Published in Water Resources Research 45, W08421, 1-13, 2009
which should be used for any reference to this work

Integrating collocated auxiliary parameters in geostatistical simulations using joint probability distributions and probability aggregation

Grégoire Mariethoz,[1] Philippe Renard,[1] and Roland Froidevaux[2]

We propose a new cosimulation algorithm for simulating a primary attribute using one or several secondary attributes known exhaustively on the domain. This problem is frequently encountered in surface and groundwater hydrology when a variable of interest is measured only at a discrete number of locations and when the secondary variable is mapped by indirect techniques such as geophysics or remote sensing. In the proposed approach, the correlation between the two variables is modeled by a joint probability distribution function. A technique to construct such relation using underlying variables and physical laws is proposed when field data are insufficient. The simulation algorithm proceeds sequentially. At each location of the domain, two conditional probability distribution functions (cpdf) are inferred. The cpdf of the main attribute is inferred in a classical way from the neighboring data and a model of spatial variability. The second cpdf is inferred directly from the joint probability distribution function of the two attributes and the value of the secondary attribute at the location to be simulated. The two distribution functions are combined by probability aggregation to obtain the local cpdf from which a value for the primary attribute is randomly drawn. Various examples using synthetic and remote sensing data demonstrate that the method is more accurate than the classical collocated cosimulation technique when a complex relation relates the two attributes.

1. Introduction

[2] There are numerous situations in surface and groundwater hydrology in which an attribute of interest is measured at a discrete number of locations and needs to be mapped accounting for exhaustive secondary information. For example, this is encountered (1) when interpolating ground based measurements of precipitation using a digital elevation model [*Goovaerts*, 2000] or radar observations [*Creutin et al.*, 1988; *Haberlandt*, 2007] as secondary information; (2) when estimating soil surface moisture and surface roughness from satellite images and ground measurements [*Makkeasorn et al.*, 2006]; (3) when characterizing the heterogeneity of an aquifer from local hydraulic conductivity data and an exhaustive geophysical survey [*Cassiani et al.*, 1998]; (4) when mapping groundwater recharge using local estimates and an exhaustive evaporation map produced from remote sensing data [*Brunner et al.*, 2004]; or (5) when mapping large scale groundwater contamination (such as Arsenic) using again local measurements in boreholes and exhaustive maps of geology and surface soil properties [*Winkel et al.*, 2008].

[3] The most simple way to solve this type of problem is to model the correlation between the primary and secondary attributes using a statistical regression (it does not need to be linear) and to rescale the exhaustive map. This is, for example, the procedure that was used by *Brunner et al.* [2004] for recharge estimate or by *Winkel et al.* [2008] to map Arsenic concentration. However, such a method does not enforce local accuracy (the interpolated value at the location of a ground based measurement may be different from the measurement itself) and does not account for the possibly known covariance structure of the primary variable. More flexible approaches are available in the framework of multivariate geostatistics [*Chilès and Delfiner*, 1999; *Wackernagel*, 2003]. Such methods were reviewed and compared in the field of hydrology, for example, by *Ahmed and De Marsily* [1987] or *Goovaerts* [2000]. They include a set of estimation techniques (providing the most likely value at any location) such as cokriging [*Matheron*, 1971], kriging with external drift [*Delhomme et al.*, 1981], collocated cokriging (W. Xu et al., Integrating seismic data in reservoir modeling: The collocated cokriging alternative, paper SPE 24742 presented at 67th Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, Washington, DC, 4–7 October, 1992), and stochastic simulation techniques (allowing to generate a set of equally

[1] G. Mariethoz and P. Renard, Centre for Hydrogeology, University of Neuchâtel, 11 Rue Emile Argand, CP 158, CH-2000 Neuchâtel, Switzerland. (gregoire.mariethoz@unine.ch)
[2] R. Froidevaux, Ephesia Consult, 9, rue Boissonnas, CH-1227 Geneva, Switzerland.

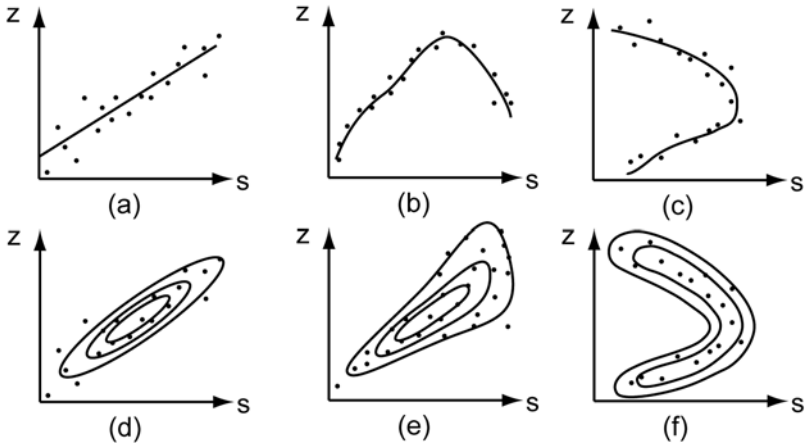

Figure 1. Schematical representation of the different types of possible relationships between a primary and secondary variable. *s*, secondary variable; *z*, primary variable. (a) Linear relationship modeled by a regression line. (b) Nonlinear relationship modeled by an analytical injective function relating $z = f(s)$. (c) Noninjective (or multivalued) relationship between *s* and *z*, which cannot be modeled by a function $z = f(s)$. (d) Linear relationship modeled by a bi-Gaussian probability distribution function. (e) Linear relationship showing a variation of the uncertainty around the regression line and which cannot be modeled by a bi-Gaussian pdf. (f) Noninjective relation modeled by a joint probability distribution.

probable maps) such as cosimulation [*Gomez-Hernandez and Journel*, 1993], collocated cosimulation, cosimulation with external drift, etc. These techniques are based on two ingredients: a model of cross covariance that relates the variability of the first variable at a given location with the variability of the secondary variable at another location, and the assumption that the relation between the two variables is essentially linear and can therefore be modeled in a multi-Gaussian framework. Furthermore, it is often found in practice that the secondary data located precisely at the location that needs to be estimated (or simulated) has a much stronger impact on the estimation (or simulation) than the data located aside. This is why when an exhaustive map of secondary information is available, a common simplification is to write a cokriging system that accounts only for the secondary variable at the location to be estimated (or simulated) and not at the neighboring nodes. These are the so-called collocated cokriging or cosimulation techniques (Xu et al., presented paper, 1992). It is an approximation that has the advantage to be much faster than the full cokriging but it is not optimal for all situations and can even be useless in certain cases. For example, when solving an inverse problem in a cokriging framework, the cross-covariance between heads and transmissivity under the uniform flow assumption shows that a head measurement has no impact (cross-covariance equal to zero) on the estimation of the transmissivity at the location of the head measurement [*Dagan*, 1989].

[4] In this paper, we do not consider the inverse problem. Our aim is to propose a new method for generating an ensemble of stochastic simulations of a primary attribute using a discrete number of local measurements of the primary attribute and an exhaustive map of the secondary attribute in a collocated cosimulation framework. The specificity of our approach is to consider complex relations between the two variables and to use a probability aggregation technique.

[5] Indeed, most papers presenting applications of geostatistical techniques (cokriging or cosimulations) to this problem assume a multi-Gaussian framework and a linear relationship (Figures 1a and 1d) between the two attributes. When the relation is not linear but can be modeled by an analytical function (e.g., Figure 1b), one can use a transformation of variable to linearize the relation. For example, this is the case when estimating the log hydraulic conductivity using geoelectrical surveys [*Cassiani et al.*, 1998; *El Idrisy and De Smedt*, 2007; *Slater and Lesmes*, 2002; *Soupios et al.*, 2007]. However, the assumptions of multi-Gaussianity and linear correlation between the variables (or their transforms) is too restrictive. It is an oversimplified description of complex physical processes and therefore it does not hold in many real-case applications. For example, we observed in a karstic coastal aquifer in Oman [*Alcolea et al.*, 2009] that karstic conduits had a high hydraulic conductivity and a low electrical resistivity if they were fully saturated with seawater. However, clayey deposits with very low hydraulic conductivity (that were also saturated with seawater) had a low electrical resistivity too. Higher electrical resistivites were associated to intermediate conductivity. A low electrical resistivity could therefore indicate either a high or a low hydraulic conductivity, but not an intermediate conductivity (Figure 1c). Such a relation was not only nonlinear (see Figure 1b for an example of nonlinear relation), it was essentially noninjective or multivalued (Figure 1c). We think that such situations are more frequent than usually admitted and that there are many cases in which there may be several possible and distinct values (or modes) of the primary parameter for a given value of the secondary parameter. This type of noninjective relations cannot be modeled by an analytical function of the secondary variable only nor by a simple change of variable. In addition, the possible variability of the primary variable knowing the secondary variable can vary as a function of the value of the secondary variable (heteroscedasticity) as shown in Figure 1e.

[6] A possibility to model the relation between the two variables in a completely general and statistical manner is to use a bivariate probability distribution function (pdf) $f(z, s)$.

This joint pdf can either be expressed analytically and parameterized by a type of distribution and its means, variances, correlation coefficient, etc., or it can be expressed as a fully nonparametric joint distribution (i.e. provided as a numerical matrix of probability values such as a bivariate histogram).

[7] To our knowledge, only a few techniques allow the use of complex bivariate models. The bilinear coregionalization model [*Wackernagel*, 2003] is a step toward integrating more complex relationships as it allows for non even covariance functions in the multi-Gaussian framework. The Successive Linear Estimator [*Yeh et al.*, 1996] is a modification of cokriging allowing to generate mean Gaussian parameters fields considering nonlinear parametric relationships. The so-called cloud transform technique [*Bashore et al.*, 1994; *Kolbjørnsen and Abrahamsen*, 2004] and the stepwise conditional transformation technique [*Leuangthong and Deutsch*, 2003], are closely related in the sense that they both impose a given bivariate correlation by performing appropriate normal score transformations of the variables. These methods rely also on a nonparametric description of the bivariate distribution between the main and secondary attributes thus allowing to accommodate situations in which the marginal distributions are multimodal and the relationship between the attributes is nonlinear. Simulated annealing [*Deutsch*, 1992] is another technique that could be used to impose various kinds of constraints, including nonparametric correlations [*Caers*, 2001]. Thanks to its flexibility, it has been applied to various practical cases [*Dafflon et al.*, 2008], even if it presents shortcomings such as a high CPU cost and a high parameterization of the cooling schedule.

[8] In all the methods cited above, a single conditional distribution function of the main attribute is estimated directly for each simulated grid node. Here we split the problem and estimate two separate probability distribution functions. The first one is estimated considering the primary variable only and a model of its spatial variability. In order to respect the nature of the first variable, that may or may not be multi-Gaussian, any suitable technique can be used (e.g., multi-Gaussian kriging [*Emery*, 2005; *Goovaerts et al.*, 2005; *Verly*, 1993], direct sequential simulation [*Soares*, 2001], indicator kriging [*Journel and Alabert*, 1990; *Journel and Isaaks*, 1984], multiple points statistics [*Hu and Chugunova*, 2008], etc.). The second one is the conditional distribution of the primary attribute computed from the bivariate distribution knowing the value of the secondary attribute. These two distributions are then combined into a single one using the concept of probability conjunction [*Tarantola*, 2005], which is a particular case of the theory of *Bordley* [1982] used in management science for aggregating expert's opinions. A similar approach was used by *Ortiz and Deutsch* [2004] to update the indicator kriging probability with multiple-points statistics.

[9] To apply the technique proposed in this paper, one needs first to model the bivariate distribution describing the relation between the two attributes. This is a prerequisite of the method. Because there are many statistical techniques available to do so, we will not focus on that aspect within the paper but just give some directions to help the user. Assuming that enough couples of values relating the primary and the secondary variable are available, one can statistically infer the parameters of any analytical joint distribution, or use techniques such as kernel smoothing to build a nonparametric distribution [*Epanechnikov*, 1969; *Kolbjørnsen and Abrahamsen*, 2004]. A thorough review of the techniques for estimating nonparametric density functions can be found in *Izenman* [1991]. Even if colocated data are available, the problem of estimating a statistical relationship between geophysical and hydrologic properties is not trivial. Several caveats remain, such as of scale issues [*Moysey and Knight*, 2004], local variations in the relationships and artifacts caused by inversion techniques [*Day-Lewis and Lane*, 2004]. Some authors have addressed these issues and proposed sophisticated and efficient techniques for estimating such relationships based on rock physics relationships and Monte Carlo simulations [*Moysey et al.*, 2005; *Mukerji et al.*, 2001]. When there is not enough data, we propose an alternative approach that consists in building the bivariate distribution from known physical laws and latent variables that indirectly relate the attributes of interest.

[10] The paper is structured as follows. The first part proposes a method for building a physically based joint probability distribution function when there is a lack of direct data. The second part describes the proposed simulation algorithm. The third part illustrates its application on a fully synthetic example. The fourth part presents an application of the method on a more realistic example based on satellite images. That case is used because, even if it has no direct application, it is analogous to real problems such as those described above and, more importantly, it is one of the rare situations that allows to test the accuracy of the method with real data showing a complex bivariate relation. This example is also used to show how to determine the optimal parameters of the simulation algorithm. Last comes a discussion of the overall method.

2. Inferring the Joint Distribution From Physical Laws and Latent Variables

[11] Denote: Note that upper cases S and Z represents random functions, while lower cases s and z represents actual values of these random functions.

$\mathbf{x}$	vector describing a location.
$Z(\mathbf{x})$	the attribute of main interest.
$S(\mathbf{x})$	the colocated attribute.
$z(\mathbf{x}_i)$, $i = [1 \ldots N]$	available conditioning data for the main attribute at location $\mathbf{x}_i$.
$f(z, s)$	bivariate joint probability density function.

[12] When there is not enough data available to statistically infer the joint distribution $f(z, s)$, one can build a realistic distribution from physical laws and latent variables. Often the variables of interest (z and s) are indirectly related to one (or several) underlying attribute t through physical laws: $z = \alpha(t)$ and $s = \beta(t)$. The underlying attribute t is the latent variable [*Bollen*, 1989]. Often there is sufficient data to estimate or assume the univariate pdf. To build the joint pdf $f(z, s)$, one can randomly sample a large number of values of t in $f(t)$, compute the corresponding values of z and s for each value of t and obtain a large number of couples $[z(t), s(t)]$ for estimating the joint pdf.

Table 1. Summary of the Parameters Used to Build the Joint pdf Displayed in Figure 2 for the Kölliken Site [*Mariethoz et al.*, 2009]

Facies	Proportion	bA_s	Mean $\log_{10} K$	σ $\log_{10} K$	Mean ϕ	Variance ϕ
RG	0.25	44,000	−5.95	1.46	0.209	0.003
DFR	0.36	166,100	−7.68	1.70	0.154	0.005
UW	0.10	700,000	−9.10	1.21	0.135	0.003
UPS	0.29	1,200,000	−9.56	0.38	0.112	0.003

[13] The following example illustrates this method. The primary attribute is the hydraulic conductivity and the secondary attribute is the electrical resistivity. To build a relation between these two attributes, we use porosity as the latent variable. The data used in this example originates from core samples of USM (Untere Süsswasser Molasse) level of the Swiss Molasse formation, in which four main facies have been identified. The overall composition of the USM is mainly marl and sandstone. Field data are described by *Mariethoz et al.* [2009].

[14] The first physical law is the Hagen-Poiseuille equation (1) relating hydraulic conductivity K [m/s] to porosity ϕ [−]:

$$K = \frac{\phi^3 \rho g}{b A_s^2 \mu}, \tag{1}$$

where b is the formation factor (usually between 10 and 20), A_s the specific contact surface between grains and water [m^2/m^3], μ the water viscosity fixed at 0.0027 [kg/m s], ρ the water density fixed at 999.7 [kg/m^3] (for freshwater at 10°C) and g the gravity acceleration, 9.81 [m^2/s]. This relationship has been investigated for the USM formation and available statistical data are shown in Table 1.

[15] The second physical law is Archie's law (equation (2)), relating the electrical resistivity of the fluid saturated rock R_t [Ωm] to the porosity [*Archie*, 1942]:

$$R_t = R_w \phi^{-m}, \tag{2}$$

where R_w is the fluid electric resistivity that is estimated to be around 5 [Ωm] in this formation [*Hug*, 2005] and m is traditionally defined as a constant usually varying between 1.5 and 2.5, depending on the geometry of the pores.

[16] The joint law is estimated by drawing samples of ϕ in the distributions presented in Table 1. For each ϕ sample, the couple [$K(\phi)$, $R_t(\phi)$] is computed. As m is unknown, it is randomly sampled for each porosity value in a uniform distribution whose bounds are 1.5 and 2.5. A large number of samples are drawn until their number is sufficient to obtain a stable joint distribution by kernel smoothing [*Epanechnikov*, 1969]. Figure 2a shows the resulting joint pdf, using 1,000,000 porosity samples. The specific characteristics of the different facies listed in Table 1 are visible in the different modes of the marginal pdf of hydraulic conductivity (Figure 2b). The marginal distribution of R_t (Figure 2c) corresponds to resistivity values measured in

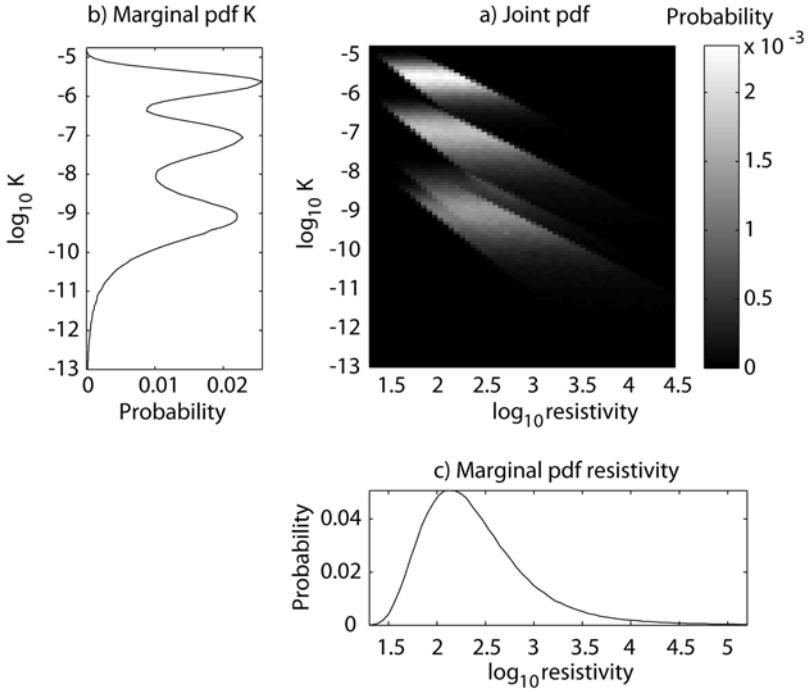

Figure 2. Inference of the joint distribution. (a) The joint pdf resulting of porosity sampling. (b) The marginal distribution of K. (c) The marginal distribution of R_t.

this part of the Molasse formation [*Jin et al.*, 1995; *Mächler*, 1994].

[17] This example shows that even if we assume very simple physical laws and statistical distributions for the porosity and basic parameters of these laws, we obtain a statistical relation between the two variables of interest which is complex and characterized by certain couples of values of log resistivity and hydraulic conductivity that are impossible while others are acceptable or even highly probable. For a given value of the resistivity such as 2, the bivariate distribution shows that a log hydraulic conductivity of -9 or -5.5 is acceptable but not a value of -7.5. It shows that it is possible to build the joint pdf with the proposed technique, but more important it shows also that multivalued relations can occur from simple physical relations and from a mixture of different rock types at a small scale. We would like also to emphasize that when considering only a single rock type, the joint relation which is obtained using that approach is very often extremely different from a multi-Gaussian distribution. There are, for example, very sharp boundaries between zones where some values are impossible and zones where the values have a certain probability of occurrence. We can expect that this type of relations are certainly very difficult to identify from a small number of samples that include measurement errors and this is why most often a multi-Gaussian model is used while it is most probably inadequate.

3. Simulation by Probability Aggregation

3.1. Outline of the Method

[18] Before discussing the details of the method, let us outline its main characteristics. First, it is a sequential simulation algorithm, and as such it is performed on a regular grid. The grid nodes are successively visited in a random order. For each successive node $\mathbf{x}$, two local cumulative conditional distribution functions (ccdf) are estimated. One, $F_1(\mathbf{x}, z)$, is the distribution function of the main attribute $Z(\mathbf{x})$ conditional to the neighboring data and the spatial correlation model:

$$F_1(\mathbf{x};z) = \text{Prob}\{Z(\mathbf{x}) \leq z|z(\mathbf{x}_1),\ldots,z(\mathbf{x}_N)\}. \quad (3)$$

This ccdf can be estimated using any suitable geostatistical method. In our examples, multi-Gaussian kriging is used.

[19] At the same location $\mathbf{x}$, the second distribution function, $F_2(\mathbf{x}; z|s)$, of $Z(\mathbf{x})$ conditional to the colocated attribute $s(\mathbf{x})$ can be obtained in a straightforward manner because $s(\mathbf{x})$ is informed at all locations of the grid. One just needs to extract it from the two-dimensional bivariate joint probability density function $f(z, s)$:

$$F_2(\mathbf{x};z|s) = \text{Prob}\{Z(\mathbf{x}) \leq z|S(\mathbf{x}) = s\} = \frac{\int_{\varsigma=-\infty}^{\varsigma=z} f(\varsigma, s(x))\, d\varsigma}{\int_{\varsigma=-\infty}^{\varsigma=+\infty} f(\varsigma, s(x))\, d\varsigma} \quad (4)$$

$F_1(\mathbf{x}; z)$ and $F_2(\mathbf{x}; z|s)$ provide two distinct pieces of information on the value that should be finally assigned to $Z(\mathbf{x})$. The issue is therefore to combine (aggregate) $F_1(\mathbf{x}; z)$ and $F_2(\mathbf{x}; z|s)$ into a single ccdf $F(\mathbf{x}; z|s)$ which would be an approximation of:

$$\text{Prob}\{Z(\mathbf{x}) \leq z|z(\mathbf{x}_1),\ldots,z(\mathbf{x}_N), S(\mathbf{x}) = s\}. \quad (5)$$

Once this ccdf is available for the location $\mathbf{x}$, a value is drawn from it and assigned to $Z(\mathbf{x})$. As usual in sequential simulation, $Z(\mathbf{x})$ is thereafter treated as conditional data for simulating the remaining unknown values of Z.

3.2. Estimating $F_1(\mathbf{x}; z)$ Using Multi-Gaussian Kriging

[20] Estimating a local ccdf $F_1(\mathbf{x}; z)$ given a set of conditioning data $z(\mathbf{x}_i)$, $i = [1 \ldots N]$ can be achieved by a variety of geostatistical techniques. Multi-Gaussian kriging [*Emery*, 2005; *Goovaerts et al.*, 2005; *Verly*, 1993] is probably the most widely used and well suited in the case of high entropy phenomena (as opposed to more structured, low entropy phenomena, that could better be described with methods such as multiple-point statistics). The main advantage of multi-Gaussian kriging is that it allows estimating for each location the complete pdf of the variable of interest even if its univariate distribution is not Gaussian. The approach consists of:

[21] 1. Performing a normal score transform of the conditioning data: $y(\mathbf{x}) = G[z(\mathbf{x})] \sim N \begin{cases} 0 \\ 1 \end{cases}$.

[22] 2. Assuming that this transformation is sufficient to ensure multi-Gaussianity.

[23] 3. Estimate the covariance function of the transformed data.

[24] 4. Estimating at node $\mathbf{x}$, by simple kriging, the mean and variance of the local conditioning Gaussian distribution m, σ^2.

[25] 5. Back-transformation of the entire local Gaussian distribution in order to obtain a distribution function that is not necessarily Gaussian. This is done numerically by applying the back-transformation $G^{-1}[y(\mathbf{x})]$ on equally spaced quantiles of the distribution found by simple kriging. The non-Gaussian local distribution is then reconstructed from its quantiles.

3.3. Probability Aggregation

[26] The problem of aggregating $F_1(\mathbf{x}; z)$ and $F_2(\mathbf{x}; z|s)$ is not trivial. Bayesian updating [e.g., *Woodbury and Ulrych*, 2000] could be used in this context, but it assumes conditional independence. Therefore it would require some knowledge of the relationship between the secondary variable at the current location $S(\mathbf{x})$ and the primary variable at all known locations, including all previously simulated nodes. According to *Journel* [2002], conditional independence is generally too strong an assumption in the context of sequential simulation. Indeed, the ccdfs defined in equations (3) and (4) are not conditionally independent because $F_1(\mathbf{x}; z)$ is based on previously simulated nodes that already integrated information on the joint distribution $f(z, s)$.

[27] Management science provides methods for aggregating expert's opinions while dealing with data interaction and without assuming conditional independence. *Bordley* [1982] proposed a formula for computing an aggregated probability density function $f^a(z)$ from n individual pdfs $f^k(z)$, $k = [1 \ldots n]$ (representing the various expert opinions) that satisfies

the basic probability properties and a series of axioms, including the weak likelihood ratio axiom:

$$f^a(z) = \frac{1}{\eta} f^0(z) \prod_{k=1}^{n} \left[\frac{f^k(z)}{f^0(z)}\right]^{w_k}, \quad (6)$$

where η is a normalizing factor. Each probability density function has a weight, $w_k \in \Re$ which can be seen as a way of quantifying redundancy or as a confidence factor. f^0 is the prior density function which is, in our case, the marginal pdf $f(z) = \int_{s=-\infty}^{s=-\infty} f(z, s)ds$ (i.e. the only information available on z when s is unknown). By construction, Bordley's formula has a certain number of important mathematical properties [see *Clemen and Winkler*, 2007, for a discussion], but nevertheless it is only an approximation allowing to combine several probabilities of the same event to occur (estimated from different sources of information) when the complete probability model including all the sources of information, and all the data interaction, is lacking. Equation (6) is closely related to the *tau* and *nu* models [*Journel*, 2002; *Krishnan*, 2005; *Polyakova and Journel*, 2007]. In the same spirit, *Tarantola* [2005] defines the conjunction of probability densities as the simplest way to aggregate probabilities in that manner:

$$f(\mathbf{x}; z|s) = f_1(\mathbf{x}; z) \wedge f_2(\mathbf{x}; z|s) = \frac{1}{\eta} \frac{f_1(\mathbf{x}; z) f_2(\mathbf{x}; z|s)}{f(z)}. \quad (7)$$

[28] Equation (7) is a particular case of equation (6) with two probabilities being aggregated with identical weights equal to 1. Nevertheless, such restrictions are not necessary. Using equation (6), it is possible to apply the method on any number of attributes and to adjust w_k in order to assign the relative weights to an expert-provided bivariate distribution model or to account for a spatially variable model of uncertainty by setting different weights values at different locations. Setting a weight to 0 at a certain location would result in the corresponding source of information having no influence.

[29] Our purpose is to aggregate $f_1(\mathbf{x}; z)$ and $f_2(\mathbf{x}; z|s)$ and we exposed various ways of achieving it. We do not want to favor a specific method among the ones mentioned above. In our examples, we illustrate the method using equations (6) and (7), but there are no restrictions regarding other techniques.

3.4. Step-by-Step Algorithm

[30] The proposed algorithm is implemented as follows:

[31] 1. Define the marginal cdf and spatial correlation model of $Z(\mathbf{x})$ as well as the bivariate model $f(z, s)$.

[32] 2. Each conditioning data is assigned to the closest grid node in the simulation grid (SG).

[33] 3. Define a path through the remaining nodes of the SG. The path is a vector containing all the indices of the grid nodes that will be simulated sequentially. Any type of path is suitable, for example, random or unilateral [*Daly*, 2004; *Pickard*, 1980].

[34] 4. For each successive location $\mathbf{x}$ in the path:

[35] a. Infer the local pdf $f_1(\mathbf{x}; z)$ from the known conditioning data in the neighborhood of $\mathbf{x}$. Any appropriate geostatistical method can be used for this purpose. In the following examples, we use multi-Gaussian kriging (see above for details).

[36] b. Extract $f_2(\mathbf{x}; z|s)$ from the bivariate model (that can be spatially dependant or not).

[37] c. Estimate $f(\mathbf{x}; z|s)$ by probability aggregation using Bordley's equation (6) or using the conjunction of probability (7) (i.e using weights equal to one). In the examples below we will first start with the conjunction of probability and later test the effect of varying the weights. Note that the *tau* or *nu* models could also be used here to aggregate $f_1(\mathbf{x}; z)$ and $f_2(\mathbf{x}; z|s)$ in order to obtain $f(\mathbf{x}; z|s)$.

[38] d. Randomly draw a sample $z'(\mathbf{x})$ from $f(\mathbf{x}; z|s)$, assign it to the location $\mathbf{x}$ in the grid and add it to the conditioning data set.

4. Synthetic Example

[39] In this section, we show an example of application of the proposed algorithm on a synthetic case in which the primary variable is related to a noisy secondary variable via a croissant shape bivariate distribution. The bivariate density function is known and the synthetic reference field for the primary attribute is created by unconditional simulation. The secondary attribute is constructed by drawing for each node $\mathbf{x}$ a value of the secondary variable s from the conditional distribution at that location knowing the previously simulated primary variable z:

$$F(\mathbf{x}; s|z) = \text{Prob}\{S(\mathbf{x}) \leq s|Z(\mathbf{x}) = z\}. \quad (8)$$

By construction, the secondary variable field is spatially correlated but rather noisy, as one can expect in real case examples. Then, the reference is sampled at 50 random locations. This is used as input conditioning data for the algorithm. The simulation grid size is 50 $\times$ 50 cells, and 100 realizations are generated. In order to evaluate the performance of the method, the simulations are compared to the reference which is known exhaustively. The comparison criteria are the reproduction of the histogram and variogram, the errors between the simulated values and the known reality, and the visual aspect of the simulations.

[40] Figure 3 illustrates the method. The bivariate distribution function is inspired from the Oman case described in the introduction. The primary attribute (Figure 3a) has a bimodal distribution (Figure 3g). Locations of randomly sampled data are marked by crosses (Figure 3a). The secondary attribute is noisy (Figure 3b) and is related to the primary attribute by a crescent-shaped joint pdf (Figure 3c). Despite the noise, the secondary attribute still contains enough information to guide the simulations (Figures 3d and 3e), where features of the reference are present at locations where no data are available (for example, the dark channel that runs through the field from left to right).

[41] The joint distribution (Figure 3f), the reference histogram (Figure 3g) and variogram (Figure 3h) are rather well reproduced (the solid red line represents the reference, dots represent the simulations and blue circles the sampled data). There is no systematic bias, as shown by the histogram of errors of the simulated attribute that is centered on 0 (Figure 3i).

[42] By construction, the relation between the secondary and primary variable is noninjective. As expected the proposed method allows to generate an ensemble of simu-

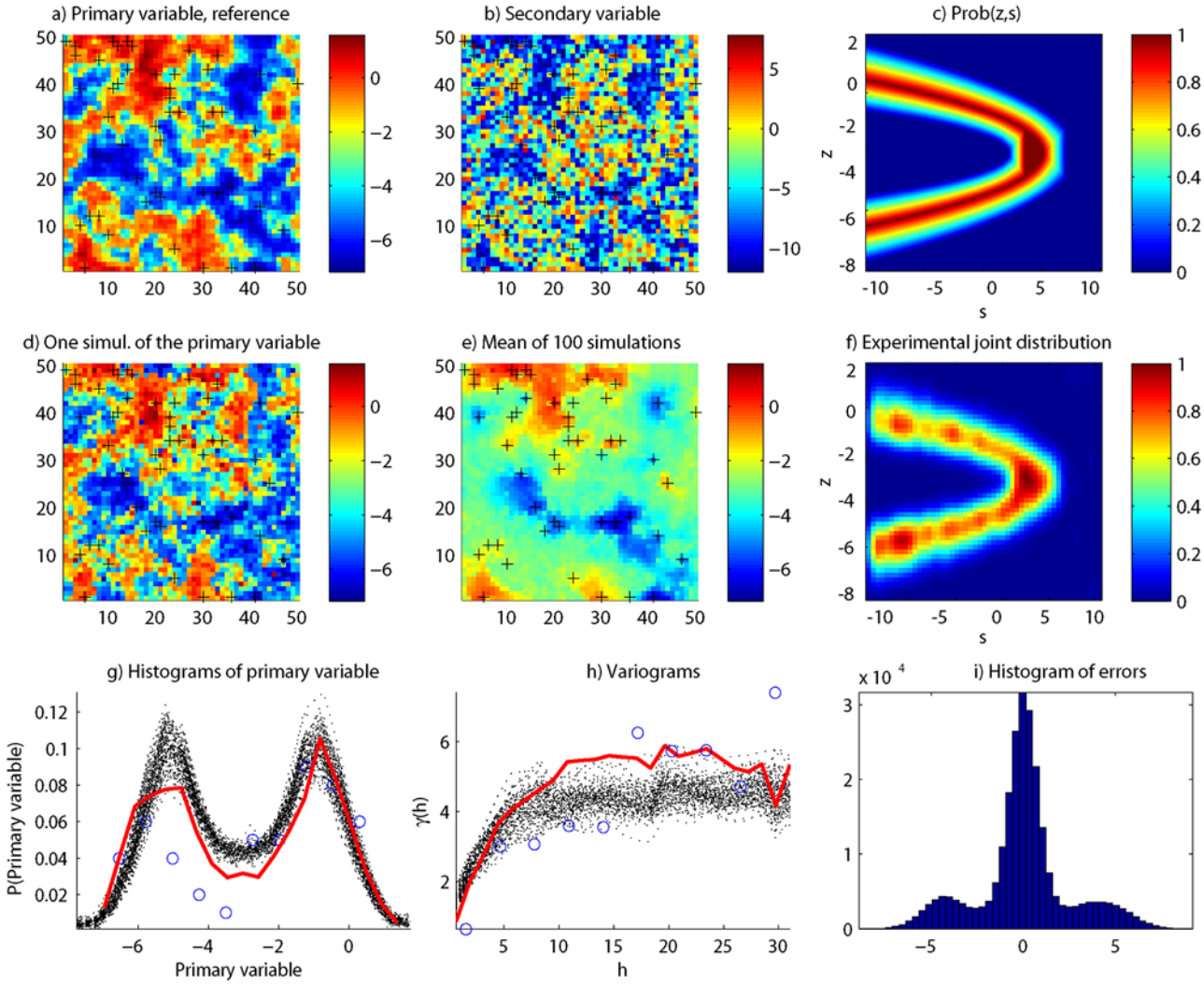

Figure 3. Synthetic example using a multivalued relation modeled by a custom joint pdf. The primary attribute has a spherical variogram model (sill = 5.3, range = 12, adjusted on the 50-sample data).

lations of the primary field that respect this very particular relation as shown by the reproduction of the joint pdf (Figure 3f). Among the simulations, at locations where one finds low values of the secondary variable, the primary variable is either extremely low or extremely high.

[43] In order to test whether the method could also be applied in more standard cases where a linear and multi-Gaussian relation holds, several tests were made with different correlation coefficients and variograms for the primary attribute. The accuracy of the method was compared to standard cosimulation algorithm. These tests are not shown in this paper for the sake of brevity, however they have all shown that the method performed as well as a traditional collocated cosimulation method. However, there is no advantage of using our method in those cases. When the multi-Gaussian assumption holds, it is wiser to use a full cosimulation or cokriging based methods because it accounts for the value of the secondary variable in the entire neighborhood, and not only at the location to simulate. Moreover, it takes advantage of a fully consistent model, whereas certain parameters of probability aggregation are subject to calibration (e.g., the weights of equation (6)).

5. Realistic Example

5.1. Testing the Method on Real Data

[44] In order to test further the method, a real data set has been used. It is based on two Landsat 7 satellite images corresponding to the same area, but taken at two different wavelengths (Figures 4a and 4b). One image is considered to be the primary attribute while the other is the secondary attribute. We decided to use such images first because satellite images are often used as secondary variables in hydrology, for example, for the estimation of soil moisture [*Makkeasorn et al.*, 2006; *Zribi et al.*, 2005] or for improving the mapping of ground based measurements of precipitation [*Haberlandt*, 2007]. However, more importantly a couple of satellite images constitutes a unique data set of two exhaustively known variables allowing to test the accuracy of the method with real data and not on a synthetic case. Indeed, exhaustive data sets are seldom available for

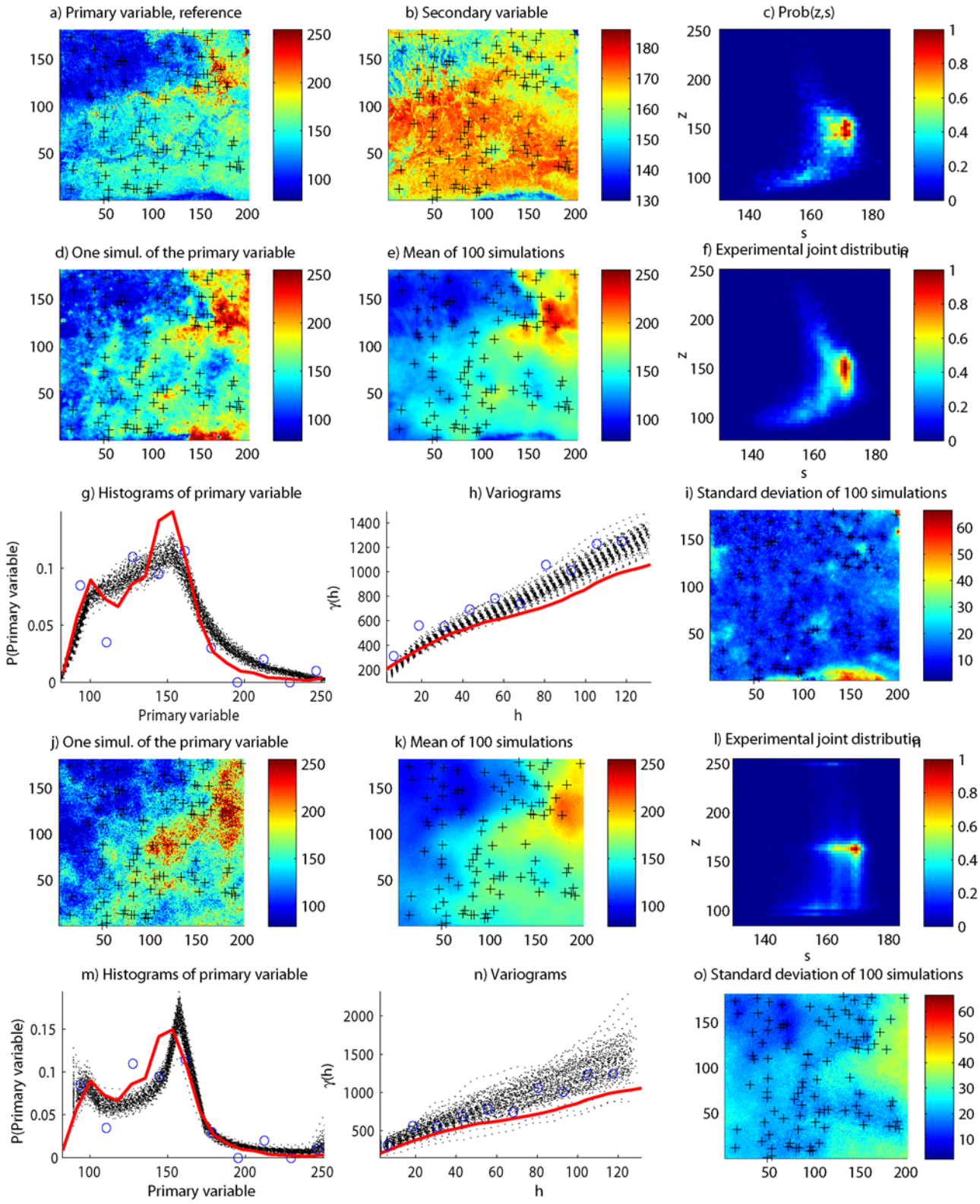

Figure 4. Real case example. (a–c) Input data. The primary attribute has an exponential variogram model (sill = 1110, range = 90, adjusted on the 100 sample data). (d–i) Results and validation criteria using the probability aggregation method. (j–o) Results and validation criteria using collocated co-simulation, using the same input data.

other variables that may be of higher interest in groundwater hydrology (exhaustive hydraulic conductivity map and exhaustive geophysical survey, for example). Of course, in practice we are generally not interested in estimating a satellite image at a given wavelength knowing the image at another wavelength and at a few discrete locations. The aim here is only to test the method on real data. Because different processes affect the absorption of light at different wavelengths, each image highlights different features of the land and the joint relationship is thus complex (Figure 4c). In summary, this data set constitutes an analogue to the real problems encountered in practice, but an analogue whose primary variable is exhaustively known and therefore an analogue allowing to evaluate the accuracy of the method.

[45] As for the synthetic examples, multi-Gaussian kriging is used for estimating $F_1(\mathbf{x}; z)$. The first image is used as a reference, sampled at 100 random locations, while the second is the auxiliary attribute. The size of the simulation grid is 181x201. Figures 4d to 4h present the results of the 100 simulations in the same fashion as the synthetic examples. Figure 4i shows the standard deviation of the stack of 100 simulations. The joint distribution used for the cosimulation (Figure 4c) is very well reproduced in the simulations (Figure 4f). The root mean square error (RMSE) of the simulated values compared to the reference field is 24.70.

[46] To compare the method against existing and well established cosimulation methods, we used the traditional collocated cosimulation technique [*Almeida and Journel*, 1994] to generate 100 simulations of the primary variable with the same input data: the same exhaustive secondary variable, the same colocated data points, adjusted cross variograms, plus the assumption of a linear correlation between the Gaussian transformed variables. The method was applied with care and all necessary Gaussian direct and back transformations were performed. The results are showed in Figures 4j to 4o, in the same manner as the previous figures, and the validation criteria are the same.

[47] The traditional cosimulation method provides good variogram and histogram reproduction, even if the reproduction is not as good as the results obtained with probability aggregation. The RMSE compared to the reference field is 29.46. Nevertheless, the reproduction of the bivariate joint probability density function is grossly inaccurate (Figure 4l). This is due to the violation of the assumption of linear correlation between both variables. The Gaussian transformations result in a relationship that is not linear, but still very far from the true relationship (Figure 4c).

[48] In terms of reproduction of the reference image, probability aggregation is able to reproduce detailed features (visible on the mean of the simulations Figure 4e) that are too specific to be inferred using only the primary variable data and its variogram. Only the secondary variable contains such detailed local information, but as the relationship is nonlinear, the traditional cosimulation approach is unable to provide a similar level of detail (compare Figures 4d and 4j to the reference, Figure 4a).

[49] The comparison of the standard deviation maps (Figures 4i and 4o) show how much information probability aggregation is able to get out of the secondary variable. Standard deviation map of traditional cosimulation (Figure 4o) is mainly related to the distance to the data points (high standard deviation when no data are present). Standard deviation maps issued from probability aggregation also show high uncertainty at locations where the secondary information carries a low information content (e.g., the lower part of the image where the secondary variable value can correspond to a wide range of values for the primary variable). At locations where the secondary variable is very informative, standard deviation is low, even in the absence of conditioning data.

5.2. Adjusting the Weights

[50] So far, both weights w_1 and w_2 have been kept identical and equal to 1. This is justified when the confidence related to each source of information as well as the data interaction are unknown [*Polyakova and Journel*, 2007]. Nevertheless, using equation (6) allows setting these parameters to modify the influence of each source of information. In the example described above, these parameters can be adjusted in order to fit the simulation on the reference field.

[51] A sensitivity analysis on w_1 and w_2 has been carried out by testing all pairs $[w_1, w_2]$, with each parameter varying from 0 to 3 with a step of 0.2, thus resulting in 256 possible pairs. A stack of 10 realizations of the primary attribute has been generated for each pair of weights. The mean error of each stack compared to the reference has been calculated. The results of this sensitivity analysis are shown graphically in Figure 5. For each stack, this error was evaluated using 4 criteria, consisting in the mean sum of squared differences between the reference and the simulations of the stack for the histogram (Figure 5a), the variogram (Figure 5b), the joint pdf (Figure 5c) and the values of the primary attribute (Figure 5d) at each grid node.

[52] The best fits are obtained by setting w_2 to a high value (about 2) and w_1 to a low value (about 0.5). This emphasizes the high local information content of the secondary attribute when the joint law is accurately estimated. However, using only $F_2(\mathbf{x}; z|s)$ (setting w_1 to 0) generates a result that depends only on the joint law, and which might be biased if this law is inaccurate. The information represented by $F_1(\mathbf{x}; z)$, although partially redundant with $F_2(\mathbf{x}; z|s)$, is also capital because it ensures a spatial consistency in the simulated field. Indeed, setting w_1 to 0 dramatically decreases the quality of the simulation. Moreover, $F_1(\mathbf{x}; z)$ contains both local and structural information, whereas $F_2(\mathbf{x}; z|s)$ contains local information only.

[53] This sensitivity analysis shows that the optimal weights depend on various factors that are difficult to foresee in practice if the true primary variable field is unknown. As a workaround, we propose to determine the goodness of a pair $[w_1, w_2]$ using cross validation [e.g., *Dubrule*, 1983]. The true and estimated values can be compared in several ways: constructing the bivariate plot of true versus estimated values, building the histogram of their differences or, simply, by mapping the differences. All these representations give us ways to assess the adequacy of the model. Two caveats, though: cross validation cannot prove that the model is right; it will only highlight its deficiencies. Secondly, in all rigor, the data to be used in the cross-validation stage should not be used for building the model.

[54] In practice, to determine the goodness of a pair $[w_1, w_2]$ using cross validation, we propose to proceed as follows:

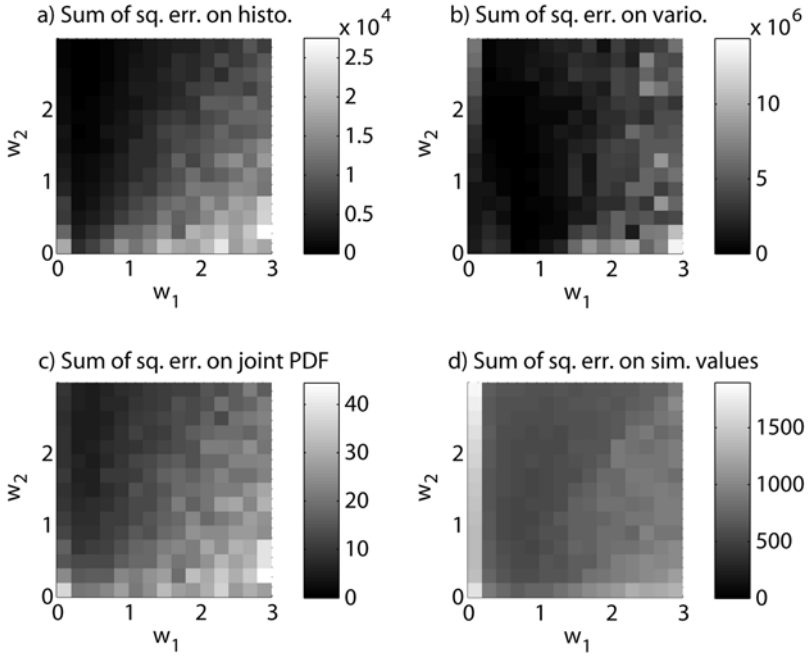

Figure 5. Sensitivity analysis of the value of the weights w_1 and w_2.

[55] 1. At the location $\mathbf{x}_i$ of each data point:

[56] a. Infer $f_1(\mathbf{x}_i; z)$ as previously, but without accounting for the fact that $z(\mathbf{x}_i)$ is actually known.

[57] b. Extract $f_2(\mathbf{x}_i; z|s)$ from the bivariate model.

[58] c. Estimate $f(\mathbf{x}_i; z|s)$ by probability aggregation using a given pair of weights $[w_1, w_2]$.

[59] d. Estimate $P^* = f(\mathbf{x}_i; z(\mathbf{x}_i)|s)$, the probability associated to the true value $z(\mathbf{x}_i)$.

[60] 2. Compute the mean probability of all data values $P^*_m = \frac{1}{N}\sum_i P^*(z(\mathbf{x}_i))$. In itself, this mean probability is not very informative, but it allows comparing various pairs of weights and determining which one yields the best results (i.e. where the true values are the most probable).

[61] P^*_m has been computed on the same pairs of weights $[w_1, w_2]$ as the sensitivity analysis. The results are displayed in Figure 6. The optimal weights found with cross validation are very similar to those found previously, but this time they were computed using the available data set only. This is very important for practical applications, where the reference field is never available.

[62] 100 additional simulations were generated using the weights obtained by cross validation, and the same input data as previously. Results are presented in Figure 7. The reference field is very well reproduced, with highly accurate histograms and variograms fit, good reproduction of the bivariate distribution and a RMSE of 25.37. The features of the reference field are well reproduced. More interesting is the standard deviation map, where all features specific to the secondary variable are highlighted. This is because more weight has been put on w_2, thus generating high variability when F_2 is not well determined.

6. Discussion and Conclusion

[63] Field observations made in a coastal aquifer in Oman [*Alcolea et al.*, 2009], statistical calculations using simple physical laws (Figure 2), or the relation between the magnitude of the signals observed at different wavelength on satellite images (Figure 4c) suggest that non multi-Gaussian and possibly noninjective relations need to be accounted for in a significant number of situations in surface and groundwater hydrology when interpolating certain primary variables using an exhaustive map of secondary information. To account for those type of relations, the simplest and most general approach is to model it using a joint probability distribution. In the examples that were used to illustrate and test the proposed methodology, we used a nonparametric joint distribution that offers a high degree of flexibility. However, one can also use an analytical expression for the joint distribution when such a model is available without any change in the algorithm.

[64] The proposed geostatistical simulation algorithm is an extension of the collocated cosimulation techniques. Its originality is that it is not based on an analytical and explicit model of the relationship between the various sources of information. Instead, it uses the joint probability density distribution to express at any location the conditional distribution of the primary attribute knowing the secondary

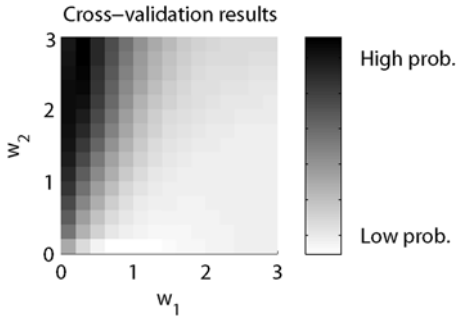

Figure 6. Optimal weights found by cross validation. The color scale represents the average probability of all measured data given by the model when a certain pair of weights $[w_1, w_2]$ is used.

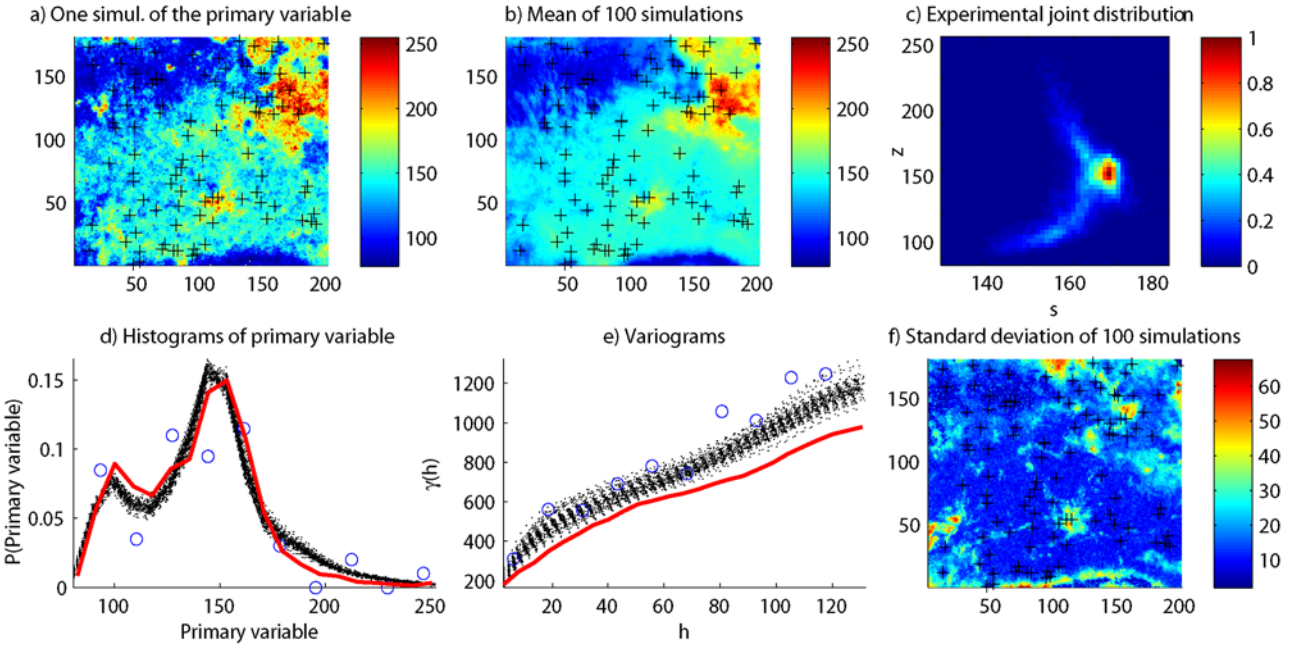

Figure 7. Real-case example using probability aggregation with optimal weights (w_1 = 0.5, w_2 = 2). The input data are the same as previously presented.

attribute and a model of spatial continuity for the primary attribute. These two models are assembled by a weighted probability aggregation technique.

[65] The main advantage of this method is that it allows to provide more accurate maps of the primary attribute than standard techniques when an exhaustive map of secondary information is available and when the relation between the two variables is better described by a joint probability distribution. Many hydrological applications could then benefit from that method. Another advantage is that it can easily be extended to multiple sources of information. A first possible extension is to use not only two local pdfs but as many local pdfs as needed. Adjusting the weights w_k is then a powerful way of parameterizing the method to aggregate secondary attributes having different information contents.

[66] The drawback of this flexibility is that finding the appropriate weights can be difficult. Keeping the weights equal to 1 does not assume conditional independence, but instead assumes the absence of data interaction, which has fewer consequences. However, a proper way of adjusting them is desired in order to improve the results. This is made difficult because the weights describe at the same time both concepts of confidence and redundancy. *Polyakova and Journel* [2007] suggest two methods to determine them: the first one is to calibrate the weights using available data, and the second one is to determine them using proxy cases. In our example, the first method gave similar weights to the ones found with a complete sensitivity analysis including a full knowledge of the true reference field. This tends to show that weights can be satisfactorily calibrated if enough hard data are present.

[67] An issue that was not considered in this work is that the relationship between the primary and the secondary variables can vary spatially [*Day-Lewis and Lane*, 2004] or as a function of a third or fourth auxiliary variables. A strength of our approach is that this additional information, when it is known, can be modeled using not only a bivariate joint probability distribution but a n-dimensional probability cube. At each location, one could compute the conditional probability density distribution of the primary variable knowing all the secondary variables known at that location. The rest of the method would remain unchanged. In doing so the method would accommodate spatially dependant statistical relationship between variables and any number of auxiliary attributes.

[68] One main limitation of the method is that it does not include an explicit model for joint spatial cross correlations between primary and secondary attributes. Only the spatial correlation of the primary attribute is modeled. We believe that this limitation is compensated by the flexibility of adjusting individual weights for an unlimited number of secondary attributes, and by the simplicity of the algorithm.

[69] The method has been tested using a multi-Gaussian model for the spatial continuity of the primary variable, but it can directly be extended to any sequential, pixel-based simulation technique that uses local conditional pdfs. Moreover, it is not limited to continuous attributes (for example, it can be used in the framework of multiple points statistics). Therefore its straightforward implementation makes it interesting to append probability aggregation on existing simulation codes.

[70] Finally, this paper has shown that the concept of probability conjunction or aggregation, originating from management science, is a precious tool for integrating information originating from diverse sources in problems related to the characterization of hydrological processes.

[71] **Acknowledgments.** Funding for this work was provided by the Swiss National Science Foundation (contract PP002-1065557). We thank Denis Allard (INRA) and three anonymous reviewers for their constructive comments, Albert Tarantola (Institut de Physique du Globe de Paris) for

enlightening lessons, and François Bertone (BCEOM Engineering) for initiating this project by giving us a real-case problem.

References

Ahmed, S., and G. De Marsily (1987), Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, *Water Resour. Res.*, *23*(9), 1717–1737.

Alcolea, A., P. Renard, G. Mariethoz, and F. Bretone (2009), Reducing the impact of a desalination plant using stochastic modeling and optimization techniques, *J. Hydrol.*, *365*(3–4), 275–288.

Almeida, A., and A. Journel (1994), Joint simulation of multiple variables with a Markov-type coregionalization model, *Math. Geol.*, *26*, 565–588.

Archie, G. (1942), The electrical resistivity log as an aid in determining some reservoir characteristics, *J. Pet. Technol.*, *5*, 1–8.

Bashore, W., U. Araktingi, M. Levy, and U. Schweller (1994), Importance of a geological framework for reservoir modelling and subsequent fluid-flow predictions, in *AAPG Computer Application in Geology*, edited by J. M. Yarus and R. L. Chambers, pp. 159–175, Am. Assoc. of Pet. Geol., Tulsa, Okla.

Bollen, K. A. (1989), *Structural Equations with Latent Variables*, 514 pp., Wiley, New York.

Bordley, R. E. (1982), A multiplicative formula for aggregating probability assessments, *Manage. Sci.*, *28*(10), 1137–1148.

Brunner, P., P. Bauer, M. Eugster, and W. Kinzelbach (2004), Using remote sensing to regionalize local precipitation recharge rates obtained from the chloride method, *J. Hydrol.*, *294*(4), 241–250, doi:10.1016/j.jhydrol.2004.02.023.

Caers, J. (2001), Automatic histogram and variogram reproduction in simulated annealing simulation, *Math. Geol.*, *33*(2), 167–190.

Cassiani, G., G. Böhm, A. Vesnaver, and R. Nicolich (1998), A geostatistical framework for incorporating seismic tomography auxiliary data into hydraulic conductivity estimation, *J. Hydrol.*, *206*(1–2), 58–74, doi:10.1016/S0022-1694(98)00084-5.

Chilès, J.-P., and P. Delfiner (1999), *Geostatistics — Modeling Spatial Uncertainty*, Wiley, New York.

Clemen, R. T., and R. L. Winkler (2007), Aggregating probability distributions, in *Advances in Decision Analysis: From Foundations to Applications*, edited by W. Edwards et al., pp. 154–176, Cambridge Univ. Press, Cambridge.

Creutin, J. D., G. Delrieu, and T. Lebel (1988), Rain measurement by raingage-radar combination: A geostatistical approach, *J. Atmos. Oceanic Technol.*, *5*(1), 102–115.

Dafflon, B., J. Irving, and K. Holliger (2008), Simulated-annealing-based conditional simulation for the local-scale characterization of heterogeneous aquifers, *J. Appl. Geophys.*, doi:10.1016/j.jappgeo.2008.09.010.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer, Berlin, Germany.

Daly, C. (2004), Higher order models using entropy, Markov random fields and sequential simulation, in *Geostatistics Banff 2004*, vol. 1, edited by O. Leuangthong and C. V. Deutsch, pp. 215–224, Kluwer Acad., Berlin.

Day-Lewis, F., and W. Lane (2004), Assessing the resolution-dependent utility of tomograms for geostatistics, *Geophys. Res. Lett.*, *31*, L07503, doi:10.1029/2004GL019617.

Delhomme, J.-P., M. Boucher, G. Meunier, and J. Jenson (1981), Apport de la géostatistique à la description des stockages de gaz en aquifère, *Rev. Inst. Fr. Pet.*, *36*(3), 209–327.

Deutsch, C. (1992), Annealing techniques applied to reservoir modeling and the integration of geological and engineering (well test) data, Ph.D. thesis, 202 pp., Stanford Univ., Stanford, Calif.

Dubrule, O. (1983), Cross-validation of kriging in a unique neighborhood, *Math. Geol.*, *15*(7), 687–699.

El Idrisy, E., and F. De Smedt (2007), A comparative study of hydraulic conductivity estimations using geostatistics, *Hydrogeol. J.*, *15*(3), 459–470.

Emery, X. (2005), Simple and ordinary multigaussian kriging for estimating recoverable reserves, *Math. Geol.*, *37*(3), 295–319.

Epanechnikov, V. A. (1969), Nonparametric estimation of a multidimensional probability density, *Theor. Probab. Appl.*, *14*, 153–158.

Gomez-Hernandez, J., and A. Journel (1993), Joint Simulation of MultiGaussian Random Variables, in *GeostatisticsTroia '92*, vol. 1, edited by A. Soares, pp. 85–94, Kluwer Acad., Dordrecht.

Goovaerts, P. (2000), Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *J. Hydrol.*, *228*(1–2), 113–129.

Goovaerts, P., G. AvRuskin, J. Meliker, M. Slotnick, G. Jacquez, and J. Nriagu (2005), Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan, *Water Resour. Res.*, *41*(7), W07013, doi:10.1029/2004WR003705.

Haberlandt, U. (2007), Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, *J. Hydrol.*, *332*(1–2), 144–157, doi:10.1016/j.jhydrol.2006.06.028.

Hu, L., and T. Chugunova (2008), Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review, *Water Resour. Res.*, *44*, W11413, doi:10.1029/2008WR006993.

Hug, R. (2005), Hydrogeologische untersuchungen im abstrom der sondermülldeponie kölliken, *Bull. Appl. Geol.*, *10*(1), 49–65.

Izenman, A. (1991), Recent developments in nonparametric density estimation, *J. Am. Stat. Assoc.*, *86*(413), 205–224.

Jin, J., T. Aigner, H. P. Luterbacher, G. H. Bachmann, and M. Müller (1995), Sequence stratigraphy and depositional history in the south-eastern German Molasse Basin, *Mar. Petrol. Geol.*, *12*(8), 929–940.

Journel, A. G. (2002), Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses, *Math. Geol.*, *34*(5), 573–596.

Journel, A., and F. Alabert (1990), New method for reservoir mapping, *J. Petrol. Technol.*, *42*(2), 212–218.

Journel, A., and E. Isaaks (1984), Conditional indicator simulation: Application to a Saskatchewan deposit, *Math. Geol.*, *16*(7), 685–718.

Kolbjørnsen, O., and P. Abrahamsen (2004), Theory of the cloud transform for applications, in *Geostatistics Banff 2004*, edited by O. L. a. C. V. Deutsch, pp. 45–54, Kluwer Acad., Berlin.

Krishnan, S. (2005), Combining diverse and partially redundant information in the earth sciences, Ph.D. thesis, Stanford Univ., Stanford, Calif.

Leuangthong, O., and C. Deutsch (2003), Stepwise conditional transformation for simulation of multiple variables, *Math. Geol.*, *35*(2), 155–173.

Mächler, E. (1994), Kombination verschiedener geophysikalischer Methoden (Refraktionsseismik und Widerstandsgeolektrik) zur Aufsuchung einer Molasserinne in Kölliken (AG), Diploma thesis, 54 pp., ETH Zürich, Zurich.

Makkeasorn, A., N. Chang, M. Beaman, C. Wyatt, and C. Slater (2006), Soil moisture estimation in a semiarid watershed using RADARSAT-1 satellite imagery and genetic programming, *Water Resour. Res.*, *42*, W09401, doi:10.1029/2005WR004033.

Mariethoz, G., P. Renard, F. Cornaton, and O. Jaquet (2009), Truncated plurigaussian simulations to characterize aquifer heterogeneity, *Ground Water*, *47*(1), 13–24, doi:10.1111/j.1745-6584.2008.00489.x.

Matheron, G. (1971), *The Theory of Regionalized Variables and Its Application*, Ecoles des Mines de Paris, Fontainebleau, France.

Moysey, S., and R. Knight (2004), Modeling the field-scale relationship between dielectric constant and water content in heterogeneous systems, *Water Resour. Res.*, *40*, W03510, doi:10.1029/2003WR002589.

Moysey, S., K. Singha, and R. Knight (2005), A framework for inferring field-scale rock physics relationships through numerical simulation, *Geophys. Res. Lett.*, *32*, L08304, doi:10.1029/2004GL022152.

Mukerji, T., A. Jorstad, P. Avseth, G. Mavko, and J. Granli (2001), Statistical rock physics: Combining rock physics, information theory, and geostatistics to reduce uncertainty in seismic reservoir characterization, *The Leading Edge*, *20*(3), 313–319, doi:10.1190/1.1438938.

Ortiz, J. M., and C. V. Deutsch (2004), Indicator simulation accounting for multiple-point statistics, *Math. Geol.*, *36*(5), 545–565.

Pickard, D. (1980), Unilateral Markov fields, *Adv. Appl. Probab.*, *12*, 655–671.

Polyakova, E., and A. Journel (2007), The Nu expression for probabilistic data integration, *Math. Geol.*, *39*(8), 715–733.

Slater, L., and D. Lesmes (2002), Electrical-hydraulic relationships observed for unconsolidated sediments, *Water Resour. Res.*, *38*(10), 1213, doi:10.1029/2001WR001075.

Soares, A. (2001), Direct sequential simulation and cosimulation, *Math. Geol.*, *33*(8), 911–926.

Soupios, P., M. Kouli, F. Vallianatos, A. Vafidis, and G. Stavroulakis (2007), Estimation of aquifer hydraulic parameters from surficial geophysical methods: A case study of Keritis Basin in Chania (Crete-Greece), *J. Hydrol.*, *338*(1–2), 122–131.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Parameter Estimation*, Soc. Ind. Appl. Math., Philadelphia, Pa.

Verly, G. (1993), Sequential Gaussian cosimulation — a simulation method integrating several types of information, in *Geostatistics Troia '92*, vol. 1, edited by A. Soares, pp. 543–554, Kluwer Acad., Dordrecht.

Wackernagel, H. (2003), *Multivariate Geostatistics: An Introduction With Applications*, 3rd ed., 387 pp., Springer-Verlag, Berlin.

Winkel, L., M. Berg, M. Amini, S. J. Hug, and C. A. Johnson (2008), Predicting groundwater arsenic contamination in Southeast Asia from surface parameters, *Nat. Geosci.*, *1*(8), 536–542, doi:10.1038/ngeo254.

Woodbury, D., and T. Ulrych (2000), A full-Bayesian approach to the groundwater inverse problem for steady state flow, *Water Resour. Res.*, *36*(8), 2081–2093.

Yeh, J., M. Jin, and S. Hanna (1996), An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields, *Water Resour. Res.*, *32*(1), 85–92.

Zribi, M., N. Baghdadi, N. Holah, and O. Fafin (2005), New methodology for soil surface moisture estimation and its application to ENVISAT-ASAR multi-incidence data inversion, *Remote Sens. Environ.*, *96*(3–4), 485–496.