Transformation Spaces for Determining Model Complexity

Orhun Aydin[1] and Jef Caers[2]

Abstract How complex should a spatial or spatial-temporal geostatistical model be in order to suit the purpose for which it is used? This is a common question to all applications of geostatistical modeling whether it is mining, petroleum, environmental or any other. How many grid-blocks, how many indicator categories should we use? Surprisingly very few general and flexible tools are available to start addressing this important question. We propose a general framework for determining a suitable spatial model complexity on the basis of distances obtained from a series of image transformations and the linear combinations of these distances thereof. This general framework is applied in two workflows, namely Top-down and Bottom-Up to reach a simple enough set of models and demonstrated in an illustrative case.

Introduction

This study is about uncertainty quantification, so model complexity will be defined and evaluated for a set of geostatistical models, not just a single realization. Two workflows have been developed to reach the "simple enough" set of models: Bottom-Up and Top-down. In the Bottom-up workflow, a set of complex models is generated then gradually simplified until the "simple enough" set of models are reached. On the other hand, Top-down workflows involve building first simple models and then gradually adding detail until a "complex enough" set of models is reached.

Bottom-up workflow

The bottom-up workflow starts with the generation of a set of complex models and stops when "a set of simple enough models" is obtained by application of a simplification method and tracking the changes of the uncertainty of a response. We consider modeling uncertainty of such response as the purpose for which these models are built, for example a mine plan, a production forecast, or any other computationally demanding decision variable. In this study we reduce the complexity of such models by means of a simplification method, i.e. a repeatable operation that renders them computationally less intensive in terms of response evaluation while attempting to preserve the features of the complex case as much as possible. In order to reach these simple enough set of models in a CPU efficient way we use distances of easily computed responses applied on a set of models to

[1] Stanford University, Energy Resources Engineering, 367 Panama St, Stanford, CA 94110, USA orhuna@stanford.edu

[2] Stanford University, Energy Resources Engineering, 367 Panama St, Stanford, CA 94110, USA jcaers@stanford.edu

estimate the distance matrix for the desired response. In order to know which easily computed response distance has similar variations to the desired response distance we use desired response distance from a selected set of models to calibrate the distances obtained for easily computed responses. We also use those selected responses to quantify uncertainty at the given complexity level. We then reduce the complexity and carry on reducing complexity until we see large changes in our quantification of uncertainty. The bottom-up workflow operates as follows:

Input: A model set with maximal complexity

Output: A model set with minimal complexity that captures the desired response variations in the most complex model set

1. Start with the most complex model set
2. Combine easy-to-compute response distances by weighted linear combination using some initial weights
3. **while** $P10, P50, P90$ is similar for simpler model set
4. Compute easy-to-compute responses for the model set with reduced complexity (skip this step for the first iteration)
5. Combine linearly the distance matrix associated with easy-to-compute responses (skip this step for the first iteration)
6. Perform model selection using the combined distance
7. Compute the hard-to-compute responses on selected models
8. Calibrate the weights for easy-to-compute response distances using selected hard-to-compute response distances
9. Compute $P10, P50, P90$ for selected response distances
10. Simplify the model set
11. **end**

As seen above the complexity is reduced until the uncertainty quantification made changed a lot. This points out to the over-simplification of the set of models.

Top-down workflow

The top-down workflow starts with generating simplistic models. Similar to the bottom-up workflow, it is terminates when "a set of simple enough models" is reached. It applies a very similar algorithm compared to bottom-up workflow.

Input: A model set with minimal complexity

Output: A model set that captures the desired response variations

1. Start with the simplest model set
2. Combine easy-to-compute response distances
3. by weighted linear combination using some initial weights
4.
5. **while** $P10, P50, P90$ is similar for simpler model set
6. Compute easy-to-compute responses for the model set with increased complexity (skip this step for the first iteration)
7. Combine linearly the distance matrix associated with easy-to-compute responses (skip this step for the first iteration)

8. Perform model selection using the combined distance
9. Compute the hard-to-compute responses on selected models
10. Calibrate the weights for easy-to-compute response distances using selected hard-to-compute response distances
11. Compute $P10, P50, P90$ for selected response distances
12. Add detail to (increase complexity of) the model set
13. **end**

Unlike bottom-up workflow, this workflow terminated when the changes of uncertainty quantification falls below a given tolerance. This is logical since we expect over-simplistic models to change considerably by addition of detail, i.e. increase in complexity.

Illustrative case

The two workflows given above are applied to a flow problem. In this illustrative case, earth models are used in the workflows. Complexity is defined as the size of the grid on which these models are built on. Accordingly the simplification method is upgridding: an operation which decreases the number of grid cells while preserving structures within an earth model to some extent.

The purpose here is to model a problem involving injection of water and quantify uncertainty associated with the water-cut.

By using the bottom-up approach we start with 100X100 grid size (most complex model set) and reduce the grid size as the simplification method. In the top-down approach we start with 10X10 grid size (simplest model set) and add detail until we reach the simple enough set of models. The models that are used for bottom-up and top-down workflows are given below.

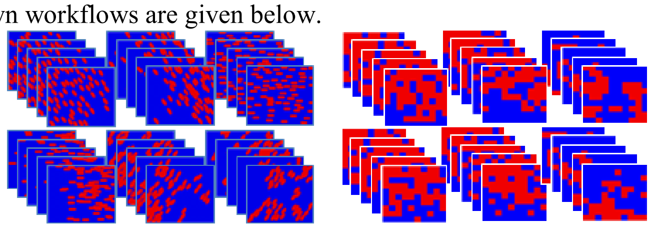

Figure 1:Earth Models, Most Complex Case (left) Simplest Case (Right)

In this case the hard-to-compute response is the water cut. In order to estimate the distances between water-cuts of earth models we propose image transforms as easily-computed responses. In this study we have used Shannon Entropy, Hough Transform, Radon Transform and Image Close as easy-to-compute responses.

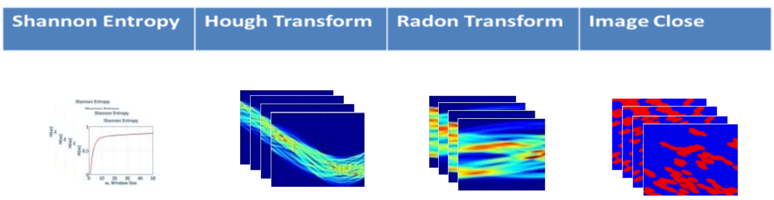

Figure 2: Image Transforms Used as Easy-to-Compute Responses

Figure 2 shows the easy-to-compute responses whose distances are used to estimate the distances between water cut values.

Applying these two workflows to the model sets given in Figure 1, we get the following uncertainty quantification results:

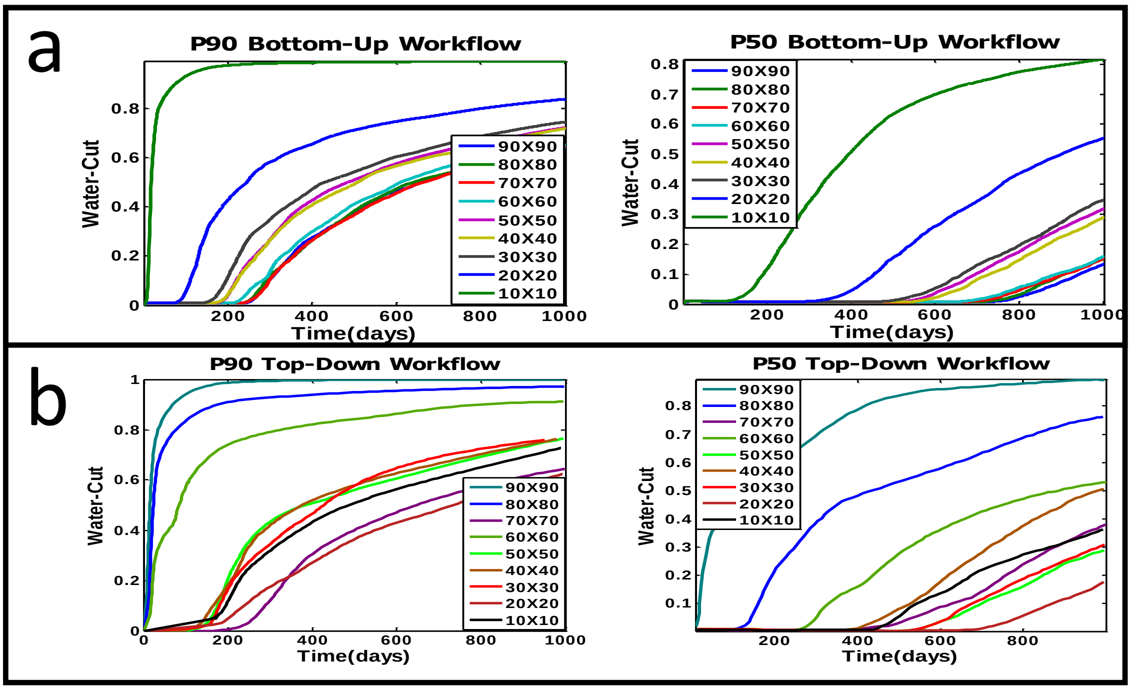

Figure 3: Uncertainty Quantification within Different Workflows

In both cases the total number of flow simulations that is need does not exceed 45. Moreover, both of the workflows point out to 50X50 Grid as the smallest grid-size that can be used for uncertainty quantification of water-cut. This value is the smallest grid size that can be used to quantify water-cut uncertainty.

Bibliography

[1] K. Park, 2011, “Modeling uncertainty in metric space”, PhD Dissertation, Stanford University, USA

[2] J. Caers, 2011, “Modeling uncertainty in Earth Sciences”, Wiley & Blackwell, Stanford University, USA