the error dynamics, but intuitively, the dynamics will depend on the initial state-space $\mathbf{w}$, the control input $\mathbf{u}$ and the final discrete time index.

An approximation method that is gaining popularity in engineering is kriging (also known as Gaussian process regression), which has been used to represent deterministic computer models in several research areas such as conceptual design, structural optimization, aerospace, and mechanical engineering [4]. Our interest in kriging is for its ability to predict a function based on local and global interpolation, its flexibility to handle multidimensional applications, and the possibility to estimate the error in the prediction. These are characteristics that most models do not have and which could be convenient for engineering applications.

Kriging had been used primarily as a technique to describe the behavior of a static system over a domain, based on the spatial distribution of sample points, but in many cases, these systems also have a temporal element [5]. The kriging method has been extended from its initial formulation to incorporate dynamic features by including time as an additional spatial variable in the model [6], or by building a space-time covariance function that can handle a time dimension separately from the state-space [7]. More relevant are the applications of kriging in the systems area, for time-series theory [8] and for discrete-time nonlinear systems [9]. These examples show the potential that this technique has to be applied in dynamic modeling and control.

The next sections present a methodology that we have developed called dynamic mapping kriging (DMK), which approximates complex dynamic systems based on a recursive space-state formulation. A brief mathematical introduction of the principal elements of kriging and the proposed approach is presented. We develop a simple case study to show the importance of the kriging elements and how these elements affect the approximation. We conclude with some recommendations for implementation and future work.

II. MODELING APPROACH

A. Mathematical Form of Kriging

Kriging was initially developed by French mathematician Georges Matheron, based on the master's thesis written by Daniel Gerhardus Krige in 1960. Krige's idea was that the predicted error of an approximated model at some point can be correlated with the distance between the point of interest and a set of sampled points in the domain.

A characteristic of kriging is the exact prediction that the approximated model makes when it is evaluated at a sample point (see Fig. 1). Moving through the domain from any sample point, the error between the kriging model and the true function increases, until the error is calculated close to another sample point, where it decreases. This behavior implies that the magnitude of the error at an unknown point is related to the distance from a measured point. This concept is not usually employed in system identification and model

Financial support for this work was provided by the Air Force Office of Scientic Research under award number FA9550-07-1-0161

Andres F. Hernandez is with the School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. andres.hernandez@chbe.gatech.edu

Martha Grover Gallivan is with the School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA. martha.gallivan@chbe.gatech.edu

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

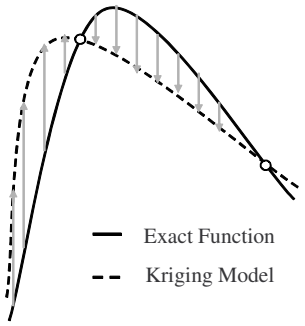

Fig. 1. Graphic representation of a Kriging model.

reduction, in which the location of the experimental points is not explicitly considered once the model had been identified.

The mathematical form of a kriging model is [10]

$$\hat{y}(\mathbf{x}) = \sum_{i=1}^{p} \beta_i h_i(\mathbf{x}) + Z(\mathbf{x}) \tag{2}$$

where $\hat{y}(\mathbf{x}) \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^d$.

The first term represents a linear combination of a basis set of p regression functions $h_i(\mathbf{x})$ with their corresponding β_i coefficients. A model that is nonlinear in the coefficients $\boldsymbol{\beta}$ could potentially be used, although this is not the common approach in kriging. The second term is modeled as a Gaussian random function that represents a correction of the regression functions. This term is characterized by

$$E[Z(\mathbf{x})] = 0 \tag{3}$$

$$Cov[Z(\mathbf{x}_i), Z(\mathbf{x}_j)] = V_{ij} = \sigma^2 \cdot R(\mathbf{x}_i, \mathbf{x}_j) \tag{4}$$

where $V \in \mathbb{R}^{n \times n}$ is the covariance matrix, σ^2 is the process variance, and $R \in \mathbb{R}^{n \times n}$ is the spatial correlation matrix for $Z(\mathbf{x})$ for the n sample points.

Suppose that there are n sample points, represented by the matrix $X \in \mathbb{R}^{n \times d}$, with an associated vector $\mathbf{Y} \in \mathbb{R}^n$ which corresponds to the measurements for X. Given these sample points, a linear predictor of the measurement for any point $\mathbf{x}$ in the domain is

$$\hat{y}(\mathbf{x}) = \boldsymbol{\lambda}^T(\mathbf{x}) \cdot \mathbf{Y} \tag{5}$$

In order to define the vector coefficient $\boldsymbol{\lambda}(\mathbf{x}) \in \mathbb{R}^n$, a minimization problem is established, where the objective function is the mean square error (MSE) of the prediction

$$MSE[\hat{y}(\mathbf{x})] = E[\boldsymbol{\lambda}^T(\mathbf{x}) \cdot \mathbf{Y} - y(\mathbf{x})]^2 \tag{6}$$

subject to the constraint

$$E[\boldsymbol{\lambda}^T(\mathbf{x}) \cdot \mathbf{Y} - y(\mathbf{x})] = 0 \tag{7}$$

The solution of this problem can be expressed using the best linear unbiased predictor with the expression [11]

$$\hat{y}(\mathbf{x}) = \mathbf{h}^T(\mathbf{x})\hat{\boldsymbol{\beta}} + \mathbf{v}^T(\mathbf{x})V^{-1}(\mathbf{Y} - H\hat{\boldsymbol{\beta}}) \tag{8}$$

where

- $\hat{\boldsymbol{\beta}} \in \mathbb{R}^p$ is the generalized least-squares estimator which corresponds to the expression

$$\hat{\boldsymbol{\beta}} = (H^T V^{-1} H)^{-1} H V^{-1} \mathbf{Y} \tag{9}$$

- $H \in \mathbb{R}^{n \times p}$ is the matrix of regression functions evaluated at the n sample points ($H_{i,j} = \mathbf{h}_j(\mathbf{x}_i)$)
- $\mathbf{h}(\mathbf{x}) \in \mathbb{R}^p$, the set of regression functions evaluated at an unknown point $\left(\mathbf{h}^T(\mathbf{x}) = [h_1(\mathbf{x}), ..., h_p(\mathbf{x})]\right)$
- $\mathbf{v}(\mathbf{x}) \in \mathbb{R}^n$, a vector that expresses the covariance of the unknown point with each sample point in X. $\left(\mathbf{v}^T(\mathbf{x}) = [Cov[Z(\mathbf{x}), Z(\mathbf{x}_1)], ..., Cov[Z(\mathbf{x}), Z(\mathbf{x}_n)]]\right)$

The expected mean square error of the prediction made by kriging (known as kriging variance) is defined as

$$MSE\left[\hat{y}(\mathbf{x})\right] = \sigma^2 - \begin{bmatrix}\mathbf{h}^T(\mathbf{x}) & \mathbf{v}^T(\mathbf{x})\end{bmatrix} \begin{bmatrix}0 & H^T \\ H & V\end{bmatrix}^{-1} \begin{bmatrix}\mathbf{h}(\mathbf{x}) \\ \mathbf{v}(\mathbf{x})\end{bmatrix} \tag{10}$$

Given that kriging returns the exact value of the true function at each sample point, the mean square error at those points is therefore equal to zero.

Based on this mathematical formulation, four elements are needed to build a kriging model:

- The location of the n sample points.
- A basis set of regression functions, which help to capture the tendencies of the function over the space.
- The spatial correlation function (SCF), which models the covariance function between points in the domain.
- The selection of the method to estimate $\hat{\boldsymbol{\beta}}, \hat{\sigma}^2$, and all parameters in the SCF to compute the kriging model and its predicted error.

B. Dynamic Mapping Kriging

An approximated model to substitute the right-hand side of (1) can be written based on the kriging structure presented in the previous section. We called this approximated model dynamic mapping kriging, (DMK) and it is represented by

$$\mathbf{w}(k+1) = \hat{f}[\mathbf{x}(k), \mathbf{Y}, X] \qquad k = 0, 1, ... \tag{11}$$

$$\mathbf{x}(k) = \begin{bmatrix}\mathbf{w}(k) & \mathbf{u}(k)\end{bmatrix}^T \tag{12}$$

$$t = k\Delta t \tag{13}$$

$$\mathbf{Y} = X(k=1) = f[X(k=0)] \tag{14}$$

where (11) is represented by the kriging model in (8). The idea is to store representative precomputed function evaluations from the full dynamic system in the time invariant vectors $X \in \mathbb{R}^{n \times d}$ and $\mathbf{Y} \in \mathbb{R}^n$. These reference vectors are used to approximate the dynamics at a new value of $\mathbf{w}$ and $\mathbf{u}$ using the Markov property. Notice that $\mathbf{x}$ and X belong to a higher dimension than the state-space $\mathbf{w}$, ($\mathbf{x} \in \mathbb{R}^d, \mathbf{w} \in \mathbb{R}^m$ and $d \geq m$) because the variables in the kriging model should contain the information of the control inputs $\mathbf{u}(k)$.

By taking the approximated value from the previous step, DMK moves among the reference vectors to approximate the dynamic behavior. Due to the recursive nature of (11), error

will propagate from one step to the next. However, kriging also provides estimates of this error via (10), which could be used to indicate where additional sampling is needed.

III. DEMONSTRATION

To demonstrate and evaluate this approach, a second order elementary reaction was selected as a test problem. This simple test problem does not require an approximated model; however, it allows us to explore the various options and relevant concepts for DMK implementation, before considering more complex models. The model is

$$\frac{dC}{dt} = -rC^2 \tag{15}$$

For this test problem, the state space $\mathbf{w}$ is one dimensional, with initial concentration $C(k=0) = C_0$ and one control input $\mathbf{u}(k) = r(k)$, the reaction rate. For this test problem, DMK is described by:

$$C(k+1) = \hat{f}[\mathbf{x}(k), \mathbf{Y}, X] \qquad k = 0, 1, ... \tag{16}$$

$$\mathbf{x}(k) = \begin{bmatrix} c(k) & r(k) \end{bmatrix}^T \tag{17}$$

$$\mathbf{Y} = f[X] \tag{18}$$

where (18) is the solution of (15). In the next section, we consider various options for implementation of DMK. The nominal settings for the simulations are

- *Spatial distribution of the sample points*: Equally spaced square grid design (11 points from 0 to 100 for the concentration dimension and 9 points from 0 to 0.4 for the reaction rate dimension).
- *Spatial correlation function (also known as covariograms)*: Gaussian function. This function corresponds to the expression [4]

$$\begin{aligned} V_{ij} &= \sigma^2 \cdot R_{ij} \\ &= \sigma^2 \cdot exp\left\{ - \left[\left(\frac{|C_j - C_i|}{\theta_C} \right)^2 + \left(\frac{|r_j - r_i|}{\theta_r} \right)^2 \right] \right\}, \end{aligned} \tag{19}$$

  where θ_C and θ_r are parameters of the SCF, known as range parameters. Notice that it is not necessary to define a numerical value for σ^2 to use (8), but to compute the predicted error using (10) requires an estimate of σ^2 (see Section IV - *E*).
- *Regression functions*: Ordinary kriging (one unknown constant regressor p = 1) and full quadratic model (p = 6).
- *Kriging parameters*: Two range parameters fixed at the size of the grid in each dimension (θ_C = 10 and θ_r = 0.05).

IV. ANALYSIS OF RESULTS

A. Characteristics of the Test Problem

Fig. 2 shows the surface to be approximated by kriging for our test problem. A large gradient of the function is found in the region of low reaction rates. This behavior of the surface

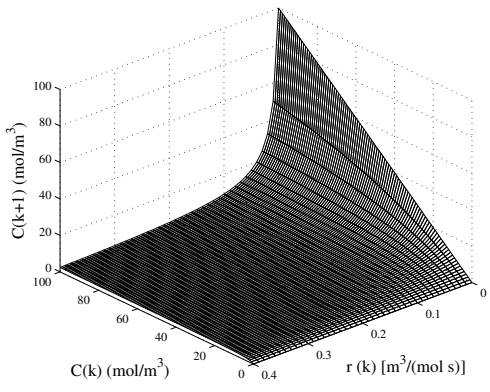

Fig. 2. Surface to be approximated by the dynamic mapping kriging for the proposed test problem.

will affect the quality of the prediction because kriging is based on local interpolation to model the *Z* term in (2).

To illustrate this point, Fig. 3 shows the root mean square error (RMSE) for complete trajectories between the exact solution and the kriging approximation up to t = 20 s with Δt = 1 s at several constant reaction rates and initial concentrations. At the boundaries of the domain, the error in the prediction increases because the number of sample points is lower, compared to the center of the domain (a point in the interior is influenced by more points, which improves the approximation for trajectories in this region). This behavior shows how the distribution of the sample points can be relevant in the model. The application of several techniques from design and analysis of computer experiments (DACE) could be used to improve the selection of the sample points [12].

Because the square grid DACE design is spatially symmetric, we originally expected a symmetric curve of the RMSE, but the magnitude and the gradient of the function have a strong influence at low values of r. This result suggests that classical space-filling DACE designs are not good enough for use in kriging approximations unless other elements in the model are adapted to match the behavior of the function. These classical techniques only consider the location of the points and not the information that these points can provide about the function to be modeled. Other aspects, such as the location of the boundaries and the gradient of the function, should also be considered to improve the design of the reference vectors by resampling the domain in certain regions [13].

To get an approximation for a specific trajectory in the domain, more points around this point can be added to improve the prediction locally. Fig. 4 shows different ways to enhance the approximation of a trajectory in the low reaction rate region, which has a high gradient. Compared to the original grid design over the reaction rate dimension, the approximation is improved by adding a complete row at r = 0.025 m^3/(mol·s), as shown in Fig. 4. Surprisingly, adding two new rows at r = 0.025 m^3/(mol·s) and r =

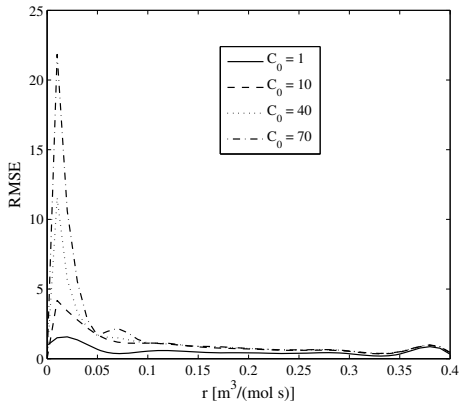

Fig. 3. Root mean square error of the kriging prediction for several trajectories. Settings: Square grid design, Gaussian SCF, ordinary kriging and fixed parameters.

0.075 m^3/(mol·s) does not generate a better approximation, showing that improvements are not necessarily achieved by adding more points in the region of interest. A more interesting result is shown when the added row is at the exact location of the reaction rate (at r = 0.04 m^3/(mol·s)). This location seems reasonable to get a better prediction, but ultimately results in a poor approximation. These results demonstrate that the design of sample points is one of the challenges in the application of DMK. As we will describe in Section IV - *E*, the kriging variance can help guide the selection of new sample points.

B. Effect of the regression functions in dynamic mapping kriging

There is little guidance about the selection of a good regression function to be used in kriging. The most frequently used function is called ordinary kriging ($h_1(x) = 1, p = 1$), which is a simple function compared to the pool of possible options. A good selection of a regression function will affect not only the quality of the approximation but also the optimization procedure used to fit the parameters of the SCF, especially in methods like maximum likelihood [14]. In particular, a better set of regression functions could reduce the possibility of obtaining an ill conditioning problem in the spatial correlation matrix R during the evaluation of the kriging model because it should generate a set of smalll range parameters that makes the R matrix closer to the identity matrix. Fig. 5 shows the effect of different regression functions used in dynamic mapping kriging. There is an improvement using the full quadratic model as the regression function.

C. Choosing a Spatial Correlation Function

Despite the diversity of spatial correlation functions proposed for a kriging model [10], [14], the Gaussian function is mostly used for its smoothness and its simplicity to be parameterized. Even though these properties are convenient, not all functions to be approximated will follow a Gaussian behavior for the covariance between the errors of the sample points. To illustrate how a Gaussian SCF would work for the test problem, a variogram curve fitting procedure[1] [15] was applied using a square grid design in Fig. 6. Clearly a Gaussian SCF does not represent the spatial relationship between the sample points.

Recently, Davis and Ierapetritou [15] used the variogram analysis to define the most appropriate function to describe the spatial relation of the points by making linear combinations of the major types of semivariogram functions. Based on these types of functions, a power SCF was fitted to our case (shown in Fig. 6). Using this fitted curve, DMK was performed but revealed ill conditioning problems when computing the inverse of the spatial correlation matrix. This problem is most likely due in part to the symmetry of the sampled points in the grid design. We have observed improved conditioning properties using the Latin Hypercube design [16], which is a stochastic space-filling methodology

[1] Variogram curve fitting is a procedure to generate an SCF which is based on the definition of a variogram and the Euclidian distance between sample points instead of absolute distances for each dimension. For more references about this procedure, see [14].

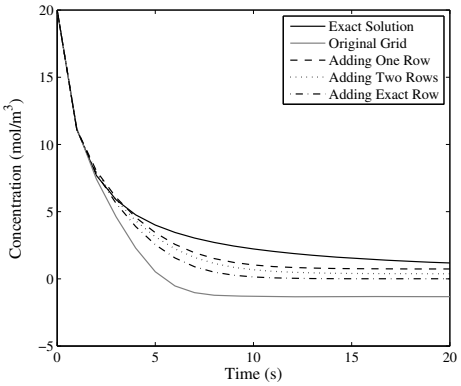

Fig. 4. Improvement of the kriging prediction by resampling. C_0 = 20 mol/m^3, r = 0.04 m^3/(mol·s). Settings: Square grid design, Gaussian SCF, full quadratic model and fixed parameters.

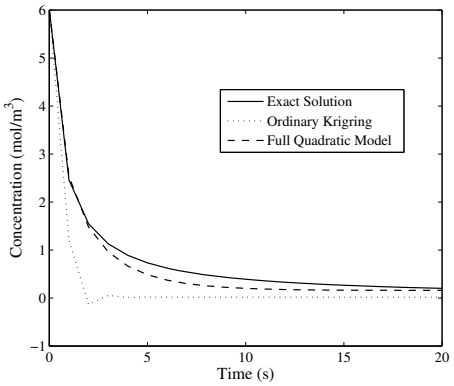

Fig. 5. Comparison between different regression functions in dynamic mapping kriging. C_0 = 25 mol/m^3, r = 0.24 m^3/(mol·s). Settings: Square grid design, Gaussian SCF and fixed parameters.

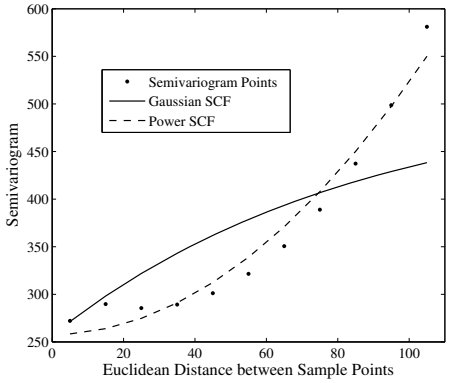

Fig. 6. Semivariogram curve fitting plot for the test problem.

for the selection of sample points. The appropriate selection of an SCF for a specific problem does not yet have a clear procedure [9].

D. Dynamic Changes in the Reaction Rate

One key aspect in control systems is the behavior of the model when the control inputs change, especially when these inputs represent a manipulated control variable. For the proposed test problem, a reaction rate profile was established to analyze how DMK behaves under changes in this control input. Fig. 7 shows the approximation made by DMK at two different initial concentrations C_0 = 10 mol/m^3 and C_0 = 60 mol/m^3. At C_0 = 10 mol/m^3, our model reacts properly to the changes in the reaction rate capturing the tendency on the exact solution. However, because the test problem tends to guide all the trajectories to values close to the domain boundaries, the error in the approximation is expected to increase at each step, affecting the prediction. This tendency could explain the increase in the predicted concentration, when the exact solution shows a decrease on it. A comparison at different initial concentrations, like C_0 = 60 mol/m^3, analyzes the effect of the location of the trajectory in the response of the model. At this initial concentration, the DMK model exhibits a better performance because the model makes a better prediction at the center of the domain.

E. Error Prediction Analysis

An analysis of the associated error that an approximated model has with respect to the true solution of the function is important to measure the uncertainty in the model, and it is a valuable information when a control system is designed. One of the attractive features of kriging is the ability to compute the prediction variance, by using (10). In order to analyze this aspect of our test problem, we selected maximum likelihood as a technique to estimate the process variance $\hat{\sigma}^2$ as a function of the actual θ_C and θ_r parameters of the SCF. The expression for this estimate is [10]

$$\hat{\sigma}^2(\theta_C, \theta_r) = \frac{(\mathbf{Y} - H\hat{\boldsymbol{\beta}})' R^{-1} (\mathbf{Y} - H\hat{\boldsymbol{\beta}})}{n}. \qquad (20)$$

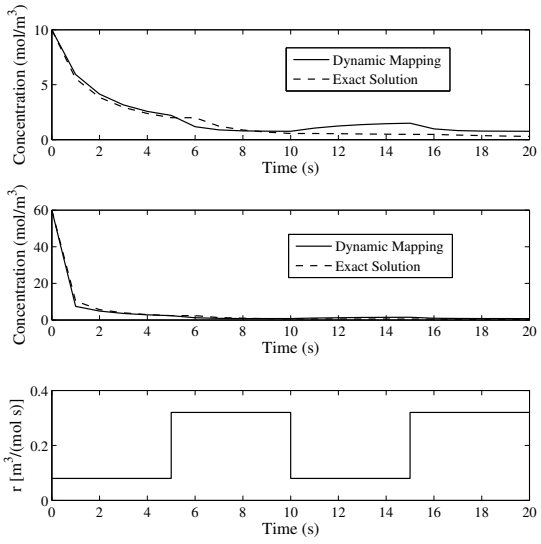

Fig. 7. Reaction rate profile for the simulation. C_0 = 10 mol/m^3 and C_0 = 60 mol/m^3. Settings: Square grid design, Gaussian SCF, full quadratic model and fixed parameters.

Fig. 8 shows the true error between the DMK prediction and the exact solution of (15) at the discrete time index $k = 1$, compared with the predicted error for DMK computed by (10) for a constant initial concentration C_0 = 20 mol/m^3 and several reaction rates. As was expected, the magnitude of the predicted error increases at the boundaries because the high uncertainty at these regions.

Even though the magnitude of the predicted error for our DMK model using the full quadratic model as a regression function is lower than for ordinary kriging, the predicted error overpredicts the true error in both cases. By changing the set of range parameters θ_C and θ_r it is possible to get a better agreement between these curves. To get this new set of range parameters, we assumed that the estimate set of range parameters will not differ substantially no matter which DACE design is used to get those estimates. Under this assumption, a Latin Hypercube DACE design with n = 99 points was used and a new set of range parameters θ_C = 17 and θ_r = 0.08 were estimated using a maximum likelihood procedure. This new set of sample points avoids the ill conditioning problem that our previous DACE design has during the optimization procedure. As seen in Fig. 8, this new set of parameters generates a better prediction error for the model, but it still needs more research in order to improve this predictions. On analyzing these error prediction results, the kriging variance can be used as a guiding element to generate new sets of sample points, by resampling where the error prediction is higher. For example, in Section IV - *A*, it was shown that adding a row at the highest error prediction (at r = 0.025 m^3/(mol·s)) generates a better prediction.

Fig. 8 also shows the impact of the regression functions in the error prediction. Depending on which regression function is used, the magnitude of the true error is different, especially

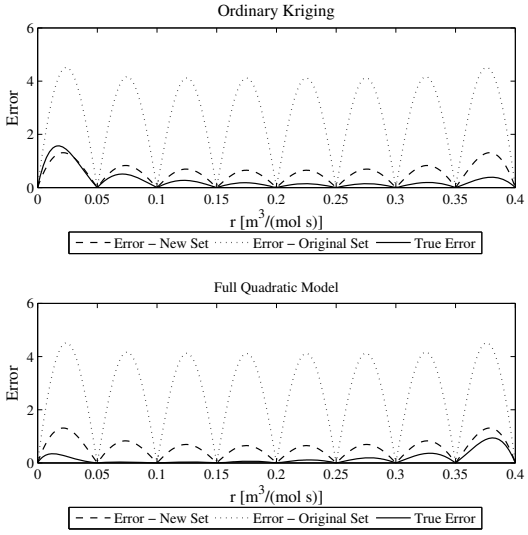

Fig. 8. Error prediction graph for dynamic mapping kriging for different regression functions. C_0 = 20 mol/m^3. Settings: Gaussian SCF and fixed parameters. The original set reffers to the square grid DACE design and the new set reffers to the Latin Hypercube DACE design.

at the boundaries of the domain. These functions help to capture the trend of the system at some specific regions in the domain. Subdividing the spatial domain in DMK could create a better approximated model. Some ideas to subdivide the domain as follows:

- Divide the domain into smaller regions with special distribution of the sample points at each region.
- Divide the application of the spatial correlation function at some regions where the function presents a different behavior.
- Define regions where some regression functions work better that others by using polynomial regression functions with compact support.

V. CONCLUSIONS

In the paper, we presented an exploratory study of the application of kriging to approximated complex systems under the framework of discrete time state-space models. The presented application of kriging requires the generation of reference vectors as previous information about the process and then uses a discrete time approach to model the evolution of the process during time. One key element that affects the performance of kriging is the design of the reference vectors. Using a simple test model, it was shown that the application of classical DACE designs in kriging are not enough to capture the necessary information for a good approximation, especially when a prediction is made close to the boundaries of the domain. These results also showed the potential of the kriging error prediction as a tool to improve the design of the reference vectors and to estimate the true error. The optimization procedure to obtain the range parameters is an important final issue to build the reduced model. Some solutions for this problem are a good selection of the regression function or a random selection of the sample points. It was also observed that the form of the SCF should be tailored for specific engineering problems.

In summary, dynamic mapping kriging exhibits a good agreement in the approximation under the influences of changes in the parameters of the system. It is a promising approach for generation of reduced models for complex simulations, which can then enable online computation and control.

VI. ACKNOWLEDGMENTS

Financial support for this work was provided by the Air Force Office of Scientic Research under award number FA9550-07-1-0161.

REFERENCES

[1] C. Oguz and M. A. Gallivan, "Identification and evaluation of a dynamic model for a thin film deposition process," in *Proceedings of the 2007 American Control Conference*, 2007, pp. 4124-4129.

[2] S. B. Pope, "Computationally efficient implementation of combustion chemestryusing in-situ adaptive tabulation," *Combustion Theory and Modelling*, Vol. 1, No. 1, 1997, pp. 41-63.

[3] C. S. Hsu, *Cell-to-Cell mapping: A method for global analysis for nonlinear systems*, Springer-Verlag, New York, 1987.

[4] J.D. Martin and T.W. Simpson, "Use of kriging models to approximate deterministic computer models," *AIAA Journal*, Vol. 43, No. 4, 2005, pp. 853-863.

[5] L. Li and P. Revesz, "Interpolation methods for spatio-temporal geographic data," *Computers, Environment and Urban Systems*, Vol. 28, 2004, pp. 201-227.

[6] E.J. Miller, "Towards a 4D GIS: four-dimensional interpolation utilizing kriging," in *Innovations in GIS 4: selected papers form the Fourth National Conference on GIS Research UK*, Z. Kemp, Editor. 1997, Taylor & Francis: London. pp. 181-197.

[7] G. Jost, G.B.M. Heuvelink and A. Papritz, "Analysing the space-time distribution of soil water storage of a forest ecosystem using spatio-temporal kriging," *Geoderma*, Vol. 128, 2005, pp. 258-273.

[8] W.S. Kerwin and J.L. Prince, "The kriging update model and recursive space-time function estimation," *IEEE Transactions on Signal Processing*, Vol. 47, No. 11, 1999, pp. 2942-2952.

[9] H. Wu and F. Sun, "Adaptive kriging control of discrete-time nonlinear systems," *IET Control Theory & Applications*, Vol. 1, No. 3, 2007, pp. 646-656.

[10] J.R. Koehlrer and A.B. Owen, *Computer experiments*, in *Handbook of Statistics 13: Design and Analysis of Experiments*, S. Ghosh and C.R. Rao, Editors. 1996, Elsevier Science: Amsterdam. p. 261-308.

[11] A.S. Goldberger, "Best linear unbiased prediction in the generalized linear regression model," *Journal of the American Statistical Association*, Vol. 57, No. 298, 1962, pp. 369-375.

[12] Crary, S.B., "Design of computer experiments for metamodel generation," *Analog Integrated Circuits and Signal Processing*, Vol. 32, 2002, pp. 7-16.

[13] J.-C. Jouhaud, et al., "A surrogate-model based multidisciplinary shape optimization method with application to a 2D subsonic airfoil," *Computers & Fluids*, Vol. 36, 2007, pp. 520-529.

[14] N.A.C. Cressie, *Statistics for spatial data, Revised Edition*, 1993, New York: John Wiley & Sons, Inc.

[15] E. Davis and M. Ierapetritou, "A kriging method for the solution of nonlinear programs with black-box functions," *AIChE Journal*, Vol. 53, No. 8, 2007, pp. 2001-2012.

[16] Wei-Liem Loh, "On latin hypercube sampling," *The Annals of Statistics*, Vol. 24, No. 5, 1996, pp. 2058-2080.