A joint model-based design of experiments approach for the identification of Gaussian Process models in geological exploration

Philipp Deussen[1], Federico Galvanin[1]

Department of Chemical Engineering, University College London, Torrington Place, London, WC1E 7JE, UK

Corresponding author: f.galvanin@ucl.ac.uk

Introduction and Motivation

Problem Metals demands from electrification, infrastructure and industrial projects is projected to rise, but productivity remains low.

- Unknown subsurface must be modelled to avoid mine planning mistakes
- **High uncertainty** in model selection and parameter estimation; **few samples**
- Surrogate models are employed (**Kriging**[5] a.k.a. **Gaussian Process** [5] (GP))
- Funding of mining projects depends on promising initial sampling results
- **Industry standard**[4] is to sample based on **qualitative expert decisions**, the **reduction of prediction variance** of one model or **space filling designs**

Solution A Model-Based Design of Experiments method is proposed that systematises exploration and model and parameter identification.

- **Parameter estimation** criterion reduces parametric uncertainty
- **Model discrimination** criterion includes prediction uncertainty at all locations (adjusted for Kriging models), model probability and prediction difference
- **Exploration criterion** minimises prediction variance and avoids local optima
- **Multi-objective** sampling procedure that optimises sampling for all objectives
- **Quantified metrics** of parameter uncertainty and model discrimination ability

Methodology

- Model consists of semivariogram and Kriging type (eq. 1-5, table 1)
- Parameter estimation objectives: estimates pass the t-test (fig. 5)
- Model discrimination objectives: correct models have low KL-divergence (fig. 4)

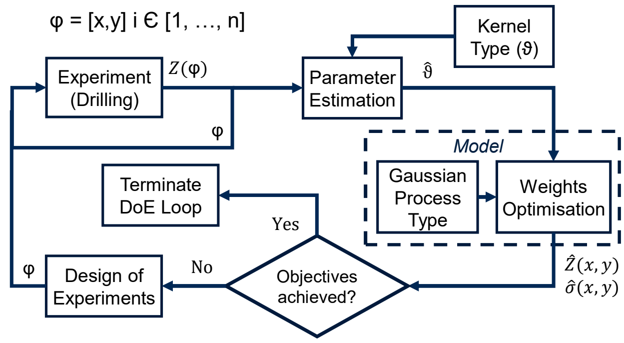

Figure 1 - Flowchart of proposed multi-objective MBDoE procedure.

Key Equations for Ordinary Kriging Models

- The **kernel of the GP**[5] is **correlation function** $R(h)$, relating semivariance $\gamma(h)$, distance h and the distribution variance σ_Z^2 of concentrations.

$$\gamma(h) = \sigma_Z^2 - R(h) \tag{1}$$

- **Ordinary Kriging** gives the **best linear unbiased estimate** of the **mean expected concentration,** Z, and its variance, σ_{OK}^2, using estimator $\hat{Z}$.

$$\sigma_{OK}^2 = E\left[\left(Z-\hat{Z}\right)^2\right] \tag{2}$$

- Estimators based on samples Z_i and their relative importance weights w_i.

$$\hat{Z} = \sum_{i=1}^{N} w_i Z_i \tag{3}$$

- **Optimal weights w_i** found from samples i and j by substitution (3 into 2):

$$\sigma_{OK}^2 = 2\sum_{i=1}^{N} w_i R(Z, Z_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j R(Z_i, Z_j) \tag{4}$$

- And then **minimising the Kriging variance**: $\partial\, \sigma_{OK}^2 \,/\, \partial\, w_i = 0$, giving matrices for: optimal weights, **W,** correlation between sampled points **A**, and sampled and unsampled locations **P**. To then find the predictions[2]:

$$\sum_{i=1}^{N} R(Z_i, Z_j)\, w_i = R(Z, Z_j) \tag{5}$$

Design Objective

- Improved Epsilon Constraint Method[3] used to optimise multi-objective MBDoE[2]
- The sensitivity matrix Q is used to determine Fisher Information, **H**.[1]
- Determinant of H is used as the scalar criterion to optimise estimability.[1]

$$\mathbf{Q} = \begin{bmatrix} \frac{\partial \hat{y}_1}{\partial \vartheta_1} & \cdots & \frac{\partial \hat{y}_1}{\partial \vartheta_{n_\vartheta}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_{n_{\hat{y}}}}{\partial \vartheta_1} & \cdots & \frac{\partial \hat{y}_1}{\partial \vartheta_{n_\vartheta}} \end{bmatrix} \approx \begin{bmatrix} \frac{y_1-\hat{y}_1}{\vartheta_1-\hat{\vartheta}_1} & \cdots & \frac{y_1-\hat{y}_1}{\vartheta_{n_\vartheta}-\hat{\vartheta}_{n_\vartheta}} \\ \vdots & \ddots & \vdots \\ \frac{y_{n_{\hat{y}}}-\hat{y}_{n_{\hat{y}}}}{\vartheta_1-\hat{\vartheta}_1} & \cdots & \frac{y_{n_{\hat{y}}}-\hat{y}_{n_{\hat{y}}}}{\vartheta_{n_\vartheta}-\hat{\vartheta}_{n_\vartheta}} \end{bmatrix} \tag{6}$$

$$\mathbf{H} = \sum_{i=1}^{n_{exp}} \sum_{j=1}^{n_m} \left[\frac{1}{\sigma_{ij}^2} Q_{ij}^{\mathrm{T}} Q_{ij}\right] \tag{7}$$

- Model discrimination: Schwaab Criterion[2] (Ψ^{MD}); variables as in eqns. 1-5
- Probability of models m and n, P_m, uses sum of Kriging variance instead of χ^2

$$\Psi^{MD}(x,y,\vartheta) = \sum_{m=1}^{M-1}\sum_{n=m+1}^{M}(P_m, P_n)\mathbf{D}_{m,n}(x,y,\vartheta)^{\mathrm{T}}\mathbf{V}_{m,n}^{-1}(x,y,\vartheta)\mathbf{D}_{m,n}(x,y,\vartheta) \tag{8}$$

$$\mathbf{D}_{m,n}(x,y,\vartheta) = \left[\hat{\mathbf{Z}}_m(x,y,\vartheta) - \hat{\mathbf{Z}}_n(x,y,\vartheta)\right] \tag{9}$$

$$V_{m,n}(x,y,\vartheta) = 2\sigma_{OK\,m,n}(x,y) + \sigma_{OK\,m}^2(x,y,\vartheta) + \sigma_{OK\,n}^2(x,y,\vartheta) \tag{10}$$

$$\phi_m(\vartheta) = {}^{1}\!/\!\sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{k=1}^{K}\sigma_{OK\,m}(x,y,k,\vartheta) \tag{11}$$

$$P_m(\vartheta) = \frac{\phi_m(\vartheta)}{\sum_{m=1}^{M}\phi_m(\vartheta)} \tag{12}$$

- For exploration, the location with maximum Kriging variance (eq. 5) is chosen.

Case study

- In-silico case data generated using M1 with five random samples
- Five samples are selected randomly which constitute the initial information
- For all kernels: $\delta_{(i,j)}(h) = \begin{cases} 1 \text{ if } h \in (i,j) \\ 0 \text{ otherwise} \end{cases}$

Table 1 – Two candidate Ordinary Kriging models with different kernel functions

Type of Kernel	Kernel Expression
M1: Spherical[5]	$\gamma(\mathrm{h}) = \left\{(s)\left(\frac{3\mathrm{h}}{2r} - \frac{1}{2}\left(\frac{\mathrm{h}}{r}\right)^3\right)\right\}\delta_{(0,r)}(\mathrm{h}) + s\,\delta_{(r,\infty)}(\mathrm{h}) + n\,\delta_{(0,\infty)}(\mathrm{h})$
M2: Gaussian[4]	$\gamma(\mathrm{h}) = \left\{(s)\left(1 - e^{-3\frac{\mathrm{h}^2}{r^2}}\right)\right\}\delta_{(0,r)}(\mathrm{h}) + s\,\delta_{(r,\infty)}(\mathrm{h}) + n\,\delta_{(0,\infty)}(\mathrm{h})$

Results

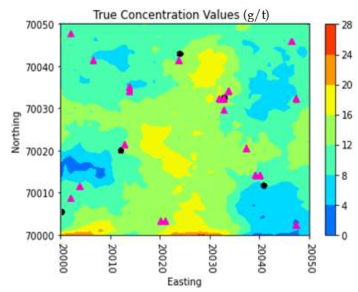

Figure 2 - Map of generated deposit with initial samples (●), MBDoE samples (▲)

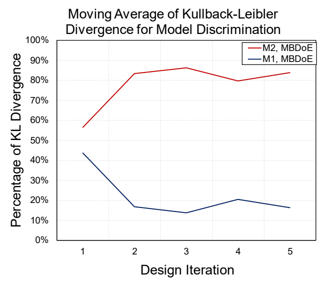

Figure 3 - Moving averages of model contributions to Kullback-Leibler Divergence

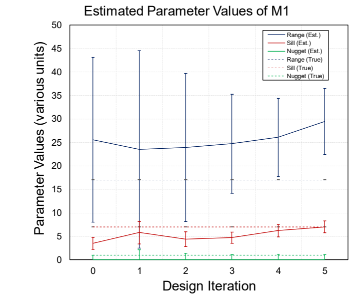

Figure 4 – Parameters: estimates and true values

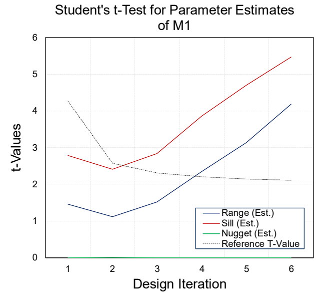

Figure 5 – Parameters: Student t-Test values

- Experimental budget of **five** design **iterations** was spent (samples in fig. 1).
- Model discrimination **identified the true model** with 85% confidence (fig. 3).
- **KL-divergence** (fig. 3) shows **distance between distributions**; contribution of candidate models to KL **moving average** is sensitive discrimination metric.
- **Correct estimates** within error margin for **two parameters** (fig. 4).
- **Statistical significance** measured by **t-test** (fig. 5) based on degrees of freedom, error and estimated value used as **metric of parameter estimates**.
- **Exploratory component** assists in **avoiding local optima** for the other two objectives and to **reduce the prediction variance** in the process.

Conclusions

- MBDoE Procedure was proposed to optimise sampling for three design objectives
- Design criteria and success metrics, e.g. KL-divergence, were applied to Kriging
- True model and two of its parameters were identified; design space was explored

Future work

- Compare to industry-standard designs (space-filling, variance minimisation)
- Explore Monte Carlo methods in model discrimination and parameter estimation
- Explore average, instead of maximum, variance minimisation for exploration

1. Franceschini, G. and Macchietto, S. (2008) Model-based design of experiments for parameter precision: State of the art. Chem Eng Sci, 63, pp 4846-4872.
2. Galvanin, F., Enhong, C., Al-Rifai, N., Gavriilidis, A. and Dua, V. (2016) A joint model-based design approach for the identification of kinetic models in continuous flow laboratory systems. Computers and Chemical Engineering, 95, pp 202-215.
3. Mavrotas, G. (2009) Effective implementation of the ε-constraint method in multi-objective mathematical programming problems. Appl. Math. Comp., 213, pp 455-465.
4. Rossi, M.E. and Deutsch, C.V. (2014) Mineral Resource Estimation, 1st Ed., Springer.
5. Sahimi, M. (2011) Flow and Transport in Porous Media. 2nd Ed., Wiley-VCH Verlag.

Author Information & Affiliations:


Philipp Deussen


Prof. Federico Galvanin

Sargent Centre for Process Systems Engineering

UCL ENGINEERING Change the world