Adaptive sequential space-filling design for geostatistical simulations

Céline Scheidt and Jef Caers
Stanford University
Uncertainty Quantification using Distances

*Scheidt et al., 2007*

Model 1 $\delta_{12}$ Model 2

Model 3 $\delta_{13}$ $\delta_{14}$ $\delta_{24}$ Model 4

2D projection - MDS

Defined by similarity distance

Single reservoir model

Clustering

**Uncertainty Estimation**

CUMOIL (MSTB)

Exhaustive Set

KKM

P10/P50/P90

Current Difficulties with the Method

Distance-based techniques for uncertainty quantification require:

1. Construction of a **large initial ensemble** (hundreds) of models
2. Definition of an application tailored distance

✓ **How many models are needed?**

   ➡ Generate **adaptively** the “right” models, only when needed

✓ **How to define the distance?**

   ➡ Use of full-physics flow simulation (perfect distance)
Adaptive Sampling for Uncertainty

- Adaptive experimental designs (ED)
  o Uncertainty quantification using proxy response
  o History matching (EnABLE)

- Limitations of ED:
  o Spatial uncertainty is difficult to account for
  o Smoothness of the response is required
  o Design is only optimal for a single response at a given time
    (e.g. cumulative oil production at 10 years)

*Distance-based techniques overcome those limitations*
A Simple Example

Objective
Uncertainty quantification in water production

Distance
Difference in water production

Training Image
10 models

2D projection - MDS

120 models

snesim

P25-P75 Interval

Water Production

TIME (days)

0 1000 2000 3000 4000 5000 6000

0 200 400 600 800 1000 1200 1400

Provides reference uncertainty

Objective
Uncertainty quantification in water production

Distance
Difference in water production

Training Image
10 models

2D projection - MDS

120 models

snesim

P25-P75 Interval

Water Production

TIME (days)

0 1000 2000 3000 4000 5000 6000

0 200 400 600 800 1000 1200 1400

Provides reference uncertainty

Objective
Uncertainty quantification in water production

Distance
Difference in water production

Training Image
10 models

2D projection - MDS

120 models

snesim

P25-P75 Interval

Water Production

TIME (days)

0 1000 2000 3000 4000 5000 6000

0 200 400 600 800 1000 1200 1400

Provides reference uncertainty
Where to Add New Models?

Desired properties of the new models:

- Stay in the prior model of uncertainty (internal consistent)
- Have responses different from those already evaluated
Where to Add New Models?

1. New models should remain in the prior uncertainty
   - Use of kernel smoothing to generate a pdf of the current points
   - Generate candidate points from this density

[Diagram showing 2D projection with PDF, Initial models, and Candidate Points]
Where to Add New Models?

2. New models should have different response from the current ones
   - Select the point which is the farthest away from current points

![Diagram showing 2D projection - MDS with points and a red X indicating new location.]

Area where new models should be added

Pdf
- Initial models
- Candidate Points
- New Location
Creation of New Models

The Post Image Problem (Park et al, 2010)

- **Post-image solution:**
  Non-linear combination of existing models

- Use of **Probability Perturbation Method** (PPM, Caers, 2003) until convergence to the desired location
Creation of New Models

6 additional models created by PPM

2D projection - MDS

Many model created in the same area

Account for PPM bias:
Accept/Reject models using Importance Sampling

14 models at current iteration
10 initial
4 additional

Current models
Models accepted
Models rejected
Target Location
Uncertainty Evaluation

Changes have occurred in the interval of uncertainty
Iteration #2

2D projection - MDS

19 models at current iteration
6 models generated by PPM
5 models accepted

Uncertainty Quantification

P25-P75 Interval

Water Production

TIME (days)

- 120 models
- 14 models
- 19 models
Iteration #3

2D projection - MDS

Uncertainty Quantification

P25-P75 Interval

Water Production

- 25 models at current iteration
- 7 models generated by PPM
- 6 models accepted

models generated by PPM
6 models accepted

0 1000 2000 3000 4000 5000
0 200 400 600 800 1000 1200 1400

TIME (days)

120 models
19 models
25 models
Iteration #4

2D projection - MDS

27 models at current iteration
2 models generated by PPM
2 models accepted

When to stop?
P25-P75 interval has stabilized

P25-P75 Interval

Water Production

TIME (days)
Comparison with Random Generation

32 simulations performed in total:
- Generation of 10 sets of 32 random models
- Evaluation of the P25-P75 interval for each set

High variance in P25-P75 interval for random simulations
Test Case # 2

Prior probabilities
- \( p(T1) = p(T2) = 0.2 \)
- \( p(T3) = p(T4) = 0.3 \)

Reference: 400 models

2D projection - MDS

- T1 – 80 models
- T2 – 80 models
- T3 – 120 models
- T4 – 120 models
Initial Set of 10 Models

2D projection - MDS

Additional Difficulty:
- Preserve the prior probabilities of each TI

Important under-estimation of uncertainty
Implication in the Workflow

1. Select **one training image** at each iteration
   - Depends on the prior probabilities

2. During bias correction, **prior probabilities** should be accounted for
   - If one TI is over-sampled, it will be penalized
   - If one TI is under-sampled, it will be preferred
Iteration #1

Selection of TI4

15 models at current iteration
7 models generated by PPM
5 models accepted
Iteration #2

Selection of TI2

19 models at current iteration
5 models generated by PPM
4 models accepted

Iteration #3

Selection of TI3

21 models at current iteration
2 models generated by PPM
2 models accepted

Water Production vs TIME (days)

P25-P75 Interval

400 models (black)
19 models (blue)
21 models (red)
Iteration #4

Selection of TI1

23 models at current iteration
2 models generated by PPM
2 models accepted

Iteration #5

Selection of TI1

25 models at current iteration
3 models generated by PPM
2 models accepted

5/4/2011
SCRF Affiliate Meeting - 2011
Comparison with Random Generation

30 simulations performed in total:
- Generation of 10 sets of 30 random models
- Evaluation of the P25-P75 interval for each set

High variance in P25-P75 interval for random simulations
Concluding Remarks

- **Workflow to construct new models in a intelligent manner for uncertainty quantification**
  - Method is general and can be applied with spatial uncertainty and with multiple responses
  - Removes the difficulty of choosing a distance

- **Preliminary results show promise to:**
  - Reduce the number of models required to construct for accurate uncertainty quantification
Future Work

- **Location of the new models has a great impact on the quantile estimations**
  - Apply kriging on the rank
    - Use kriging value and kriging variance to define the new location to sample

- **Limit the time required to generate new models**
  - Use of a proxy flow model during the PPM procedure

- **Improve the stopping criteria**
Thank You!

Questions?
Taking into account possible bias (2/3)

**Importance Sampling with Rejection Control:**

- For each model generated by PPM, compute:
  
  \[ w_i = \frac{f(x_i)}{f^*(x_i)}, i = 1, \ldots, N_{PPM} \]

- Draw \( U_i \) from a uniform distribution on \([0,1]\)

- Accept the model if
  
  \[ U_i \leq \min \left\{ 1, \frac{w_i}{c_r} \right\}, i = 1, \ldots, N_{PPM} \]