GEOMODELING: A TEAM EFFORT TO BETTER UNDERSTAND OUR RESERVOIRS

Part 3: Geostatistics

By Thomas Jerome, Suzanne Gentile and Jun Yang, RPS

INTRODUCTION

The previous paper introduced the general reservoir modeling workflow. One topic was left aside though: geostatistics. It is the focus of the present paper. Geostatistics are a whole set of techniques that allow modeling properties in three dimensions (3D) by taking into account the spatial variations of these properties. The topic is large and can't be covered in one paper. We narrowed down to kriging and simulation techniques, which are the more popular techniques by far. We also narrowed down to the modeling of rock types – and by generalization of discrete properties. Similar techniques exist for continuous properties like porosity. Once the concepts presented hereafter are assimilated, the reader will have no problem transferring them to the equivalent techniques for continuous properties.

RULE No 1: FIRST TRUST YOUR BRAIN AND ONLY THEN THE MACHINE

We are interpreting reservoirs from a limited amount of data – wells and seismic mostly. To palliate to this problem, we evaluate the reservoir characteristics between data points using interpolation and extrapolation techniques. Numerous mathematical techniques exist and it is up to us to select the one(s) most appropriate to a given property type (discrete/continuous), to a specific property (facies, porosity, permeability…), to the specific geological characteristics of the studied reservoir (clastic, carbonates, channels, reefs…) and to the specific purpose of the model (deterministic model / quantifying the uncertainties).

Interpolation means evaluating the property between the available data points. It is usually a well-defined problem, as the data points limit the possible range of the property. On the contrary, extrapolation means interpreting the property beyond the last data point. It is a much more difficult problem as one can't be sure if the trend observed around the last set of data can be propagated far past the last known value. The last section will illustrate this problem. Extrapolation problems can be turned into interpolation problems by including data on the immediate surroundings of the zone of interest (see the Figure 3 of the March paper for an example). All evaluation techniques interpolate and extrapolate at the same time. We have us to keep in mind which areas correspond more to extrapolation than interpolation, so as to be more skeptical about the model where extrapolation prevails.

Evaluation techniques can be deterministic or probabilistic. The former give a unique solution, such as the orange geometry for horizon A (Figure 1). The later will provide multiple solutions, such as the set of possible black geometries (Figure 1). Each realization respects the input parameters, here the well picks, while showing variations between the data points. Probabilistic techniques allow taking into account the uncertainty. In Figure 1, we will never know exactly where the horizon lays between the well picks. But at least we can, and we should, quantify the level of uncertainty whenever possible.

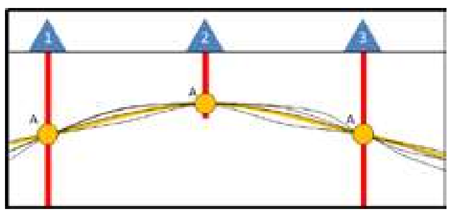

Figure 1. Evaluating the geometry of a horizon between wells in a deterministic way (orange line) or in a probabilistic way (black lines).

Mathematical evaluation techniques available on our computers are useful. For example, they allow quick testing of multiple models. Also, all the input parameters can be archived and the method rerun at a later stage. But, we must never forget that these techniques are the automation of the manual evaluation techniques that we, scientists, master. As such, we should never trust blindly what computers compute for us. If the results don't seem to make sense based on what we know about the reservoir (geological context, typical fluid characteristics, statistics at the wells…), then we must first, double-check how we used the software before eventually changing our vision of the reservoir. It must never be done the other way around. Maybe we simply didn't use the most appropriate evaluation technique or we didn't set its parameters correctly. Of course, no need to be extreme the other way. If everything ran as it should and the results still can't back up the assumptions, our hypotheses might need to be updated. Figures 2 to 13 and the accompanying text illustrate this point.

Geostatistics is the largest evaluation toolbox available to us, thanks to several main types of algorithms, which can, in turn, take multiple different types of input, from the most basic to the most sophisticated. Geostatistics are powerful because these techniques not only take into account the univariate statistics (mean value, min/max values, standard deviation…), but they also take into account how the property is varying spatially between the data points. This is perfect for modelers, as many reservoir properties vary spatially. For example, rock types will have accumulated differently in different parts of the reservoir, depending on the geological context (fluvial, marine…). Porosity might be increasing with depth because of the increasing compaction. As another example, water saturation will vary spatially depending on the fluid zone (gas, oil, water) and it might also vary depending on the distance to the contact itself (transition zone above an oil-water contact).

Variograms are the key mathematical objects used to capture the spatial variability of the data. They are input to kriging and simulation techniques. Variograms are to the understanding of spatial variability as histograms are to the understanding of univariate statistics: essential. For this reason, variograms are explained in the next section to some details so that every asset team member can understand how their reservoir modeler defined them in their project. As promised in the introduction paper though, the next section is free of any equation.

Once the notion of variogram is explained, the remainder of this paper goes through a simplified 2D dataset of a fluvial system to illustrate the results obtained by these two types of techniques.

VARIOGRAMS, AT THE HEART OF GEOSTATISTICS

In the next two sections, we'll go through the modeling of a sand/shale facies distribution, first using a dense dataset (Figure 2A) and then a limited dataset extracted from the dense dataset (Figure 2B). In this section, we are focusing on the variograms that will be used with this dataset.

Figure 2 shows the different variograms that will be used in the next sections. Variograms are represented on a map either as circles (black circles, Figure 2) or as ellipses (red and green ellipses, Figure 2). By extension, 3D variograms are represented as spheres or ellipsoids. A circular variogram means that there is no preferred orientation in the data. On the contrary, the more anisotropic the ellipse is, the more elongated and narrow the facies will be distributed in that direction. When the property is evaluated at a given empty location, the variogram being used (circle or ellipse) is centered on this location. The data points found inside the circle will have an influence on the value that will be computed at the new location. The data points outside of the variogram won't have any impact.

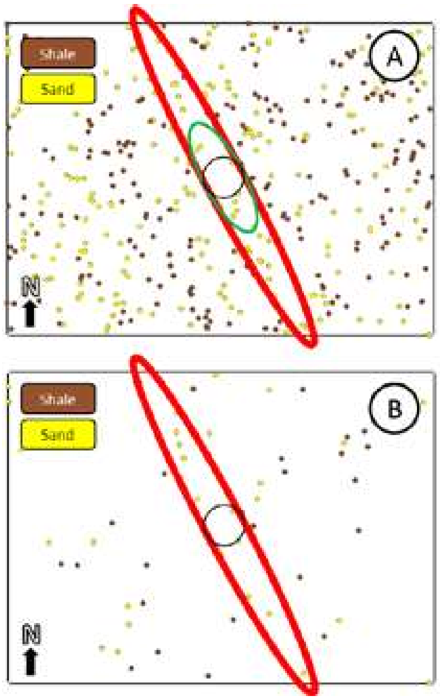

Figure 2. Dense (A) and limited (B) input dataset of sand and shale + variograms used as input for kriging and simulation.

Kriging algorithms use only the input data points, which is why kriging is a deterministic technique. Simulation algorithms, on the contrary, use both the input data points and the values that were computed before moving to this location. Simulations are probabilistic in nature because the empty nodes are not populated in the same order from one realization to the next. As a result, when the time comes to populate a given location, the surrounding available data will be different from realization to realization. For more details on how the surrounding data are used, please refer to (Pyrcz and Deutsch, 2014) for example.

2D isotropic variograms are defined by their range and their sill. The range represents the radius of the circle/ellipse. The sill will be defined in the next paragraph. 2D anisotropic variograms are defined by their maximum and minimum ranges, represented respectively by the ellipse semi-major and semi-minor axes. They are also defined by a sill, as for isotropic variogram, and by the azimuth of the semi-major axis (referenced to the North; 150 degrees on Figure 2 for example). A 3D variogram is usually defined as the combination of a 2D horizontal isotropic or anisotropic variogram and a vertical range. The vertical range is much smaller than the horizontal ranges. It reflects the fact that geological properties are continuous over a large area, while they rapidly change in the direction of deposition (here referred to as vertical). A true 3D variogram implies that the ellipsoid can have a dip and a plunge. True 3D variograms are used when we assume that the plane of deposition is not horizontal but inclined. True 3D variograms are not commonly used, but they are gaining some traction for example in oil sands project to model dipping IHS.

Variograms are defined using variogram analyzers (Figure 3 and Figure 4). The correlation found in different orientations (azimuths) is analyzed to identify the directions of the maximum and minimum horizontal ranges (Azimuths 150 and 60 degrees respectively in our dataset). For a given azimuth, the analyzer superimposes two objects: the experimental variogram and the variogram model. The experimental variogram is a succession of points computed from the input data. The variogram model is a mathematical equation that we have to adjust to the points of the experimental variogram. The circle/ellipses (Figure 2) are the spatial representations of the corresponding variogram models.

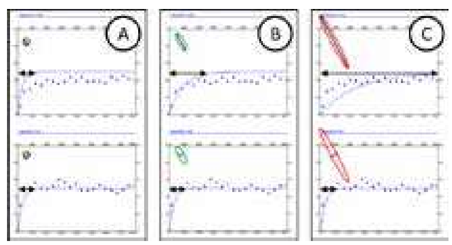

Figure 3. Variogram analyzers showing the experimental variograms for the dense dataset (Figure 2A) along the azimuths 60 and 150 degrees + variogram model for the circular variogram (A), the slightly elongated variogram (B) and the flattened variogram (C).

The modeling expert feeds two main datasets to the variogram analyzer: a set of azimuths and a set of distances between data points. The horizontal axis of the variogram analyzer represents these distances. For our dataset, we decided to compare each data point with the nearby point, if any, 400 meters away. We then do the same for a distance of 800 meters and so on until distances of 8000 meters. As a result, our experimental variograms have one point every 400 m. We did so in 10 different azimuths, of which we show only azimuths 60 and 150, the azimuths of the axes of the variogram model. For a given azimuth and a given distance, the goal is to check how two data points (= a pair) are similar. If the values of the two points making every pair are the same, the correlation is perfect and the corresponding point of the experimental variogram will be at Y=0 on the variogram analyzer. This only happens at the origin of the graph, where the distance is zero and each node is compared to itself. The bigger the distance, the lower the correlation will get, until a distance (the range) is reached beyond which there is no more correlation. At this stage, the points of the experimental variogram plateau. This plateau is the sill. For stationary and ergodic properties, the sill is the variance of the data.

A good variogram model will be one that starts at the origin, climbs progressively until reaching a plateau equal to the sill of the experimental variogram. It is essential to properly fit the experimental variogram between the origin and the range, as this is the part of the variogram model which will have the higher influence on the results of kriging and simulation. It is also essential to capture the anisotropy of the experimental variogram: keeping an isotropic (circular) variogram while the data show anisotropy will lead to missing some important information about the property we want to predict. Figure 3 A, B and C were used for kriging and the results are respectively shown on Figure 6, Figure 7 and Figure 8. As can be seen with this dataset, different variogram shapes do indeed give some drastically different models.

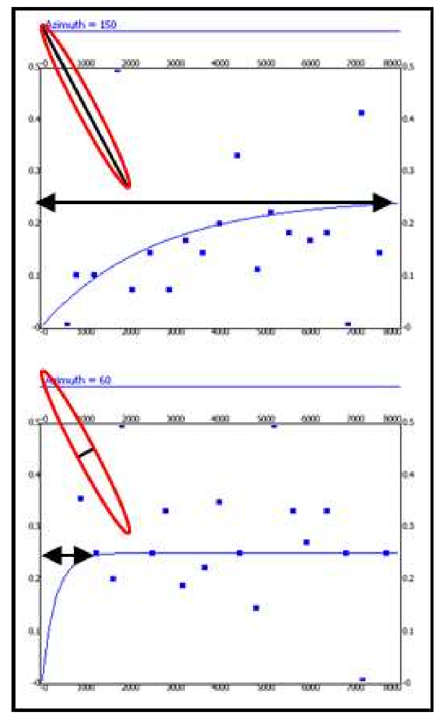

Figure 4. Variogram analyzer showing the experimental variograms for the limited dataset (Figure 2B) along the azimuths 60 and 150 degrees + variogram model for the flattened variogram (C).

Adjusting a variogram model is often challenging. It is rare to have a dataset as dense as the one used here (Figure 2A). As a result, it is rare to have horizontal experimental variograms as clean as in Figure 3. Often, the data is limited and the experimental variogram difficult to interpret (Figure 4). In this example, the experimental variogram in azimuth 60 degrees even looks as if it's a perfect sill: there are no points dipping down progressively to the origin. If this were true, it would mean that even for very short distances, there is no correlation between the values. While true for some ore deposits, this is rarely – if ever – the case in sedimentary rocks. The issue is not the geology but the dataset: the facies distribution is under-sampled and as such the first few points are not representative. In petroleum

(... Continued on page 24

(... Continued from page 23)
studies, variogram models should always start at the origin unless it can be backed up otherwise with solid geological evidences. In technical terms, we should never have any nugget effect (= variogram model not starting at zero).

KRIGING & SIMULATION – DENSE DATASET

The datasets and the variograms introduced in the previous section were used as input for kriging and simulation. The dense dataset is used in this section, while the limited dataset is used in the next section. Figure 5 shows the conceptual model from which the dense dataset was extracted. The limited dataset is a subset of the large one. The area represents a set of fluvial channels which flow from the North to the South along the azimuth 150 degrees. Naturally, in a real study, the truth is not known. Here, we are assuming that the well data and the geological context lead the geologist to see it is a fluvial system and the dipmeter data helped identifying the main azimuth of 150 degrees. Kriging was applied first with different variograms (Figure 6, Figure 7 and Figure 8) before simulation was run using the most elongated ellipse (Figure 9 and Figure 10).

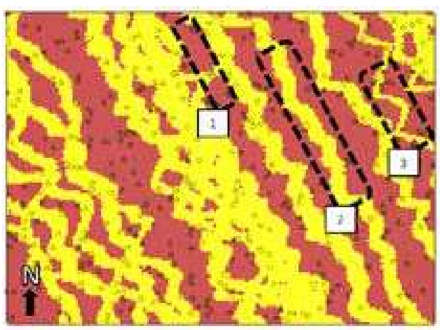

Figure 5. Conceptual sand/shale distribution from which a dataset was extracted and used in this study.

Kriging was first done using an isotropic variogram (Figure 6), even if the variogram analyzer showed that this variogram has too short a range in the azimuth 150 (Figure 3A). After all, with such a dense dataset, why shall we worry about the preferred orientation of the facies distribution? The data will take care of everything for us with some simple interpolation! The result is good overall and the sand facies does align along North-South geobodies, which might be interpreted as large channels. These channels are, nevertheless, wider than the input ones we know the dataset is coming from.

Then, kriging is done using a variogram model matching the experimental variogram (Figure 3B). One might argue that the plateau for the azimuth 150 is lower than the plateau at azimuth 60. In a real study, this would be investigated further. The resulting model (Figure 7) is closer to what we expected. The sand geobodies are more continuous along the azimuth 150 than in the first model. As our variogram is fitting to the experimental points, we could stop here.

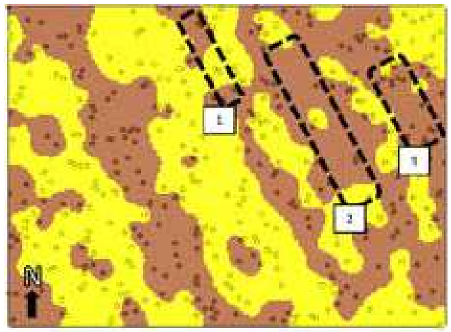

Figure 6. Kriging on the dense dataset – isotropic variogram.

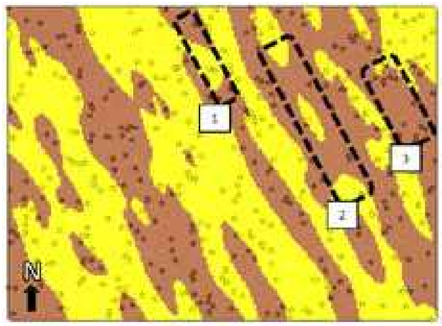

Figure 7. Kriging on the dense dataset – slightly anistropic variogram.

The geologist insisted though that he expected the channels to be even more continuous than they are now. He asked us to see if we could find a way to make it happen. After some testing, we decided to run kriging with a highly anisotropic variogram (Figure 3C). Of course, this variogram no longer matches the experimental variogram in the azimuth 150 - our new range is much too large. The kriging results pleased the geologist though (Figure 8) as the channels are now much better defined than before, as can be seen in the rectangle area labelled 2 noted in Figure 5 to Figure 8.. In the meantime though, we start creating channel geobodies where none should exist (rectangle labelled 1, same pictures). Also, we still can't get some channels right (rectangle labelled 3, same pictures). This channel was not sampled well enough by our wells for kriging to be able to track it

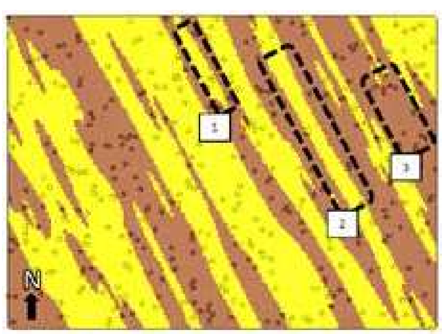

Figure 8. Kriging on the dense dataset – highly anistropic variogram.

In a real project, it would make sense to carry forward at least the two anisotropic variograms as it is impossible to know which one is the most reasonable one. The shape, dimensions and orientations of the variograms is a major source of uncertainty. In many reservoir modeling projects though, studying the variogram uncertainty is not done. Instead, modelers tend to pick one variogram – here the highly anisotropic one for example – and they run numerous simulation models with it. Figure 9 and Figure 10 are examples of two such simulation realizations. Each realization respects the input data, the input facies proportion and the input variogram. But each does it by distributing the sand and shale slightly differently.

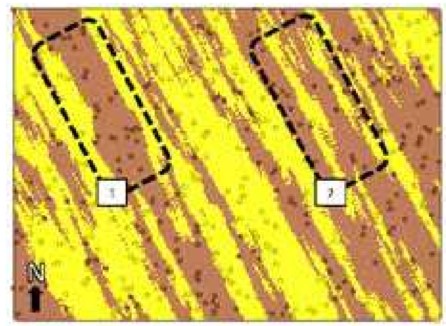

Figure 9. One possible simulation realization among many, created using the highly anisotropic variogram – dense dataset.

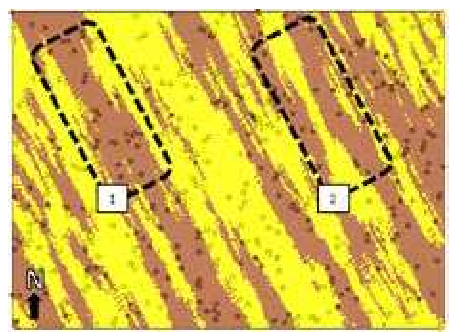

Figure 10. Second possible simulation realization among many, created using the highly anisotropic variogram – dense dataset.

Realizations defined by simulation have considerable value as together they build a range of possible rock distributions for the reservoir. This range can then be used to run sensitivity analysis while doing well planning or reserve computations for example. Ideally, modelers should first spend time understanding the uncertainty hidden in the variograms but also in the proportions they are using. In a second step, they can use simulation to generate multiple realizations in which these different key sources of uncertainty are taken into account.

At last, modelers should not limit themselves in matching the data strictly. It often makes sense to adjust our data analyses in light of the extra information provided to them by their team. General geological knowledge must be used to transform data into information.

KRIGING & SIMULATION – LIMITED DATASET

As mentioned earlier, the dense dataset is not realistic and one might even argue based on it that isotropic variograms are in fact good enough. To test this hypothesis, a subset made of 1/8th of points of the dense dataset has been randomly picked and kriging and simulation was run on it.

Firstly, kriging was run using an isotropic variogram. If we use the same small range than for the dense dataset, one gets an ocean of sand with a few patches of shale (Figure 11B). This is mathematically correct, but geologically implausible: it doesn't look anything like the fluvial system we know we have. Kriging is assigning an average value – sand in this case – at all the locations too far from the input points for the variogram to include them. This is an example of problematic extrapolation that is up to us to spot and fix by changing the kriging parameters. Using a very range 10 times the size of the initial one fixes this problem (Figure 11A). Nevertheless, the model still doesn't show any channel.

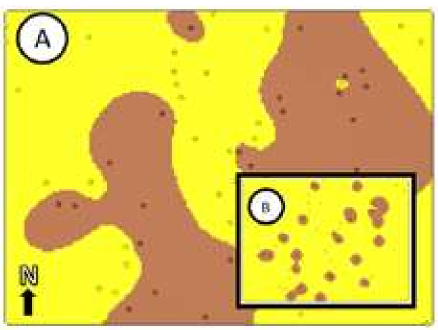

Figure 11. Kriging on the limited dataset – isotropic variogram with a long (A) and a short (B) range.

If we use the highly isotropic variogram, the model is showing some elongated geobodies that start looking like channels (Figure 12). But we are still far from the level of detail that we obtained with kriging the dense dataset (Figure 8).

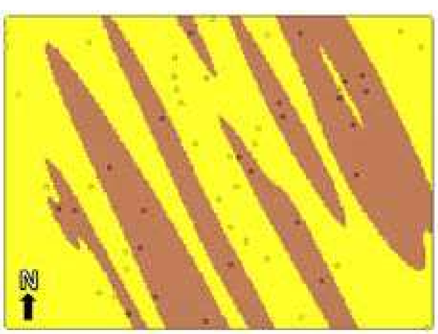

Figure 12. Kriging on the limited dataset – highly anisotropic variogram

On the other hand, the results of running simulation with this anisotropic variogram are very interesting (Figure 13 and Figure 14). The sand distribution in these two realizations is similar to the ones computed from the dense dataset (Figure 9 and Figure 10). It means that with a good variogram and simulation, even this limited dataset allows us to show possible geometries for the channels that our team knows must be present. Naturally, the local variations between these two realizations are much more important than with the two realizations of the dense dataset. For example, with this dataset (Figure 13 and Figure 14), the areas in rectangles 1 and 2 change from sand to shale drastically while with the two realizations ran on the dense dataset are very similar in these areas (Figure 9 and Figure 10)

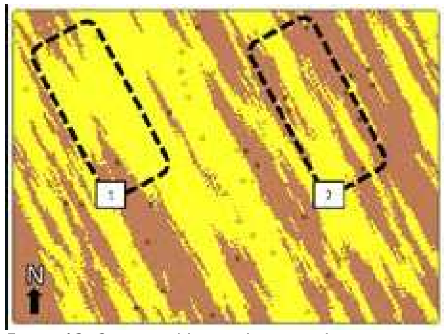

Figure 13. One possible simulation realization among many, created using the highly anisotropic variogram – limited dataset.

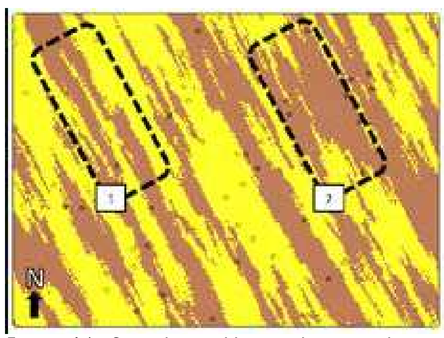

Figure 14. Second possible simulation realization among many, created using the highly anisotropic variogram – limited dataset.

This example shows that geostatistics have the potential to create realistic models even from a small dataset.

CONCLUSION

Geostatistics techniques are powerful because they take into account both the statistics and the spatial variability of the data. They are an essential part of every reservoir modeling workflow.

Having reviewed the reservoir modeling workflow in this paper and the previous one, the next three papers will focus on the interaction between reservoir modeling and geology, petrophysics and geophysics respectively. After this, the focus will shift to the interaction between reservoir modeling and engineering.

TO GO BEYOND

Geostatistics are a vast topic that is impossible to cover in a short introduction paper. Aspects of vertical, horizontal and 3D trends as well as the declustering of input data will be discussed in the papers on geology, petrophysics and geophysics.

Several important categories of geostatistical techniques could not be presented either by lack of space. Readers interested in plurigaussian simulations can refer to (Armstrong and als, 2011), while those eager to know more about multipoint geostatistics should have a loot at (Mariethoz and Caers, 2014).

(Isaaks and Srivastava, 1990) is a good introduction on geostatistics, as are the different courses on the topic that the CSPG offers every year.

Lastly, Alberta has the chance to host one of the world's leading teams in geostatistics: the Center for Computational Geostatistics in Edmonton, led by Professor Clayton Deutsch (www.ccgalberta.com). Each of their publications is a valuable source of information and of new ideas on geostatistics.

REFERENCES

Armstrong, M., Galli, A., Beucher, H., Loc'h, G., Renard, D., Doligez, B., Eschard, R. and Geffroy, F., 2011. Plurigaussian Simulations in Geosciences. Springer, 2nd edition. 176 pages.

Chilès, J.-P. and Delfiner, P., 2012. Geostatistics: Modeling Spatial Uncertainty. Wiley, 2nd edition. 734 pages.

Isaaks, E.H. and Srivastava, R.M., 1990. An Introduction to Applied Geostatistics. Oxford University Press. 592 pages.

Mariethoz, G. and Caers, J, 2014. Multiple-Point Geostatistics. Wiley-Blackwell. 376 pages.

Pyrcz, M.J. and Deutsch, C.V., 2014. Geostatistical Reservoir Modeling. Oxford University Press, 2nd edition. 448 pages.

TO CONTACT THE AUTHORS

Feel free to contact us if you have questions about this paper or about the series.

Thomas Jerome, Reservoir Modeling Manager and Geologist,
RPS: Thomas.Jerome@rpsgroup.com

Suzanne Gentile, Reservoir Modeler and Geophysicist,
RPS: Suzanne.Gentile@rpsgroup.com

Jun Yang, Reservoir Modeler and Geologist,
RPS: Jun.Yang@rpsgroup.com