

Chapter 1

Kriging, Splines, Conditional Simulation, Bayesian Inversion and Ensemble Kalman Filtering



Olivier Dubrule

Abstract This chapter discusses, from a theoretical point of view, how the geostatistical approach relates to other commonly-used models for inversion or data assimilation in the petroleum industry. The formal relationship between point Kriging and splines or radial basis functions is first presented. The generalizations of Kriging to the estimation of average values or values affected by measurement errors are also addressed. Two algorithms are often used for conditional simulation: the “rough plus smooth” approach consists of adding a smooth correction to a non-conditional simulation, whilst sequential Gaussian simulation allows the point-by-point construction of the realizations. As with Kriging, conditional simulation can be applied to average values or to data affected by measurement errors. Geostatistical inversion generates high-resolution realizations of vertical impedance traces constrained by seismic amplitudes. If the relationship between impedance and amplitude data is linearized, geostatistical inversion is a particular case of Bayesian inversion. Because of the non-linearity of production data vis-à-vis the variables of the earth model, their assimilation is harder than that of seismic data. Ensemble Kalman filtering, if considered from a geostatistical viewpoint, consists of using a large number—or ensemble—of realizations to calculate empirical covariances between the dynamic data and the parameters of the geostatistical model. These covariances are then used in the equations for interpolating the mismatch between simulated and new production data using a coKriging-like formalism. Interestingly, most of these techniques can be expressed using the same generic equation by which an initial model not honouring some newly arrived data is made conditional to these data by adding a (co-)Kriged interpolation of the data mismatches to the initial model. In spite of their similar equations, Bayesian inversion, geostatistics and ensemble Kalman filtering have a different approach to the inference of the covariance models used by these equations.

Keywords Dual Kriging • Radial basis functions • Geostatistical inversion
Energy-based methods • Prediction error filter

O. Dubrule (✉)
Imperial College London, London, UK
e-mail: o.dubrule@imperial.ac.uk

© The Author(s) 2018
B. S. Daya Sagar et al. (eds.), *Handbook of Mathematical Geosciences*,
https://doi.org/10.1007/978-3-319-78999-6_1

1.1 Introduction

Fifty years ago, when geostatistics was pioneered by Matheron (1971), its main applications were Kriging and the change of support for mining applications. At the time, geostatistics was presented as a new discipline, without much reference to its relationships with other mathematical interpolation and modeling techniques. This has now changed as the relationships between geostatistics and such techniques as splines, regularization, Bayesian inversion, or ensemble Kalman filtering have become clearer. This convergence is fascinating and has led to many significant developments allowing the integration of multi-disciplinary data into 3-D geostatistical earth models.

This chapter discusses approaches for generating 2-D or 3-D subsurface models constrained by geological (wells), seismic or dynamic data. In spite of the wealth of data available, the uncertainty on the 3-D earth model remains high in most cases. Approaches that are designed to generate one unique “deterministic” model often pick the smoothest one. This is not realistic in situations where the Earth Model is used for flow simulation, as the results are biased if the model heterogeneities are not representative of that of the actual reservoir. More generally, non-linear operations, such as the application of cut-offs, may give biased results when applied to deterministic smooth models such as those produced by Kriging.

The multi-realization approach is now routinely applied to subsurface parameters inversion. *Looking at the mean provides much less information than looking at a movie of realizations. ...By construction, each of the realizations captures the essential random fluctuations of the actual field from which the data were extracted* (Tarantola 2005). This is a fundamental change. The traditional inversion approach could be formulated as “How to find an estimate of the spatial parameters which is as close as possible to the first guess values of these parameters and which provides, through forward modeling, an output which is as close as possible to the available data” (modified from Evensen 2007). These first guess values are usually a smooth (Kriging-like) spatial model of these parameters. Now the question has changed to “Find the probability density function (pdf) of 3-D models constrained by all the existing data, and provide techniques for sampling realizations from this pdf”.

This chapter, written from a geostatistical perspective, discusses the convergence between the existing techniques.

Deterministic approaches such as Kriging, splines, regularization- or energy-based methods generate a single model of the subsurface, which usually minimizes or maximizes an optimisation criterion. These approaches are closely related and their formal relationships are discussed.

Geostatistical simulation is then revisited, and two key simulation algorithms are discussed; The first one is sequential Gaussian simulation and the second one is the “rough plus smooth” combination of an unconditional simulation plus a smooth correction term. These two algorithms have helped bridge the gap between geostatistics and inversion.

Two successful approaches are then discussed for integrating seismic and dynamic data into the earth model. Rather than using an approach merely based on statistical correlations between data and model parameters, it is assumed that there exists a deterministic relationship (or forward model) between model parameters and data, possibly including a random error.

The first approach, geostatistical inversion, produces reservoir-scale models of acoustic or elastic parameters constrained by single- or multi-offset seismic amplitude data. The value of using sequential Gaussian simulation to calculate seismically-constrained realizations is discussed. In situations where the forward model is linear, geostatistical inversion can be formulated as a particular case of Bayesian seismic inversion.

The second approach, ensemble Kalman filtering, consists of sequentially updating an “ensemble” of geostatistical realizations using dynamic data as they are acquired in time. The key idea here is to statistically derive the covariance terms of the equation used in Bayesian inversion from an ensemble of realizations rather than from a theoretical covariance model. The formal relationship between ensemble Kalman filtering and co-Kriging is discussed.

Most of the above techniques can be shown to use the same kind of formalism, where the mismatch between newly arrived data and the current model is interpolated and used to update this model.

One of the conclusions of this chapter is that the equations of Bayes, geostatistics or ensemble Kalman filtering are closely related. However, this relationship is mostly formal as the three techniques differ in their approach to the covariances used in the equations. Geostatisticians first fit a model to the data, whilst Bayesians start from a model based on general “prior” information. Only later in the process do they introduce the well data. And ensemble Kalman filtering directly uses the experimental covariances calculated from the realizations of the ensemble.

The topic of joint inversion of seismic and dynamic data is not discussed here, in spite of the interesting on-going developments in 4-D seismic data inversion. This is because the objective of this chapter is to address formal relationships between the different formalisms rather than discuss specific applications.

1.2 Deterministic Aspects of Geostatistics

1.2.1 *Simple Stationary Kriging*

The basic model used by geostatistics is that of stationary random functions of order 2: a spatial property $z(\mathbf{x})$ at location \mathbf{x} is represented by a random function $Z(\mathbf{x})$, which is assumed to follow a trend $m(\mathbf{x})$ and a stationary covariance $C(\mathbf{h})$

$$m(\mathbf{x}) = E(Z(\mathbf{x})) \quad (1.1a)$$

$$C(\mathbf{h}) = E(Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h})) - E(Z(\mathbf{x}))E(Z(\mathbf{x} + \mathbf{h})) \quad (1.1b)$$

At each unsampled location \mathbf{x} , the value of $Z(\mathbf{x})$ is estimated by a linear combination $Z_k(\mathbf{x})$ of the values $Z_i = Z(\mathbf{x}_i)$ at the n data points $(\mathbf{x}_i)_{i=1, \dots, n}$. Kriging is the best linear unbiased estimator, in the sense that it is unbiased and that it minimizes the estimation variance. If the trend $m(\mathbf{x})$ is known at each location \mathbf{x} , the simple Kriging (Chilès and Delfiner 2012, p. 151) system of equations is obtained

$$Z_k(\mathbf{x}) - m(\mathbf{x}) = \sum_{i=1}^n \lambda_i (Z_i - m(\mathbf{x}_i)) \quad (1.2a)$$

$$\text{with } \sum_{i=1}^n \lambda_i C(\mathbf{x}_i - \mathbf{x}_j) = C(\mathbf{x} - \mathbf{x}_j) \text{ for } j \in (1, \dots, n) \quad (1.2b)$$

1.2.2 Kriging with Intrinsic Random Functions of Order k

Matheron (1973) generalized the above model to that of Intrinsic Random Functions of Order k (IRF- k), where the definition of the variogram as a generalized covariance of order zero and of generalized covariances of order k leads to a model based on the stationarity of generalized increments of order k .

With k -IRFs, the model only considers linear combinations of $Z(\mathbf{x})$ that filter polynomials of order k (such polynomials being likely to represent a trend). Simple Kriging is not applicable any more. For instance, if $k = 1$ in two dimensions, and if $K(\mathbf{h})$ designates the generalized covariance of order k (GC- k), the kriging system becomes

$$Z_k(\mathbf{x}) = \sum_{i=1}^n \lambda_i Z_i \quad (1.3a)$$

$$\text{with } \sum_{i=1}^n \lambda_i K(\mathbf{x}_i - \mathbf{x}_j) + \mu_0 + \mu_1 x_{j1} + \mu_2 x_{j2} = K(\mathbf{x} - \mathbf{x}_j) \text{ for } j \in (1, \dots, n) \quad (1.3b)$$

$$\text{and } \sum_{i=1}^n \lambda_i = 1 \quad \sum_{i=1}^n \lambda_i x_{i1} = x_1 \quad \sum_{i=1}^n \lambda_i x_{i2} = x_2$$

where the coordinates of each point \mathbf{x} of the plane are written as $\mathbf{x} = (x_1, x_2)$.

1.2.3 Kriging Extensions

The goal here is not to discuss the details of Kriging, as there are plenty of excellent textbooks for this (Chilès and Delfiner 2012, p. 150). However, two features of Kriging deserve to be discussed, as they facilitate the understanding of the relationship between Kriging, splines and Bayesian approaches.

1.2.3.1 Generalization of Kriging to the Interpolation of Average Values

Kriging is a linear interpolator. The data used by Kriging do not have to be point values, but they can be any linear function of the parameters of interest; Hansen et al. (2006) call these “volume support data”. In particular, Kriging can be used to estimate the average value of a parameter $Z(v_{\mathbf{x}})$ at a location \mathbf{x} by a linear combination volume of support data $Z(v_{\mathbf{x}_i})$ (Chilès and Delfiner 2012, p. 198)

$$Z_k(v_{\mathbf{x}}) = \sum_{i=1}^n \lambda_i Z(v_{\mathbf{x}_i}) \quad (1.4)$$

This property of Kriging, extensively used in mining applications, is of significant interest in the context of linear inversion of volume support data (Hansen et al. 2006). The Kriging equations associated with Eq. 1.4 are not given here, as they are a bit heavy, but conceptually simple thanks to the linear property of Kriging.

1.2.3.2 Error CoKriging

Error coKriging (Dubrule 2003) is a generalization of Kriging to the situation where measurements Y_i of the parameter Z_i at data points \mathbf{x}_i are affected by an unbiased random error

$$Y_i = Z_i + \varepsilon_i \text{ with } E(\varepsilon_i) = 0 \text{ and } \text{Var}(\varepsilon_i) = C_{\varepsilon_i} \quad (1.5)$$

In this situation, error coKriging allows the estimation of $Z(\mathbf{x})$ at any unsampled location \mathbf{x} from a linear combination of values Y_i (the random measurement error attached to each data can be zero or not) (Dubrule 2003; Hansen et al. 2006; Chilès and Delfiner 2012, p. 216)

$$Z_k(\mathbf{x}) = \sum_{i=1}^n \lambda_i Y_i \quad (1.6)$$

1.2.3.3 Dual Kriging

If a global neighborhood is used, that is if all the available data are used to estimate $Z(\mathbf{x})$ at every single location \mathbf{x} , the Kriging equations (Eq. 1.3) can be inverted to obtain the dual Kriging system (for interpolation in the case of Kriging and smoothing in the case of error coKriging). For example, in two dimensions for a k-IRF of order 1

$$z_k(\mathbf{x}) = z_k(x_1, x_2) = a_0 + a_1x_1 + a_2x_2 + \sum_{i=1}^n b_i K(\mathbf{x} - \mathbf{x}_i) \quad (1.7)$$

where the conditions on the coefficients $(a_0, a_1, a_2, b_1, \dots, b_n)$ are different for Kriging and error coKriging (Dubrule 1983)

$$\text{Kriging: } \sum_{i=1}^n b_i = \sum_{i=1}^n b_i x_{i1} = \sum_{i=1}^n b_i x_{i2} = 0 \quad \text{and} \quad z_k(x_{i1}, x_{i2}) = z_i \quad (1.8)$$

$$\text{Error coKriging: } \sum_{i=1}^n b_i = \sum_{i=1}^n b_i x_{i1} = \sum_{i=1}^n b_i x_{i2} = 0 \quad \text{and} \quad z_k(x_{i1}, x_{i2}) + b_i C_{\varepsilon_i} = y_i \quad (1.9)$$

1.2.4 Kriging and Splines

1.2.4.1 Interpolating Splines

Splines are a popular method for deterministic interpolation and approximation (Micula and Micula 1999). In 2-D, interpolating splines calculate a function honouring the data and minimizing an energy functional. Harmonic splines minimize the stretching energy of a membrane while biharmonic splines minimize the bending energy of an elastic plate. The biharmonic spline function can be written using a similar expression as Eq. 1.7 (Duchon 1975), but with a specific model for the generalized covariance function

$$K(\mathbf{x} - \mathbf{x}_i) = \left((x_1 - x_{i1})^2 + (x_2 - x_{i2})^2 \right) \text{Log} \left(\sqrt{(x_1 - x_{i1})^2 + (x_2 - x_{i2})^2} \right) \quad (1.10)$$

Splines and Kriging are a particular case of a more general class of interpolators, called radial basis functions (Billings et al. 2002a, b). With splines, the polynomial in Eq. 1.7 belongs to the kernel of the operator T that is minimized by the spline function (T is the gradient for harmonic splines and the laplacian for biharmonic

splines), whilst the function $K(\mathbf{h})$ is the Green function associated with the operator $T'T$, where T' is the transposed operator of T (Matheron 1981a)

$$T'TK(\mathbf{h}) = \delta \quad (1.11)$$

where δ is the Dirac Function. Choosing the energy functional minimized by splines is equivalent to fixing the degree of the trend function and the generalized covariance model for Kriging. For harmonic splines, these are respectively a constant and the De Wijs variogram in *Logh* (Chilès and Delfiner 2012, p. 94).

The consequence of Eq. 1.11 on the spectral density of the generalized covariance $K(\mathbf{h})$ is straightforward. For example, the spectral densities associated with the harmonic and biharmonic splines are power laws, representing fractal models. Szeliski and Terzopoulos (1989) and Micula and Micula (1999) discuss this relationship between Splines and fractals.

1.2.4.2 Smoothing Splines

Smoothing splines are used in situations where measurements at data points are affected by a random error (Eq. 1.5). In two dimensions, they compute a function $f(x_1, x_2)$ minimizing the sum of a spline energy functional plus a weighted distance to the n data

$$\|Tf\|^2 + \theta \sum_{i=1}^n \frac{(f(x_{i1}, x_{i2}) - y_i)^2}{C_{\varepsilon_i}} \quad (1.12)$$

The smoothing biharmonic spline function has the same expression as that of Kriging and error coKriging (Eq. 1.7) but with the following relationships

$$\sum_{i=1}^n b_i = \sum_{i=1}^n b_i x_{i1} = \sum_{i=1}^n b_i x_{i2} = 0 \quad \text{and} \quad f(x_{1i}, x_{2i}) + b_i \frac{C_{\varepsilon_i}}{\theta} = y_i \quad (1.13)$$

Smoothing biharmonic splines are identical to error Cokriging as long as the generalized covariance used by error Cokriging is the function $\theta K(\mathbf{x} - \mathbf{x}_i)$, where $K(\mathbf{x} - \mathbf{x}_i)$ is given by Eq. 1.10 (Matheron 1981a; Dubrule 2003). This is a general relationship between smoothing splines and coKriging, which are formally equivalent if the generalized covariance $K(\mathbf{h})$ is that satisfying Eq. 1.11, with the coefficient of $K(\mathbf{h})$ equal to the smoothing parameter θ of Eq. 1.12.

1.2.4.3 Kriging and Regularization—The Discrete Case

The discrete case is the situation where interpolation is performed at the nodes of a regular grid and each data point is located at one of the nodes of this grid. If p is the total number of grid nodes, the number n of data points is such that $n < p$.

In the discrete case, Matheron (1981b) also demonstrated the equivalence between splines and Kriging, and between smoothing splines and error coKriging. Both the Kriged and spline values z_u minimize

$$\sum_{u,v=1}^p z_u B_{uv} z_v + \sum_{i=1}^n \frac{(z_i - y_i)^2}{C_{\epsilon_i}} \quad (1.14)$$

where the u and v indices designate all the p grid points where the interpolation takes place, whilst i indices designate the n data points. The minimization of Eq. 1.14 is performed according to the unknown values z_u at all grid nodes (including those unknown values z_i where a data point with measured value y_i is present). The first term of Eq. 1.14 can be interpreted as a quadratic energy function traditionally used in inverse problems. In the regularization context, the choice of this quadratic form is driven by smoothing considerations, often using Briggs' finite difference Laplacian (or spline) "roughening" operator (Briggs 1974; Bolondi et al. 1976). Seen from the geostatistical perspective, B_{uv} is the inverse of the covariance matrix in the stationary case and a pseudo-inverse of the generalized covariance matrix in the k-IRF case (Matheron 1981b). Equation 1.14 confirms the clear relationship between the inverse of the (generalized) covariance and the spline differential operator.

Kriging can thus be formalized in the frame of energy-based estimation techniques such as splines. This comes from the relationship between the inverse of the covariance function and the roughening filter implicit in the quadratic regularization term. It will be shown below that the regularization term can also be regarded, in the Bayesian inversion context, as an expression of the prior knowledge about the variable under study.

1.2.5 Kriging and Bayesian Inversion

1.2.5.1 Bayesian Linear Inversion

Here it may be useful to recall the general expression of the posterior mean and covariance in the case of Bayesian linear inversion of a multigaussian function. A very good reference for this is Tarantola (2005).

In the discrete case, consider a stationary multigaussian random vector z of dimension p containing the grid values z_u over a two or three-dimensional regular grid of size p . Assume also that a vector y contains the n data y_i . It is assumed again

that the data are affected by an error vector ε of dimension n , and also that these data are a linear function of the p values of z over the grid

$$y = Fz + \varepsilon \quad (1.15)$$

where the vector ε has mean zero and covariance matrix C_ε and F is a matrix of dimension $n \times p$. In the multigaussian case, thanks to the Bayes formula relating the posterior pdf $f_{post}(z)$ to the prior pdf $f_{prio}(z)$ and the likelihood function $g(y|z)$, the prior mean vector m (dimension p) and covariance matrix C (dimension $p \times p$) of z are updated using the information brought by the data vector y

$$f_{post}(z) \propto f_{prio}(z) g(y/z) \propto \exp\left[(z-m)'C^{-1}(z-m)\right] \times \exp\left[(y-Fz)'C_\varepsilon^{-1}(y-Fz)\right] \quad (1.16)$$

$f_{post}(z)$ is a multigaussian function with the mean vector

$$m_{post} = m + CF'(FCF' + C_\varepsilon)^{-1}(y - Fm) \quad (1.17)$$

and the covariance matrix

$$C_{post} = C - CF'(FCF' + C_\varepsilon)^{-1}FC \quad (1.18)$$

1.2.5.2 Kriging and Bayesian Inversion

Equation 1.17 can also be written

$$m_{post} = m + \Lambda(y - Fm) = (I - \Lambda F)m + \Lambda y \quad (1.19)$$

with

$$\Lambda = CF'(FCF' + C_\varepsilon)^{-1} \quad (1.20)$$

It can be checked that Λ is also the $p \times n$ matrix giving at each line u the n simple Kriging (or error coKriging) weights associated with the Kriging of the value z_u at node u . Comparing the first part of Eq. 1.19 with Eq. 1.2 shows that, in the multigaussian case, m_{post} is equal to simple Kriging and that the matrix C_{post} contains the variances and covariances of simple Kriging at each node u of the regular grid.

1.2.6 Energy-Based Versus Probabilistic Estimates

The minimization of Eq. 1.14 leads to either Kriging or splines if the (inverse of) the covariance (Kriging) and the differential operator (splines) are properly chosen. Minimizing the expression in Eq. 1.14 is equivalent to maximizing

$$\exp\left(-\left(\sum_{u,v=1}^p z_u B_{uv} z_v + \sum_{i=1}^n \frac{(z_i - y_i)^2}{C_{\varepsilon_i}}\right)\right) \quad (1.21)$$

This is also the expression (up to a multiplicative constant) of the conditional multivariate distribution in the multigaussian case, as given by Bayes theorem (Eq. 1.16), in the case where $m=0$, where the matrix C_ε is diagonal and where the data are point values. The first term represents the prior pdf and the second the likelihood function. Kriging which is equal to the mean of the posterior pdf, also maximizes this pdf in the multigaussian case.

Expression (1.21) relates the world of energy functionals (such as splines) with that of probability functions (such as Kriging). More generally regularization and maximum a posteriori Bayesian estimates are identical if the prior covariance used in Bayesian inversion is properly chosen. The equivalence between an energy function and a probability distribution is also used in statistical mechanics, as the probability of a particular configuration is inversely related to its energy. Suppose that the vector z minimizes an energy functional $E(z)$. Using the results of Geman and Geman (1984), Szeliski and Terzopoulos (1989) associate a probability to this energy through the Boltzmann (or Gibbs) distribution $p(z)$ defined as

$$p(z) = \frac{1}{Z} \exp\left(-\frac{E(z)}{T}\right) \quad (1.22)$$

where Z and T are positive constants. If Bayes' theorem is applied to the above prior pdf $p(z)$ and the posterior pdf is maximized, the formalism of splines is obtained.

1.2.7 Conclusion on Kriging

Three different ways of calculating a Kriging interpolator have been discussed

- using the basic approach where Kriging is calculated at each location as a linear combination of the data (Eq. 1.2)
- using Eq. 1.7, where the expression of dual Kriging, or more generally of radial basis functions, is used

- minimizing Eq. 1.14 in the discrete case, where Kriging values are calculated on a discrete grid by a minimization incorporating a regularization and a distance to the data term.

Kriging, although derived using a probabilistic formalism, is still a deterministic technique, in the sense that one unique or “best” model is produced. In most cases, Kriging provides a representation that is very smooth. As a result the application of non-linear operators to Kriged models will provide biased results (Dubrule 2003). This is one of the reasons for the success of conditional simulation.

1.3 Stochastic Aspects of Geostatistics: Conditional Simulation

With conditional simulation, the approach is stochastic. A large number of realizations are generated, which match the data (if the simulation is conditional) and share the first (mean) and second order (stationary covariance or generalized covariance) moments of the modeled random function. The main benefit of conditional simulation is that it produces realizations that behave away from the well data the same way as the well data themselves (Dubrule 2003). This is not true with Kriging, which produces a model that is smoother away from the wells than it is at the wells.

Conditional simulation can also be regarded as a technique for generating realizations of the conditional multigaussian pdf fully characterized by Eqs. 1.17 and 1.18. In other words, the realizations “vibrate” around their Kriging mean with a variance at each location equal to the Kriging variance.

A number of conditional simulation algorithms have been developed (Chilès and Delfiner 2012, p. 478). Among them, two are routinely used in the petroleum industry and are particularly interesting in relation with the inversion of seismic and production data.

1.3.1 Method 1: “Smooth Plus Rough” or “Rough Plus Smooth” Algorithm

$Z(\mathbf{x})$ can be simply written as the sum of Kriging plus the Kriging error

$$Z(\mathbf{x}) = Z_k(\mathbf{x}) + (Z(\mathbf{x}) - Z_k(\mathbf{x})) \quad (1.23)$$

The “smooth plus rough” (Oliver 1996) simulation method writes a conditional simulation $Z_{cs}(\mathbf{x})$ as the sum of Kriging plus a simulation of the Kriging error. A non-conditional simulation $Z_{ncs}(\mathbf{x})$ of $Z(\mathbf{x})$ is generated first, which honors the

mean and the covariance of $Z(\mathbf{x})$, then the conditional simulation $Z_{cs}(\mathbf{x})$ is calculated as

$$Z_{cs}(\mathbf{x}) = Z_k(\mathbf{x}) + (Z_{ncs}(\mathbf{x}) - Z_{ncsk}(\mathbf{x})) \quad (1.24)$$

where $Z_{ncsk}(\mathbf{x})$ designates Kriging of $Z_{ncs}(\mathbf{x})$ using as data the values $Z_{ncs}(\mathbf{x}_i)$ of the non-conditional simulation at the conditioning data locations. Thus to the smooth term $Z_k(\mathbf{x})$ is added the rough term $(Z_{ncs}(\mathbf{x}) - Z_{ncsk}(\mathbf{x}))$. Chilès and Delfiner (2012, p. 495) show that $Z_{cs}(\mathbf{x})$ honors the data and has the same (generalized) covariance as $Z_{ncs}(\mathbf{x})$ (and hence as $Z(\mathbf{x})$).

Equation 1.24 can also be expressed in the form of a “rough plus smooth” equation

$$Z_{cs}(\mathbf{x}) = Z_{ncs}(\mathbf{x}) + (Z_k(\mathbf{x}) - Z_{ncsk}(\mathbf{x})) \quad (1.25)$$

Using Eq. 1.17, Eq. 1.25 can be written in the discrete case, assuming that the data are average values of the gridded values and are affected by a measurement error. At location u of the discrete grid

$$z_{ucs} = z_{uncs} + CF' (FCF' + C_e)^{-1} (y - Fz_{uncs}) \quad (1.26)$$

Equation 1.26 shows that conditional simulation is obtained by adding to a non-conditional simulation a Kriging of the mismatch $(y - Fz_{uncs})$ between the data and the unconditional simulation at the data location. This formalism will appear to be quite general and will facilitate the understanding of the relationship between conditional simulation and Kalman Filtering.

1.3.2 Method 2: Sequential Gaussian Simulation (SGS)

SGS (Deutsch and Journel 1998) is probably the most popular and flexible conditional simulation technique used in applications. SGS works under the multi-gaussian assumption and sequentially draws random locations within the simulated grid. At each new random location, the value is first Kriged from the previously simulated values and the well data. Then, a random value is sampled from the Gaussian pdf with mean equal to the Kriged value and variance equal to the Kriging variance (SGS uses the property that, in the multivariate normal case, univariate conditional distributions are also Gaussian). Then the sampled value is merged with the rest of the dataset, and a new random location is chosen within the simulated grid. The grid points where a data point is present are treated the same way as grid points with no data if the error ε affecting the data is different from zero. If all the data are exact, then the grid nodes with data points are left unchanged. The result is

a Gaussian realization constrained by the data values and satisfying the input statistics (mean and covariance function).

The main difference between “rough plus smooth” and SGS is that SGS works sequentially, grid point by grid point. The sequential nature of SGS is well suited to the geostatistical inversion of seismic data. Indeed, at each grid node, the sequential approach can make sure that the sampled value is compatible with both the previously generated points and the seismic data at the same location, thus combining the advantage of single trace inversion with that of spatial coupling. This will be discussed in Sect. 1.4.

1.3.3 Spectrum and Conditional Simulation

Since the frequency spectrum is the Fourier transform of the covariance (Chilès and Delfiner 2012, p. 66), the spectrum of a conditional simulation is the same as that of the data. Conditional simulation addresses the following statement from Claerbout (2002) about seismic data interpolation: *Of all the assumptions we could make to fill empty bins, one that people usually find easiest to agree with is that the spectrum should be the same in the empty-bin regions as where bins are filled.*

Claerbout (2002) also defines the Prediction Error Filter (PEF) as the linear operator T that transforms the data into a white noise. In other words, $T'T$ is the inverse of the covariance. Based on Eq. 1.11, this also means that T is the spline operator associated with the covariance of the data. Claerbout (2002) shows that unconditional simulations can be generated by applying T^{-1} to a white noise. This is the same technique as that used by Oliver (1988) and Oliver (1995) who applies what he calls the square root of the covariance function to a white noise.

1.4 Geostatistical Inversion of Seismic Data

1.4.1 Deterministic Seismic Inversion

Until the mid-nineties or so, most seismic inversion studies were deterministic, in the sense that they generated a single “best” model, usually at the same resolution as the seismic data. Often, regularization-based or Bayesian methods were used, which led to the generation of one “maximum posterior” or “optimal for a given norm (often L2)” 3-D acoustic impedance model (Tarantola 2005).

If the seismic inversion problem is linearized as with Fatti et al.’s (1994) model, the reflection coefficient $r(\theta)$ at seismic time t for a seismic block of offset θ can be written

$$r(\theta) = a_1(\theta) \frac{\partial \text{Log} I_p(t)}{\partial t} + a_2(\theta) \frac{\partial \text{Log} I_s(t)}{\partial t} \quad (1.27a)$$

$$\text{and } y(\theta) = w(\theta) * r(\theta) + \varepsilon(\theta) \quad (1.27b)$$

where $I_p(t)$ and $I_s(t)$ are the compressive and shear impedances at time t , $a_1(\theta)$ and $a_2(\theta)$ are offset-related parameters, $w(\theta)$ is the seismic wavelet for offset θ and $\varepsilon(\theta)$ is noise. This model is linear in the logarithm of $I_p(t)$ and $I_s(t)$. Thus, as long as the logarithms of impedances are inverted, the seismic amplitudes can be written as in Eq. 1.15 as a linear function of the logarithms of impedances, and the posterior mean obtained by multigaussian Bayesian seismic inversion (Eqs. 1.17 and 1.18) is identical to Kriging. The solution can also be regarded as a regularization-based solution, where the norm controlling the smoothness is derived from the inverse of the covariance.

At the time when only deterministic inversion was used, geostatisticians often treated seismic data as “soft” information, making use only of statistical correlations between seismic and reservoir parameters in order to constrain the earth models. This “soft” approach to seismic data allowed the development of some interesting interpolation techniques such as external drift or collocated coKriging (Dubrule 2003). However it also led to reservoir models not fully compatible with the seismic data as, if a seismic forward model such as that of Eq. 1.27 was applied to them, the actual seismic data was not recovered.

The above approaches proved sufficient until the late eighties or so, as seismic data were used at rather large scale. Thanks to the development of 3-D earth modeling at the reservoir scale in the early nineties, it became necessary to work with models at higher resolution than seismic data, and hence to quantify the uncertainty attached to these models. Then the availability of 4D seismic data also called for new technology to better constrain the earth models. Geostatistical inversion, described below, was developed with these issues in mind.

1.4.2 Geostatistical Inversion (GI)

The original GI algorithm (Bortoli et al. 1992; Haas and Dubrule 1994) used SGS to simulate high-resolution acoustic impedance traces constrained by seismic data. SGS starts by picking a random cell within a regular two-dimensional grid. At this cell, a large number of possible acoustic impedance vertical traces are generated by SGS, then the trace that best matches the actual seismic trace at this location is selected. Then SGS moves to another random location of the two-dimensional grid, etc. until the whole model is filled with high-resolution impedance traces. Initially SGS appeared to be well suited to this application, as it allowed the use of any kind of forward model—linear or not—relating the acoustic impedance trace generated by SGS to the seismic amplitude trace at the same location. The acoustic

impedance vertical traces simulated by SGS typically have higher frequency content than the seismic amplitudes, which makes them non-unique. This uncertainty can be quantified by generating multiple conditional simulations. Unfortunately the use of SGS proved to take too much computer time for large seismic datasets.

By revisiting the above GI algorithm in a Bayesian framework and in the linear context of Fatti's model (Eq. 1.27), authors such as Buland and Omre (2003) or Escobar et al. (2006) not only clarified the GI formalism but also provided a straightforward conditional simulation algorithm based on Eqs. 1.17 and 1.18 which was more efficient than SGS for sampling acoustic impedance traces compatible with seismic amplitudes. Whilst Bayesian inversion provided an expression of the posterior mean and covariance of the impedances multiGaussian pdf, GI allowed the sampling of reservoir-scale impedance realizations from this pdf.

As convincingly shown by Francis (2006a, b) or Escobar et al. (2006), cut-off operations such as those used to translate acoustic impedance into facies can be applied to GI realizations, thus avoiding statistical bias if these cut-offs were applied to Kriging.

1.5 Kalman Filtering and Ensemble Kalman Filtering

1.5.1 Kalman Filtering (KF)

Suppose that a Gaussian random vector Z_{t-1} has evolved until time $(t-1)$ and that Z_{t-1} is an unbiased estimate of the unknown true state vector z_{t-1} at time $(t-1)$

$$Z_{t-1} = z_{t-1} + R_{t-1} \text{ with } E(R_{t-1}) = 0 \text{ and } \text{Var}(Z_{t-1}) = \text{Var}(R_{t-1}) = C_{t-1} \quad (1.28)$$

If the model error is neglected, the forward model relating the true state vector at time $(t-1)$ with the state vector at time t is assumed to be a linear function L_t

$$z_t = L_t z_{t-1} \quad (1.29)$$

At time step t , the unknown true state of the system has evolved according to Eq. 1.29 and a vector d_t of n new data may also be available. Assume that these data are linear functions of the state vector z_t , and can be expressed as in Eq. 1.15

$$d_t = F_t z_t + \varepsilon_t \quad (1.30)$$

where the error vector ε_t has mean zero and covariance matrix C_{ε_t} .

KF (Kalman 1960) aims to combine the information provided about z_t by the forward model L_t applied to the estimate Z_{t-1} (Eq. 1.29) with the information provided by the data d_t (Eq. 1.30). Bayes can be used for this, $L_t Z_{t-1}$ playing the role of the prior distribution. It is easy to verify that the covariance of the random

vector $L_t Z_{t-1}$ is $C_t = L_t C_{t-1} L_t'$. Hence, under Gaussian assumptions the best estimate is (from Eq. 1.19)

$$Z_t = L_t Z_{t-1} + \Lambda_t (d_t - F_t L_t Z_{t-1}) \quad (1.31)$$

where the kriging weights matrix Λ_t (as in Eq. 1.20) is now called the Kalman gain

$$\Lambda_t = C_t F_t' (F_t C_t F_t' + C_{\varepsilon_t})^{-1} \quad (1.32)$$

Z_{t-1} as defined in Eq. 1.28 can represent any kind of unbiased estimate based on all the information available at time $(t-1)$. Kriging and conditional simulation are both unbiased estimates of z_t , only their variance is different and is of course minimum if Z_{t-1} is Kriging and larger if Z_{t-1} is simulation. Chilès and Delfiner (2012) (p. 497) show that the variance of the difference between a random function and its conditional simulation is twice the Kriging variance. In case Z_{t-1} is simulation, Eq. 1.31 looks like the “rough plus smooth” method (Eq. 1.26) with $L_t Z_{t-1}$ playing the role of the non conditional simulation. Equation 1.31 makes the estimate $L_t Z_{t-1}$ conditional to the new data d_t by adding an interpolation of the mismatch between $F_t L_t Z_{t-1}$ and the data.

In standard geostatistical applications, the observations are often spatial and hence assimilated simultaneously, while KF processes information sequentially, time step after time step. Tarantola (2005) (in Appendix 6.18) shows that, if in a linear least-squares problem the dataset can be divided into subsets with zero covariance between them, then solving one global inverse problem is equivalent to solving a series of smaller problems using the posterior state and covariance matrix of each partial problem as prior information for the next. Oliver et al. (2008) also show (in Chap. 11) that, under the same assumptions as Tarantola (2005), the step by step computation of KF provides (in the multigaussian case) the same result as would be obtained by integrating all the data in one single step. In the case where L_t is the identity function, these two results also imply that simple Kriging would provide the same result if data were incorporated sequentially into the Kriging system, or in one single batch (under the assumptions that each batch of data has zero covariance with the others).

1.5.2 Constraining Reservoir Models by Production Data

Fluid flow models are strongly non-linear, and linear approximations such as those already discussed for seismic modeling or KF cannot be used.

A distinction must be made between “history-matching”, where a single reservoir model is modified until the flow simulation matches the production data, and “constraining reservoir models by production data”, where reservoir model

realizations compatible with production data are generated. Here the discussion focuses on the second objective rather than the first one. Some techniques to address this objective are based on rigorous approaches such as Markov-Chain Monte Carlo (MCMC) or Genetic Algorithms (GA) (Oliver et al. 2008). But these are very time-consuming and often unpractical. Ensemble Kalman filtering appears to be a more practical approach for incorporating production data into the reservoir model.

1.5.3 Ensemble Kalman Filtering (EnKF) Versus Conditional Simulation

EnKF (Evensen 2007; Oliver et al. 2008) starts with an ensemble of initial realizations that are not constrained by production data. Typically the state vector z_t at time t contains permeabilities, porosities, saturations, pressures, and thermodynamic variables at the simulator grid nodes followed by a vector of predicted production data at each well i at time t .

The following notation is used for a given state vector z_t

$$z_t = (z_{ut}^1, z_{ut}^2, \dots, z_{ut}^k, q_{it}^{1*}, \dots, q_{it}^{l*}) \quad (1.33)$$

It is assumed that there are k gridded variables in z_t , that the simulator grid is composed of p cells u , and that there are n wells i each with l new production data at time t . The total size of the state vector z_t is $kp + nl$. The predicted data vector d_t^* is the vector of size nl

$$d_t^* = (q_{1t}^{1*}, \dots, q_{nt}^{1*}, q_{1t}^{2*}, \dots, q_{nt}^{2*}, \dots, q_{1t}^{l*}, \dots, q_{nt}^{l*}) \quad (1.34)$$

The relation between state vector and predicted data is

$$d_t^* = Pz_t \text{ with } P = (O_{nl \times kp}, I_{nl \times nl}) \quad (1.35)$$

P is a $nl \times (kp + nl)$ matrix. The function f_t , which represents the flow simulator, is non-linear. If the model errors are neglected

$$z_t = f_t(z_{t-1}) \quad (1.36)$$

does not modify the rock properties (unless they are affected by changes in pressure and saturation), but replaces the pressure, saturation, and simulated data with new values at time t .

The problem is now to calculate the best estimate of the state vector z_t combining the information provided by the flow simulation forward model $f_t(z_{t-1})$ and that provided by the new data d_t .

If f_t is a linear function, this is the standard KF domain of application and Eq. 1.31 applies, L_t playing the role of f_t . But now f_t is non-linear. It would still be convenient to update the state vector through a generalization of Eq. 1.31

$$z_t = f_t(z_{t-1}) + \Lambda_t(d_t - Pf_t(z_{t-1})) \quad (1.37)$$

where the Kalman gain Λ_t is obtained using Eq. 1.32. Assuming that there is no error associated with the data, Eq. 1.32 can be simplified into

$$\Lambda_t = C_t P' (P C_t P')^{-1} \quad (1.38)$$

Equation 1.38 requires the knowledge of the covariance C_t of $f_t(z_{t-1})$, in other words the covariance of the image of the state vector after application of the flow simulation model f_t . f_t is non-linear and this covariance cannot be simply calculated—as in the linear case—from the covariance at the previous step. EnKF addresses this issue by statistically deriving this covariance using the information from the multiple realizations, typically about a hundred of them. This is the key idea behind EnKF.

There are of course a number of issues resulting from the fact that the covariances are calculated from a finite number of realizations of the ensemble. The first one is spurious correlation, because the ensemble members are not independent except in the starting ensemble. The second one is that if the number of realizations in the ensemble is not large enough, then the covariances are poorly estimated. Standard geostatistics addresses this by fitting mathematical models to the experimental covariances, in order to smooth the spurious correlations.

1.5.4 Ensemble Kalman Filtering and Its Relationship with CoKriging

In Eq. 1.37, focus now on the rock properties in the state vector. $f_t(z_{t-1})$ leaves the rock properties unchanged, as only the time-dependent state vectors in the simulator grid are calculated by one time-step of the flow simulator, whilst $\Lambda_t(d_t - Pf_t(z_{t-1}))$ is a linear combination of the differences between observed and predicted production data at each well. Thus EnKF interpolates between the wells by calculating a linear combination of these differences across the field, then adds these interpolated difference to the rock properties model. Is it possible to reformulate EnKF as a well by well geostatistical approach?

The term Λ_t in Eq. 1.37 is the Kalman gain as given by Eq. 1.38. In the case where there is no error affecting the data, Eqs. 1.37 and 1.38 can be written

$$z_t - f_t(z_{t-1}) = C_t P' (P C_t P')^{-1} (d_t - P f_t(z_{t-1})) \quad (1.39)$$

The left-hand side is the update calculated by EnKF for the property of interest as the time step evolves from $(t-1)$ to t . The Kalman gain coefficients of the right-hand side are nothing else than the simple coKriging weights (see for instance Chilès and Delfiner 2012, p. 303).

Thus, each estimate of a 3-D spatial parameter such as porosity or permeability at time $(t-1)$ is updated at time t by a linear combination of all the inconsistencies generated by this parameter at the data points. Since, in the case of flow simulations, many parameters are involved in the production profiles prediction, all the individual parameters' 3-D models must be corrected in a consistent way, which is why multivariate coKriging—and not univariate Kriging applies here.

1.6 Beyond the Formal Relationship Between Geostatistics and Bayes

1.6.1 Two Identical Formalisms but Different Assumptions

The above developments show that techniques such as conditional simulation, Bayesian inversion, geostatistical inversion and ensemble Kalman filtering follow a similar mathematical formalism.

However, their philosophy of application differs in the way the covariance is approached. This can be understood by looking again at Bayes rule as presented in Eq. 1.16

$$f_{post}(z) \propto f_{prio}(z) g(y/z) \quad (1.40)$$

With geostatistics, the experimental (generalized) covariance calculated on the data y is fitted by a model which becomes the covariance of the unconditional distribution $f_{prio}(z)$. Then the data y are used a second time through the simulation conditioning process of Eq. 1.26.

With Bayes, the covariance model associated with $f_{prio}(z)$ is a prior based on local or analog knowledge, but not on the data themselves (Tarantola 2005). This prior is transformed into a posterior covariance through the conditioning process of Eq. 1.40.

With geostatistics, the aim of conditional simulation is to generate realizations that match the data and satisfy the input covariance; the SGS and rough plus smooth algorithms work only if the data themselves satisfy this input covariance. But the random function $Z_{cs}(\mathbf{x})$ of Eqs. 1.24 and 1.25 is not an ergodic or even a stationary random function; its variance at each location \mathbf{x} is equal to the Kriging variance and changes with \mathbf{x} , as it is zero at the data points. In other words, the covariance of the random function $Z(\mathbf{x})$ is different from that of $Z_{cs}(\mathbf{x})$ conditionally to the data

(Chilès and Delfiner 2012, p. 497). But the covariance calculated on a single conditional realization does not “see” any difference between the grid cells associated with data points and those not associated with data points. It is only as the realizations change, leaving the data unchanged, that the covariance across realizations appears non-stationary and hence non-ergodic.

On the other hand, Bayes combines a prior covariance—usually different from that of the data—with a data-based likelihood, resulting into a posterior pdf that sits somewhere between the prior and the likelihood. Bayes *updates* prior covariances based on new data whilst conditional simulation *anchors* the realizations against the hard data (Escobar, personal communication).

1.6.2 Model Falsifiability

Tarantola (2006) challenges the geostatistical and Bayes formalisms if models are to be falsifiable or have a scientific meaning: *I suggest that the setting, in principle, for an inverse problem should be as follows: use all available prior information to sequentially create models of the system, potentially an infinite number of them. For each model, solve the forward modeling problem, compare the predictions to the actual observations and use some criterion to decide if the fit is acceptable or unacceptable, given the uncertainties in the observations and, perhaps, in the physical theory being used. The unacceptable models have been falsified, and must be dropped. The collection of all the models that have not been falsified represent the solution of the inverse problem.* Thus, Tarantola (2006) offers to keep all the prior realizations that are compatible with the data. Thus the data are used to validate or reject the prior realizations, rather than update the prior pdf into the posterior.

1.6.3 Looking Ahead: Machine Learning and Falsifiability

The fast growth in machine learning algorithms (Goodfellow et al. 2016) is challenging the geostatistical and Bayesian formalisms in situations where data are plenty. Thanks to this large number of data, the approach used to falsify a convolutional neural network model (for instance) relating input parameters to data is often to test whether the convolutional model works as well on a training (or calibration) dataset as on a test dataset not used for training. The prior model itself is completely data-driven, which contradicts Tarantola (2006) but the validation step is along the lines of his above recommendations! This topic is likely to generate interesting discussions in the future.

1.7 Conclusion

The objective of this chapter was to discuss the convergence observed over the last fifty years between geostatistics and other modelling and inversion techniques.

A formal convergence exists between the main techniques used to constrain reservoir models by multi-disciplinary data. Kriging, splines, conditional simulation, geostatistical inversion and ensemble Kalman filtering can be interpreted using either the geostatistical formalism or Bayes.

Most of these techniques amount to the same approach where an initial model is updated by using a linear combination of the mismatches between the new data and their prediction from the initial model (Eqs. 1.19, 1.26, 1.31 and 1.39).

However the methods above have a different philosophy towards the inference of the covariances used in these calculations. Bayes uses the data to update a prior pdf which is independent of the data. Geostatistics generate realizations of conditional simulations that reproduce the modeled covariance—or the spectrum—of the data. EnKF does not model a covariance but directly uses the empirical covariances derived from the ensemble realizations and their flow simulations.

Acknowledgements The author would like to thank Imperial College London, and Total for seconding him there as a Visiting Professor. Igor Escobar and Danila Kuznetsov are thanked for the fruitful discussions held at the Total Geoscience Research Centre in Aberdeen (UK). The author will also never forget the passionate and illuminating discussions with Albert Tarantola at the Café Beaubourg in Paris!

References

- Billings SD, Beaton RK, Newsam GN (2002a) Interpolation of geophysical data using continuous global surfaces. *Geophysics* 67(6):1810–1822
- Billings SD, Newsam GN, Beaton RK (2002b) Smooth fitting of geophysical data using continuous global surfaces. *Geophysics* 67(6):1823–1834
- Bolondi G, Rocca F, Zanoletti S (1976) Automatic contouring of faulted subsurfaces. *Geophysics* 41(6):1377–1393
- Bortoli LJ, Alabert F, Haas, A, Journel AG (1992) Constraining stochastic images to seismic data. In: Soares A (ed) *Proceedings of the international geostatistics congress, Troia 1992*. Kluwer Publishers, Dordrecht, Netherlands
- Briggs IC (1974) Machine contouring using minimum curvature. *Geophysics* 39(1):39–48
- Buland A, Omre H (2003) Bayesian linearized AVO inversion. *Geophysics* 68(1):185–198
- Chilès JP, Delfiner P (2012) *Geostatistics. Modeling spatial uncertainty*. Wiley series in probability and statistics, 2nd edn. Wiley
- Claerbout J (2002) Image estimation by example: geophysical soundings image construction: multidimensional autoregression. <http://sepwww.stanford.edu/sep/prof>
- Deutsch CV, Journel AG (1998) *GSLIB, geostatistical software library and user's guide*, 2nd edn. Oxford University Press
- Dubrule O (1983) Two methods with different objectives: splines and Kriging. *Math Geol* 15(2): 245–257
- Dubrule O (2003) *Geostatistics for seismic data integration in 3-D earth models, SEG/EAGE Course Notes*. SEG Publications

- Duchon J (1975) Fonctions Splines du Type Plaque Mince en Dimension 2, Séminaire d'Analyse Numérique, No. 231, U.S.M.G, Grenoble, France
- Escobar I, Williamson P, Cherett A, Doyen PM, Bornard R, Moyen R, Crozat T (2006) Fast geostatistical inversion in a stratigraphic grid, SEG Annual Meeting, New Orleans
- Evensen G (2007) Data assimilation. The ensemble Kalman filter. Springer
- Fatti JL, Smith GC, Vail PJ, Strauss PJ, Levitt PR (1994) Detection of gas in sandstone reservoirs using AVO analysis: a 3-D seismic case history using the geostack technique. *Geophysics* 59:1362–1376
- Francis A (2006a) Understanding stochastic inversion part 1. *First Break* 24:69–77
- Francis A (2006b) Understanding stochastic inversion part 2. *First Break* 24:79–84
- Geman D, Geman S (1984) Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell PAMI-6*(6):721–741
- Goodfellow I, Bengio Y, Courville A (2016) *Deep learning*. MIT Press
- Haas A, Dubrule O (1994) Geostatistical inversion—a sequential method of stochastic reservoir modeling constrained by seismic data. *First Break* 12(11):561–569
- Hansen TM, Journel AG, Tarantola A, Mosegaard K (2006) Linear inverse Gaussian theory and geostatistics. *Geophysics* 71(6):101–111
- Kalman RE (1960) A new approach to linear filter and prediction problems. *J Basic Eng* 82:35–45
- Matheron G (1971) The theory of regionalized variables and its applications. http://cg.ensmp.fr/bibliotheque/public/MATHERON_Ouvrage_00167.pdf
- Matheron G (1973) The intrinsic random functions and their applications. *Adv Appl Probab* 5:439–468
- Matheron G (1981a) Splines and Kriging, their formal equivalence. Syracuse University, Geol Contrib 8:77–95
- Matheron G (1981b) Splines et Krigeage: Le Cas Fini. http://cg.ensmp.fr/bibliotheque/public/MATHERON_Rapport_00228.pdf
- Micula G, Micula S (1999) *Handbook of splines*. Springer Science & Business Media
- Oliver DS (1988) Calculation of the inverse of the covariance. *Math Geol* 30(7):911–933
- Oliver DS (1995) Moving averages for Gaussian simulation in two and three dimensions. *Math Geol* 27(8):939–960
- Oliver DS (1996) On conditional simulation to inaccurate data. *Math Geol* 28(6):811–817
- Oliver D, Reynolds AC, Liu N (2008) *Inverse theory for petroleum reservoir characterization and history-matching*. Cambridge University Press
- Szeliski R, Terzopoulos D (1989) From splines to fractals. *Comput Graph* 23(3):51–60
- Tarantola A (2005) *inverse problem theory and model parameter estimation*. SIAM, Philadelphia
- Tarantola A (2006) Popper, Bayes and the inverse problem. *Nat Phys* 2:492–494

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

