Multi-point Geostatistical Simulation by Applying Quadratic Optimization Algorithm

Snehamoy Chatterjee

National Institute of Technology Rourkela
Department of Mining Engineering

Multi-point simulation

- Statistical-based algorithm
  - Neural Network (Caers 2001)
  - SNESIM (Strebelle 2002)
  - Markov Random field (Daly 2004)
  - Markov Mesh Model (Kjønsberg & Kolbjørnsen 2008)
  - HOSIM (Mustapha and Dimitrakopoulos 2010)

- Pattern-based algorithm
  - Simpat (Arpat and Cares 2007)
  - Filtersim (Zhang et al. 2006;)
  - Multi-dimensional scaling (Honarkhah & Caers 2010)
  - Wavesim (Chatterjee et al. 2012)
  - CDF-based (Mustapha et al. 2011)
  - SOM-based (Chatterjee & Roussos, 2011)

Multi-point simulation

- The problem with all pattern-based simulation algorithm is that it is highly dependence on pattern frequencies in the training image
- The problem is not with the patterns seen in the training image rather, with how the method treats patterns that are not present in the training image
- Statistical models can interpolate between observed patterns to compute the probability of patterns that are not present in the training image

Regression model using SVM

1	1	1
1	1	0
1	1	0
0	1	0
0	1	0
0	1	0

X

0
1
0
1
0
0

y

$$f(\mathbf{x}, w) = \sum_{j=1}^{k} w_j g_j(\mathbf{x}) + b$$

$$L_\varepsilon(y, f(\mathbf{x}, w)) = \begin{cases} 0 & if\ |y - f(\mathbf{x}, w)| \leq \varepsilon \\ |y - f(\mathbf{x}, w)| - \varepsilon & otherwise \end{cases}$$

$$\mathrm{Min}\ \ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

$$\begin{cases} y_i - f(\mathbf{x}_i, \omega) \leq \varepsilon + \xi_i^* \\ f(\mathbf{x}_i, \omega) - y_i \leq \varepsilon + \xi_i \\ \xi_i\ , \xi_i^* \geq 0, i = 1, \ldots, N \end{cases}$$

$$\max\ \Psi(\alpha) = \sum_{i=1}^{N}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i \alpha_i y_j \alpha_j K_{ij}$$

$$\sum_{i=1}^{N} y_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C$$

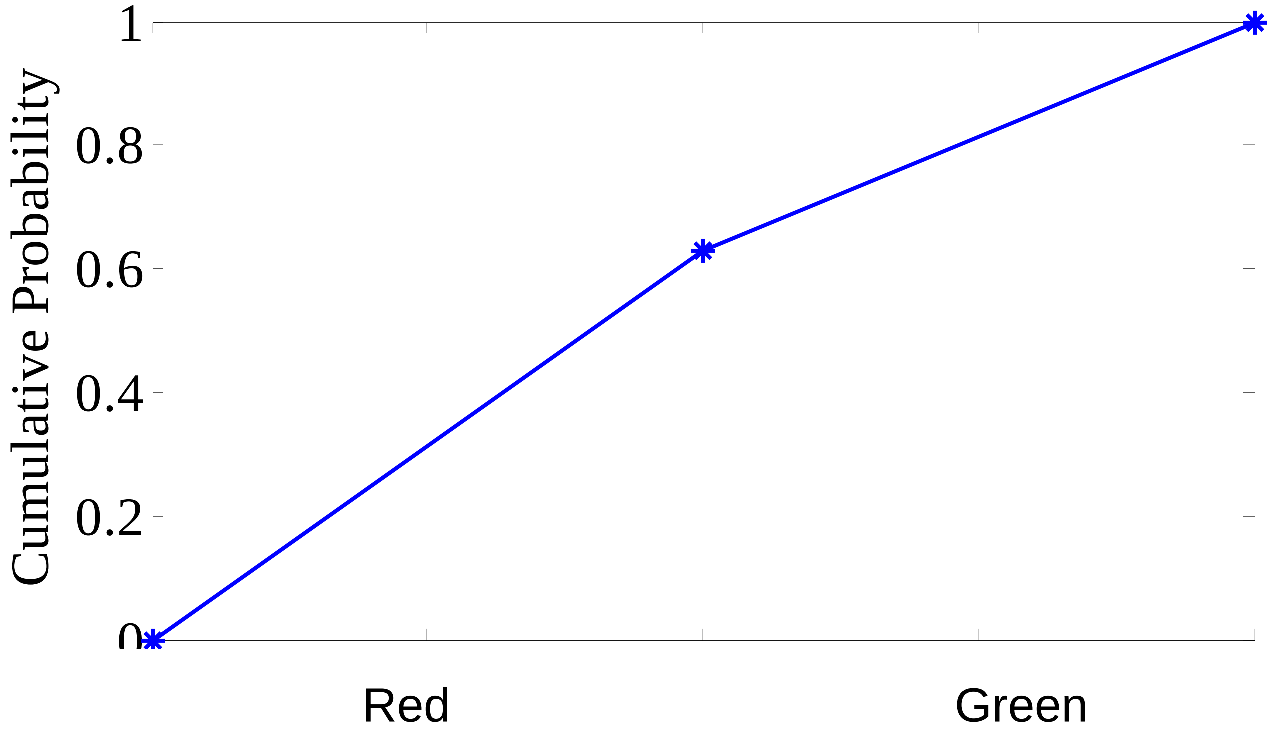
$$K_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$$

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / \sigma^2)$$

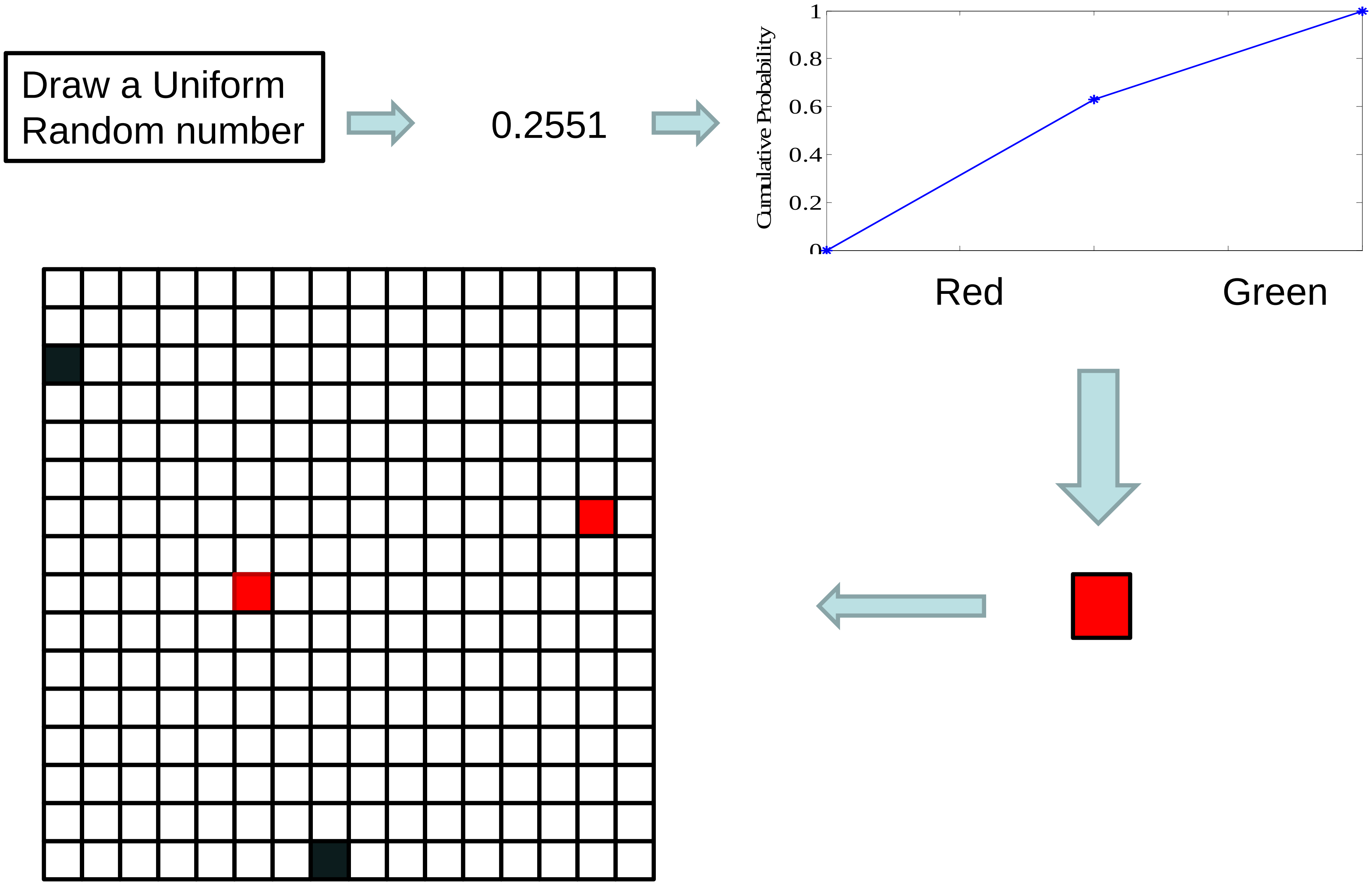
$$f(\mathbf{x}) = \sum_{i=1}^{n_{SV}} (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x})$$

Calculating the CCDF

$f(x) = 0.631$

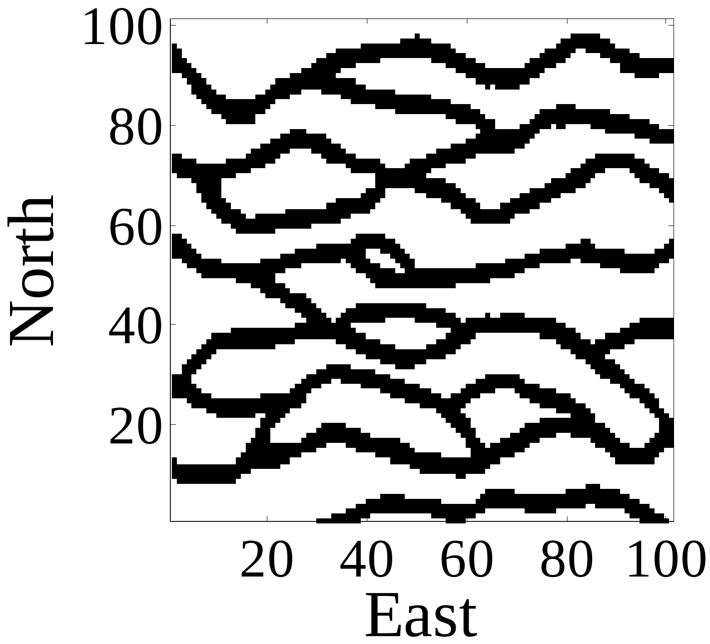


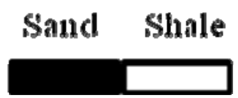
Calculating the CCDF



Validation of the method

Unconditional simulation with binary image

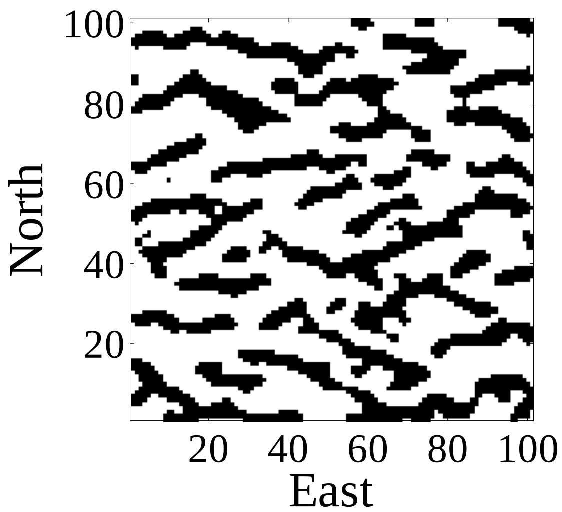




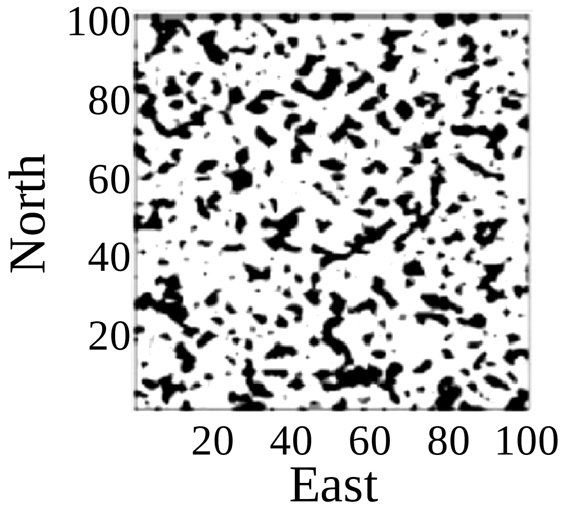
Training image

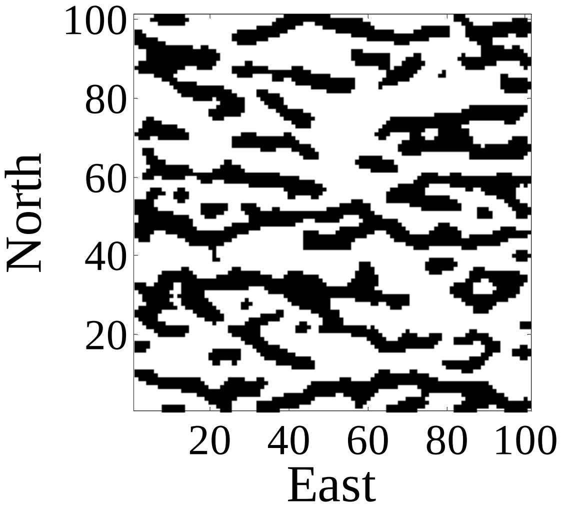
Validation of the method

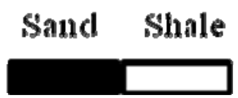
Our proposed approach



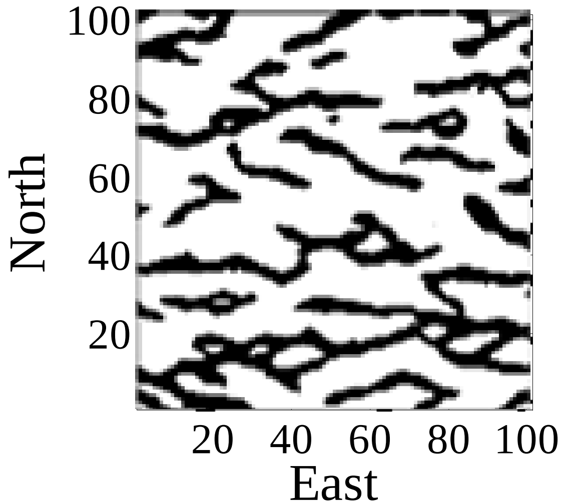
Snesim with single grid







Snesim with 3 multi-grid

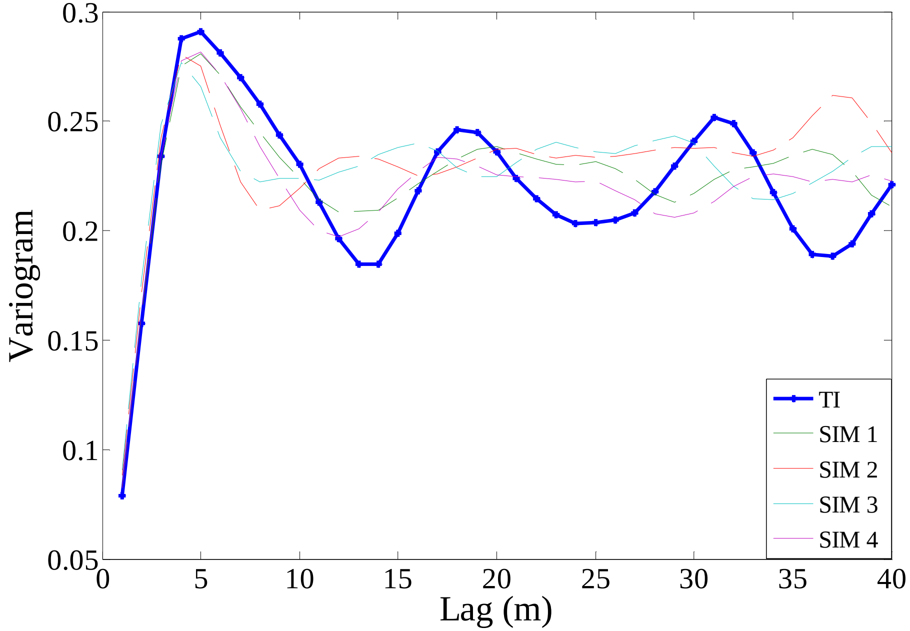


Reproduction of statistics

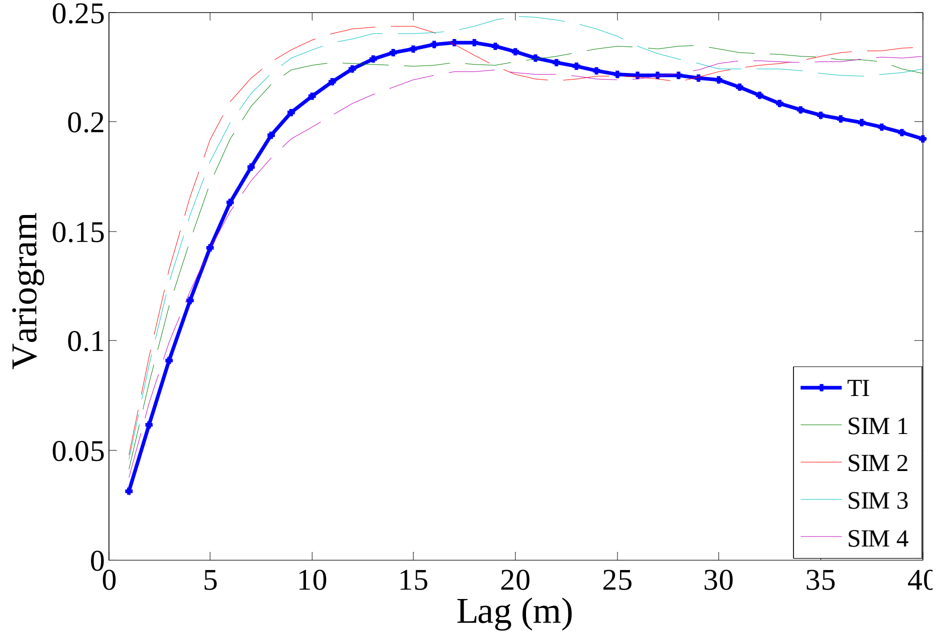
- The proportion of sand and shale in TI are 0.325, and 0.675

- The proportions of sand in 4 realisations of our proposed method are 0.317, 0.329, 0.321, and 0.33

- The proportions of sand in four realisations of snesim are 0.28, 0.289, 0.31, and 0.298



(a) Variogram E-W



(b) Variogram N-S

Conclusions

- A new multi-pint simulation algorithm is proposed
- SVM model was solved by QP to generate the *ccdf* at unknown point
- The indicators transformed of the pattern database are used to calculate the probability value of specific threshold
- The probability values at different threshold are used to calculate the *ccdf*.
- The algorithm is verified by two- dimensional conditional and unconditional simulation

Conclusions

- The advantage of the algorithm is that the chances of reproduction of statistics are much better than filtersim as well as snesim.

- The main limitation of the proposed algorithm is that the computational time is significantly high

- Mapping the high dimensional conditioning data into another domain with significantly less number of mapped data while preserving the data variability may reduce the computational time

Thank you