geoconvention
Virtual Event
September 13 - 15
2021

Modeling three ways from electro-facies workflow elements: categorical, e-facies probabilities, and petrophysics with assignment

David Garner
TerraMod Consulting

Introduction

Geomodeling for petroleum reservoirs is conventionally done hierarchically. A 3D model of geological domains establishes regions within which rock and fluid properties can be considered "stationary", i.e., spatial statistical parameters such as histograms and variograms do not change within a single domain but may change from one domain to another. Although these geological domains are often referred to as "facies", and intended to characterize the depositional environment, many reservoir models do not directly use facies descriptions developed through sequence stratigraphic analysis and examination of core and image logs. Instead, they use "electro-facies" (or "e-facies"), which are categorical codes developed from petrophysical log curves and designed to capture the major facies groupings through some type of cluster analysis usually guided by the visually interpreted depositional-concept facies. Typically, a subset of wells that have been carefully geologically logged are used as training data to develop a set of rules that convert the common petrophysical logs (gamma ray, density, porosity, etc.) to a probabilistic interpretation of facies for each interval in every well. The e-facies classification workflow, a guided machine learning, has 3 main elements: 1) Training set of petrophysical curves and interpreted facies; 2) Probabilities of each facies at each location; 3) Assigned most probable facies, the resultant e-facies log. Implicit in this paper is that application of the e-facies strategy has enough geological verisimilitude that 3D e-facies models can provide a good basis for modeling reservoir architecture, i.e., that they capture reasonable geological spatial and stacking patterns.

This paper describes and compares three approaches to the development of 3D e-facies geomodels using the 3 workflow elements as conditioning data. Each approach arrives at the same type of result: a 3D model of categorical codes that can be used as stationary domains within which rock and fluid properties can be simulated. But they differ in the data used, and in their workflows. The first approach takes e-facies codes (element 3) developed through cluster analysis as conditioning data and uses a method for simulating categorical variables, such as Plurigaussian simulation or sequential indicator simulation, to directly build a 3D model of the categorical e-facies codes. This is the standard facies simulation workflow. The second method works directly with the e-facies probabilities (element 2) that most cluster analysis techniques can provide. These probabilities are simulated, following a series of spatial transforms, as continuous variables in 3D. A unique e-facies code is assigned, typically by taking the e-facies with maximum probability at each location. The third approach takes petrophysical logs (element 1) as conditioning data and uses a method for simulating jointly spatially correlated continuous variables, such as turning bands or sequential Gaussian simulation, to build 3D models of the log responses; these are then converted to e-facies using rules developed through cluster analysis.

Facies and e-facies

As discussed in Garner et al. (2009) and Garner (2020), the concept of facies is imprecise, with practitioners from different disciplines tending to interpret the meaning of the term according to their own background and expertise. To some, it is a term that describes lithological variations. To others, the facies codes capture variations in the ichnology, palynology, and depositional environment, giving location in the deposit and facies associations. Some use seismic data to discriminate "seismic facies", while others use electric logs to identify "electro-facies". Characterization factors may include diagenesis, fossils, organic composition, or common capillary behaviour. These many perspectives on the fundamental meaning of "facies" overlap and merge within the technical literature. None of these viewpoints is more correct than any other for modeling; each finds good use in different reservoir studies and for the ultimate purpose of each study.

For integrated reservoir teams and for modeling purposes, the term "facies" refers to a categorical code, typically available as conditioning data at well locations. An important conceptual point relates to the observational scale and the type of measurements we have access to. Core facies descriptions are based on interpretations from the surface area or face of small core samples. They are inches wide, represent a fine view of the rock at a specific location, and are typically regarded as the most reliable data. Petrophysical logs such as gamma ray, bulk density, and neutron logs measure volumes of rock two orders of magnitude larger than core, typically on the scale of a cubic meter. Although these observations are at a coarser scale than core, and often regarded as less precise, their scale is usually a better match to the cell heights and effective properties commonly used for geomodeling; this makes e-facies obtained from these logs good for characterizing 1D heterogeneity along the well bore, and useful as conditioning data for 3D geostatistical models.

In general, e-facies are developed using some type of cluster or discriminant analysis. The e-facies classification workflow is a guided machine learning (Garner, 2020). Training data are used to calculate the probability that a particular suite of log responses belongs to each of the chosen facies. This is not a trivial undertaking. Data quality, data vintages, and available curves have a strong impact on results. Geological concepts, interpretation and reasoning are often necessary for clarifying ambiguities and anomalies in e-facies predictions. Tradecraft is applied to guide the process. With the primary meaning of facies being imprecise, it is not possible to develop a perfect or unambiguously correct version of e-facies. The goal of the exercise should be to identify logical spatial groupings that correlate in a geologically plausible manner from well to well, and that can serve as conditioning data for creating reasonably stationary domains useful to define the reservoir's broad architecture with the structural and stratigraphic framework.

When used to condition geomodels, facies codes in a well often need to be "blocked", i.e., combined into a single representative value over an interval that is approximately equal to the height of geomodel cells. There are common rules to choose among methods for blocking categorical variables. The approach taken to blocking logs can introduce statistical artefacts, and checks are important for ensuring that blocked logs honour initial facies proportions. To account for statistical artefacts when using blocking by majority facies, Babak et al (2013) considers use of an auxiliary variable based on Shannon's entropy, capturing the degree of mixing of blocked

facies numerically. The artefacts can be reduced if the probability of each e-facies is retained from the discriminant or cluster analysis, instead of reducing the probability information to a single choice. The e-facies probabilities can be spatially modelled directly then post-processed to create a single simulated facies at each location. The degree of mixing variable can be computed later from the 3D probabilities for use in co-simulating effective scaled properties. The use of continuous probability variables, rather than single-choice categorical codes, creates flexibility when models need to be upscaled. We will use the Shannon's entropy measure to assess consistency of a set of realizations similar to the degree of mixedness at a location defined by Babak et al (2013).

A case study example from the North Sea

Data from a North Sea reservoir has been used to build e-facies models using three different approaches. A non-parametric classification method similar to Garner et al (2009) was used to calculate the probability distribution of four lithofacies, using a training data set that consists of a geologist's assignment of lithofacies along with gamma ray (GR), density (RHOB) and neutron porosity (NPHI) logs. In Figure 1, the full data base of assigned lithofacies was cleaned, with about two-thirds of the available data being retained as training data. This cleaning improves the rate of correct classification because it removes from the training data set many intervals whose petrophysical signature is inconsistent with the visual lithofacies assignment. The geologist's intention to create large contiguous intervals with the same lithofacies code often causes inconsistencies between petrophysics and lithofacies at a fine scale. The consequence of removing short-scale anomalies from the training data is that the e-facies classification (on the far right of Figure 1) re-acquires much of the short-scale variation that the geologist intentionally removed in the manual facies interpretation (the fourth strip chart from the left).

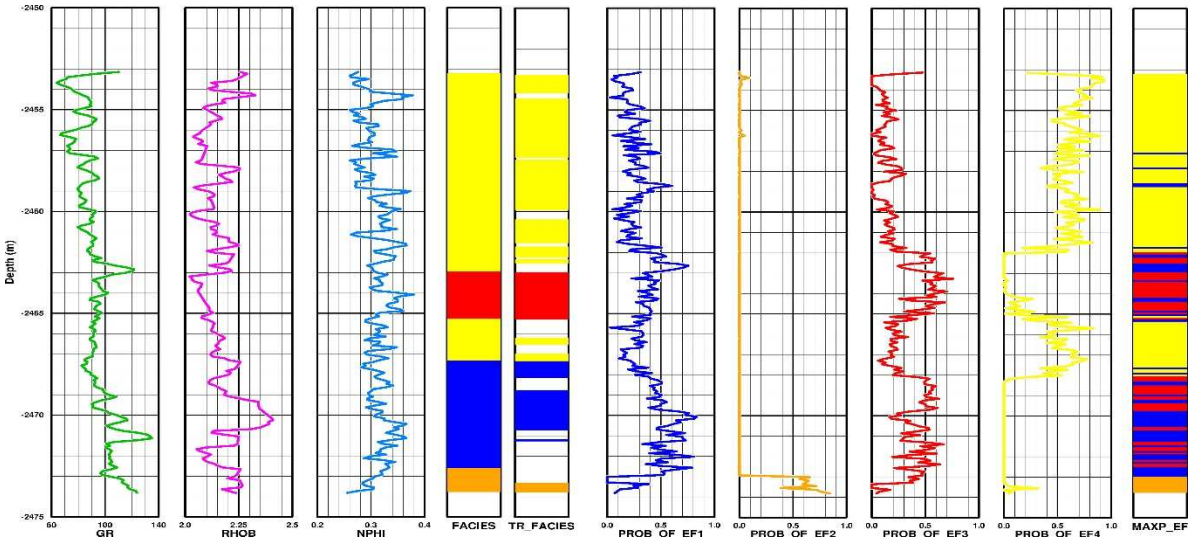

Figure 1 *The suite of logs available for e-facies classification: GR, RHOB, and NPHI, facies code (FACIES), training subset (TR_FACIES), with calculated probabilities of each e-facies (PROB_OF_EF1 to PROB_OF_EF4), and the e-facies assigned from the highest probability facies (MaxP_EF).*

Direct simulation of categorical e-facies (element 3) at well locations

Plurigaussian simulation, a common facies simulation method since the 1990's, was used to build 3D models of the e-facies, conditioned by well data. This approach makes use of 3D proportion curves and allows explicit control of transitions between simulated facies. In Figure 2a, lithotype #2 (orange) nests inside lithotype #4 (yellow), rarely touching lithotype #1 (blue).

Simulation of e-facies probabilities (element 2) at well locations

Turning bands simulation, a form of sequential simulation, was used to build 3D models of the e-facies probabilities, the 6th through 9th strip charts from the left on Figure 1. Since the four probabilities sum to one, they are not independent. Bivariate and multivariate scatter plots exhibit complex non-linear relationships which are difficult to directly simulate while preserving statistical properties of distribution shapes. Using a multi-step workflow allows for independent simulations of each transformed probability variable. The workflow consists of application of normal scores transforms, min-max autocorrelations (MAF), followed by normal scores transforms of the MAF factors as input to the 3D simulations. The three transforms are reversed after independent simulation to restore the shapes and correlations in the original space.

The normal scores transform is a form of Gaussian anamorphosis applied to obtain a variable with a Gaussian shape, i.e., mean of 0, and variance (and standard deviation) equal to 1 as a useful assumption for input to other steps. The MAF method is an extension of principal components analysis (PCA) which includes an additional rotation matrix to account for the spatial correlation or h-scatter of the variables. PCA provides statistically uncorrelated factors (at lag, h=0), but which are still spatially correlated. MAF provides spatially statistically uncorrelated factors (at lag h=h1) given a set lag value, h1. The second normal scores transform reinforces the Gaussian assumption prior to the standard 3D sequential simulation. The goal here was to have decorrelated variables to allow for independent simulations, and preserve the multivariate distributional relationships and shapes.

Treating the probabilities as continuous variables amenable to simulation violates some underlying properties of probabilities: that they lie between 0 and 1, and that they sum to 1. To ensure that the results could be regarded as probability-like variables, the simulated values were truncated to a minimum of 0 and then renormalized such that the four probabilities are bounded, and at any location, sum to 1. The reverse normal scores transform applied served to enforce the zero bounds. The inputs and simulated probabilities after truncation and renormalization are checked to ensure that the procedure has not introduced a bias. Then the e-facies with the maximum probability at each location is assigned, Figure 2b. Because the probabilities are calculated from spatially correlated e-logs, they inherit spatial continuity characteristics of the underlying e-logs. The method results do clearly exhibit the transition behaviour implicit in the conditioning data.

Direct simulation of e-log response (element 1) at well locations

Turning bands co-simulation was used to simulate the petrophysical responses in 3D, using normal scores and a linear model of coregionalization since the three attributes are correlated with each other. At each cell, the three simulated logs — GR, RHOB and NPHI — can be used in the same way that the original log values were used to develop e-facies at the wells. Using the same training data set used to create MAXP_EF, discriminant analysis was used to calculate the probability of each e-facies for the cell and the e-facies with the maximum probability was retained as the simulated value.

Figure 2c shows the results of this procedure. It acquires some of the nesting characteristics noted earlier for the Plurigaussian simulation. This is not because the user has any direct control over transition probabilities, but because the nesting is implicit in the spatial structure of the three logs on which the e-facies classification is based. The multivariate petrophysical definition of e-facies leads to distinct spatial categories, and is a reasonable classifier, in contrast to single petrophysical parameter simulations which are known to give soft boundaries in 3D models.

Figure 2c also shows one of the drawbacks of choosing the maximum likelihood e-facies as the simulated value: this causes the common facies (#4) to be over-represented and the uncommon facies (#3) to be under-represented. Other approaches to the assignment of a unique facies code in each cell could better control the honouring of global proportions. Alternatively, the procedure may require additional steps to ensure the simulated 3D petrophysical statistical distributions better honour the original training subset distributions used to create the 1D e-facies.

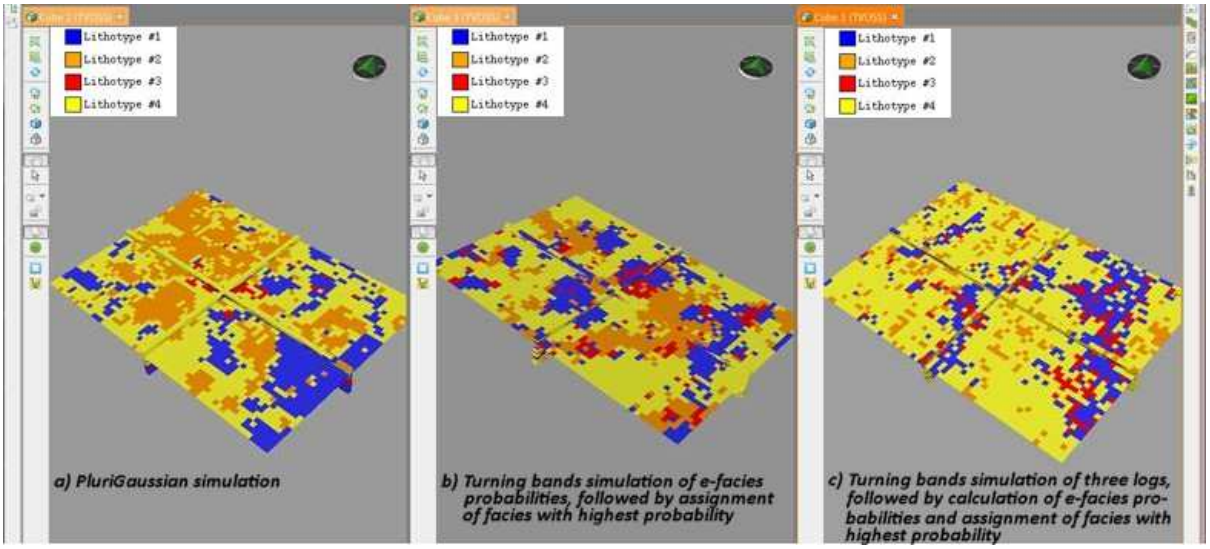

Figure 2 *Results from three approaches to simulation of e-facies.*

A comparison was made of the three approaches on their ability to honour the global proportions of the four e-facies. All the simple methods compared here tend to over-represent the most common e-facies (#4) and to under-represent the least common (#3). This effect is most noticeable when the e-facies classification is the final step following 3D simulation of petrophysical log responses. Additional steps to better honour multivariate parameter distributions may improve results. Direct simulation of e-facies codes does a better job of honouring global proportions, in

part because the Plurigaussian technique, like most other simulation methods for categorical variables, allows the user more direct control of global and local proportions.

Conclusions

The comparative studies in this paper demonstrate a variety of methods for creating 3D simulations of categorical facies. All the methods rely on a procedure for creating e-facies from petrophysical logs, a step that benefits from data cleaning and from sound geological interpretation of the resulting facies codes. Improvements in this first step of any of the procedures would lead to better results, although this paper used a direct classification method without common adjustments to optimize the results. Each of the described 3D modeling workflows can introduce biases at various steps. Satisfactory results can be assured only by thorough checking of the spatial statistical properties of input data and simulated results.

The three different approaches to e-facies simulation offer the user different types of control over the spatial properties. For example, the spatial relationships between the e-facies are controlled by the facies transition probabilities in the Plurigaussian approach, through the spatial structure preserved when probabilities are transformed appropriately for independent co-simulations, and when the petrophysical log responses are co-simulated with a full multivariate model. All three methods honoured the facies transition probabilities to a large degree. That is the stacking patterns of the e-facies where generally preserved.

Interestingly, the two continuous variable simulation procedures can be performed independent of grids and allow for estimates of the precision of the local e-facies using entropy, the mixture of facies, directly from the assignment probabilities. Entropy measures can be used as quality checks on sets of facies simulations. Lastly, the workflows demonstrate that using continuous property modeling to define stationary domains such as facies categories can address scale issues in special situations where new variables that do not arithmetically scale may be computed in 3D, i.e., geomechanical rock properties computed from petrophysical curves. The outcome considered for this special case is a model of geomechanical stratigraphy.

Acknowledgements

R. Mohan Srivastava, TriStar Gold

References

Babak, O., Manchuk, J.G., and Deutsch, C.V. [2013] Accounting for Non-Exclusivity in Sequential Indicator Simulation of Categorical Variables. *Computers & Geosciences*, **51**, 118 – 128.

Garner, D., [2020] Electrofacies, a guided machine learning, as a foundation for improving facies logs for the practice of geomodeling, (abstract) Geoconvention, Calgary, September 21-23.

Garner, D.L., Srivastava, R.M., and Yarus, J.Y. [2015] Modeling Three Ways from Electro-facies - Categorical, E-facies Probabilities, and Petrophysics with Assignment (abstract) EAGE Petroleum Geostatistics, Biarritz, September 7-11.

Garner, D. L., Woo, A. and Broughton P. [2009] Applications of 1D Electro-Facies Modeling. Paper presented at Frontiers + Innovation – 2009 CSPG CSEG CWLS Convention, Calgary, May 10 – 13.