III. THE BASIC EQUATIONS OF THE BME MAPPING METHOD

The methods of classical geostatistics were primarily developed under a set of constraining assumptions (linear estimator, measurements are exacts, etc.) and they lack the theoretical underpinnings and practical flexibility to incorporate the many sources of information available in modern days sciences (such as physical laws, empirical models, higher statistical moments, uncertain information). The physical knowledge base available in modern geostatistics was defined in the previous chapter as the union of general knowledge describing the Space/Time Random Field (S/TRF) $X(\boldsymbol{p})$, and specificatory knowledge which include exact observed values (hard data) and uncertain measurements (soft data). In this chapter I present the epistemological tools of the Bayesian Maximum Entropy (BME; Christakos, 1990, 1992) method of modern geostatistics which are used to process the total knowledge base available. The epistemological approach of BME may be represented by the use of two operators; the first for processing general knowledge, and the second to process specificatory knowledge. The result of the knowledge processing operators is a BME posterior pdf, which describes completely the S/TRF at the estimation point. The BME posterior pdf provides a complete picture of the mapping situation, and it provides as well different estimators of the S/TRF and their associated estimation uncertainty, which are useful for mapping purposes.

3.1. A Rapid Tour of the BME Approach

As we saw in the previous chapter, the spatiotemporal mapping problem is, generally, concerned with the estimation of the natural process $X(\boldsymbol{p})$ at point $\boldsymbol{p}_k$, using the total physical knowledge $K = G \cup S$ available. Due to its sound epistemological background, mathematical rigor and considerable flexibility in incorporating various sources of physical knowledge, BME is a powerful technique of spatiotemporal analysis and mapping. BME distinguishes three essential stages of analysis, as follow.

(i) The *prior* stage starts with a basic set of assumptions and general knowledge G. The epistemological goal is informativeness (prior information maximization given G).

(ii) The *meta-prior* or *pre-posterior* stage considers the specificatory knowledge S including hard and soft physical data.

(iii) The *integration* or *posterior* stage, in which the knowledge bases considered in stages (i) and (ii) are integrated. The epistemological goal here is cogency (posterior probability maximization given both G and S).

The mathematical analysis involved in these three stages leads to the posterior pdf as follows (Christakos, 1990, 1992,1998)

$$f_K(\chi_k) = A^{-1} Y_S[\boldsymbol{\chi}_{\text{soft}}, Y_G(\boldsymbol{\chi}_{\text{map}})], \tag{3.1}$$

where Y_G represents the operator processing general knowledge G, Y_S represents the operator processing specificatory knowledge S, and A is a normalization constant. The operators Y_G and Y_S takes on different form depending on the type of general and specificatory knowledge considered.

When the general knowledge G consists of constraints expressed in the form of Eqs. (2.23), and using the epistemological goal of informativeness (also called entropy maximization), the Y_G operator can be expressed as

$$Y_G = \sum_{\alpha=1}^{N_\alpha} \mu_\alpha(\boldsymbol{p}_{\text{map}}) g_\alpha(\boldsymbol{\chi}_{\text{map}}), \tag{3.2}$$

where μ_α ($\alpha = 1,..,N_\alpha$) are Lagrange multipliers calculated by substituting Eqs. (2.23) into Eqs. (3.2), so that $f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1}\exp[Y_G(\boldsymbol{\chi}_{\text{map}})]$ is the prior pdf associated with the general knowledge G, and $Z = \exp[-\mu_0]$ is a normalization constant.

Similarly, a variety of Y_S operators can be developed, which express knowledge related to uncertain evidence, probability distributions, functional relationships, etc. In the case of interval soft data described by Eq. (2.26), we will show in chapter 4. that the Y_S may be written as (Christakos, 1990, 1992)

$$Y_S = Z^{-1}\int_{\boldsymbol{l}}^{\boldsymbol{u}} d\boldsymbol{\chi}_{\text{soft}} \exp[Y_G(\boldsymbol{\chi}_{\text{map}})], \tag{3.3}$$

where the multi-dimensional integration is over a hyper-cube bounded by the lower bound vector $\boldsymbol{l} = [l_{m_h+1}\ldots l_m]^T$ and the upper bound vector $\boldsymbol{u} = [u_{m_h+1}\ldots u_m]^T$. In the case of soft data of the probabilistic type as described in Eq. (2.27), we will show in chapter 5 that the operator is given by (Christakos, 1990, 1992, 1998),

$$Y_S = Z^{-1}\int d\boldsymbol{\chi}_{\text{soft}}\, f_S(\boldsymbol{\chi}_{\text{soft}})\exp[Y_G(\boldsymbol{\chi}_{\text{map}})]. \tag{3.4}$$

A more detailed discussion of the epistemological principles underlying the derivation of the BME formulas is now in order.

3.2. Different Forms of the General Knowledge Operator

When the general knowledge G is expressed in terms of the general constraints of Eqs. (2.23), the Y_G operator takes the form of Eq. (3.2). This powerful result is obtained using the epistemological goal of informativeness by maximizing the expected information given the general constraints, i.e. by maximizing the Entropy of the multivariate (pdf) $f_G(\boldsymbol{\chi}_{\text{map}})$, $\boldsymbol{\chi}_{\text{map}} = [\chi_1,\ldots,\chi_m,\chi_k]^T$, of the random variables $\boldsymbol{x}_{\text{map}} = [x_1,\ldots,x_m,x_k]^T$. The pdf

$f_G(\boldsymbol{\chi}_{\text{map}})$ is named the "prior" pdf relative to the specificatory knowledge, which has not been used yet. In the following we present the general solution to the maximum entropy problem, and we then derive some specific form of the prior pdf for some specific types of general knowledge. The epistemological goal of informativeness is powerful because it may be applied to a large class of general knowledge, that goes beyond the specific forms presented here.

3.2.1. A General Solution of the Entropy Maximization Problem

Let $f_G(\boldsymbol{\chi}_{\text{map}})$, $\boldsymbol{\chi}_{\text{map}} = [\chi_1, \ldots, \chi_m, \chi_k]^T$, be the multivariate probability density function (pdf) of the random variables $\boldsymbol{x}_{\text{map}} = [x_1, \ldots, x_m, x_k]^T$ before any data (hard or soft) have been taken into consideration.

We obtain the prior probability density function by maximizing the expected information -also called the *entropy function*- given by

$$S[f_x] = \overline{Inf(\boldsymbol{x}_{\text{map}})} = -\int d\boldsymbol{\chi}_{\text{map}} f_G(\boldsymbol{\chi}_{\text{map}}) \ell og\, f_G(\boldsymbol{\chi}_{\text{map}}), \tag{3.5}$$

for the set of prior constraints (Eqs 2.23), expressed here again for convenience as

$$\overline{g_\alpha} = \int d\boldsymbol{\chi}_{\text{map}}\, f_G(\boldsymbol{\chi}_{\text{map}}) g_\alpha(\boldsymbol{\chi}_{\text{map}}), \tag{3.6}$$

where $\alpha = 0,1,\ldots,N_c$ and the $g_\alpha(\boldsymbol{\chi}_{\text{map}})$ are given functions of χ_i's. We introduce the modified entropy functional $S_c[f_x]$ involving the Lagrange multipliers μ_α, i.e.,

$$S_c[f_G] = -\int d\boldsymbol{\chi}_{\text{map}}\, f_G(\boldsymbol{\chi}_{\text{map}}) \log f_G(\boldsymbol{\chi}_{\text{map}}) + \sum_{\alpha=0}^{N_c} \mu_\alpha \overline{g_\alpha}[f_G]. \tag{3.7}$$

The prior pdf is obtained by maximizing the above functional, which leads to the following solution (see Appendix B.)

$$f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1} \exp[\sum_{\alpha=1}^{N_c} \mu_\alpha \, g_\alpha(\boldsymbol{\chi}_{\text{map}})], \tag{3.8}$$

where $$Z = \int d\boldsymbol{\chi}_{\text{map}} \exp[\sum_{\alpha=1}^{N_c} \mu_\alpha \, g_\alpha(\boldsymbol{\chi}_{\text{map}})]. \tag{3.9}$$

The prior pdf of Eq. (3.8) is quite general since it accounts for any type of general knowledge as long as that knowledge may be expressed in terms of the general constraints of Eqs (3.6). One case of interest in spatiotemporal mapping --which will be the case for the numerical BME implementations of subsequent chapters-- is when the general knowledge consist of the mean and covariance function. This case is presented next.

3.2.2. The Gaussian Form: When Only the Mean and Covariance Functions are Known

We now assume that the general knowledge includes only the mean m_i at each data points $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, m, k)$, and the covariance's $c_{ij} = \overline{[X(\boldsymbol{p}_i) - m_i][X(\boldsymbol{p}_j) - m_j]}$, for $(i\ ,\ j = 1, 2, \ldots, m, k)$. We choose the constraints so they incorporate that information to the prior pdf. The information about the mean is incorporated by writing the following $g_\alpha(\boldsymbol{\chi}_{\text{map}})$ functions

$$g_i(\boldsymbol{\chi}_{\text{map}}) = \chi_i \quad \text{for } i = 1, 2, \ldots, m, k \tag{3.10}$$

and requiring that $\overline{g_i(\boldsymbol{\chi}_{\text{map}})} = m_i$. Similarly the covariance information is incorporated by writing

$$g_{ij}(\boldsymbol{\chi}_{\text{map}}) = (\chi_i - m_i)(\chi_j - m_j) \quad \text{for} \quad i \;, \; j = 1,2,\dots,m,k \tag{3.11}$$

and requiring that $\overline{g_{ij}(\boldsymbol{\chi}_{\text{map}})} = c_{ij}$.

Substituting these expressions for the functions $g_\alpha(\boldsymbol{\chi}_{\text{map}})$ we get the following prior pdf

$$f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1} \exp \Im[\boldsymbol{\chi}_{\text{map}}], \tag{3.12}$$

where the $\Im$ function is given by (3.13)

$$\Im[\boldsymbol{\chi}_{\text{map}}] = \sum_{i=1}^{m,k} \mu_i \, \chi_i + \sum_{i,j=1}^{m.k} \mu_{ij} \, (\chi_i - m_i)(\chi_j - m_j), \tag{3.14}$$

and the constraints are written as

$$m_i = \int d\boldsymbol{\chi}_{\text{map}} \, \chi_i \; f_G(\boldsymbol{\chi}_{\text{map}}) \quad \text{for} \; i = 1,2,\dots,m,k \tag{3.15}$$

and

$$c_{ij} = \int d\boldsymbol{\chi}_{\text{map}} \, (\chi_i - m_i)(\chi_j - m_j) \, f_G(\boldsymbol{\chi}_{\text{map}}) \quad \text{for} \quad i \;, \; j = 1,2,\dots,m,k \tag{3.16}$$

In the above constraint equations the means m_i and the covariance's c_{ij} are known, and we are solving for the Lagrange coefficients μ_i $(i = 1,2,\dots,m,k)$ and $\mu_{i,j}$ $(i \;, \; j = 1,2,\dots,m,k)$

Solving first for μ_i $(i = 1,2,\dots,m,k)$, we note that the constraints for these unknowns can be rewritten as

$$m_i = \int d\boldsymbol{\chi}_{\text{map}} \, \chi_i \; f_G(\boldsymbol{\chi}_{\text{map}}) = Z^{-1} \frac{\partial}{\partial \mu_i} \int d\boldsymbol{\chi}_{\text{map}} \exp \Im[\boldsymbol{\chi}_{\text{map}}] = Z^{-1} \frac{\partial Z}{\partial \mu_i} = \frac{\partial \ln Z}{\partial \mu_i} \tag{3.17}$$

In the following calculations it will be convenient to define the coefficients

$$\lambda_{ij} = -2\mu_{ij}, \tag{3.18}$$

which let us rewrite the prior pdf as

$$\begin{aligned} f_G(\boldsymbol{\chi}_{\text{map}}) &= Z^{-1}\exp[\sum_{i=1}^{m,k}\mu_i\,\chi_i - \tfrac{1}{2}\sum_{i,j=1}^{m,k}\lambda_{ij}\,(\chi_i - m_i)(\chi_j - m_j)] \\ &= Z^{-1}\exp[\boldsymbol{\mu}^T\boldsymbol{\chi}_{\text{map}} - \tfrac{1}{2}(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})^T\,\boldsymbol{\lambda}\,(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})], \end{aligned} \tag{3.19}$$

where $\boldsymbol{\mu} = [\mu_1 \dots \mu_m, \mu_k]^T$ and

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1\text{m}} & \lambda_{1\text{k}} \\ \vdots & & & \\ \lambda_{\text{m}1} & \dots & \lambda_{\text{mm}} & \lambda_{\text{mk}} \\ \lambda_{\text{k}1} & \dots & \lambda_{\text{km}} & \lambda_{\text{kk}} \end{bmatrix}. \tag{3.20}$$

By letting $\boldsymbol{\psi}_{\text{map}} = \boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}}$, Z can be written as

$$\begin{aligned} Z &= \exp[\boldsymbol{\mu}^T\mathbf{m}_{\text{map}}]\int d\boldsymbol{\chi}_{\text{map}}\,\exp[\boldsymbol{\mu}^T\boldsymbol{\psi}_{\text{map}} - \tfrac{1}{2}\boldsymbol{\psi}_{\text{map}}^T\boldsymbol{\lambda}\,\boldsymbol{\psi}_{\text{map}}] \\ &= \exp[\boldsymbol{\mu}^T\mathbf{m}_{\text{map}}]\int d\boldsymbol{\chi}_{\text{map}}\,\exp[-\tfrac{1}{2}(\boldsymbol{\psi}_{\text{map}}^T - \boldsymbol{\lambda}^{-1}\boldsymbol{\mu})^T\,\boldsymbol{\lambda}\,(\boldsymbol{\psi}_{\text{map}} - \boldsymbol{\lambda}^{-1}\boldsymbol{\mu}) + \tfrac{1}{2}\boldsymbol{\mu}^T\boldsymbol{\lambda}^{-1}\boldsymbol{\mu}] \\ &= \exp[-\tfrac{1}{2}\boldsymbol{\mu}^T(\boldsymbol{\lambda}^{-1}\boldsymbol{\mu} - \mathbf{2m}_{\text{map}})](2\pi)^{\frac{m+1}{2}}|\boldsymbol{\lambda}|^{-\frac{1}{2}} \\ &= (2\pi)^{\frac{m+1}{2}}|\boldsymbol{\lambda}|^{-\frac{1}{2}}\exp[-\tfrac{1}{2}\boldsymbol{\mu}^T(\boldsymbol{\lambda}^{-1}\boldsymbol{\mu} - \mathbf{2m}_{\text{map}})]. \end{aligned} \tag{3.21}$$

Using this result we write the constraint for μ_i as

$m_i = Z^{-1}\dfrac{\partial Z}{\partial \mu_i} = \sum_{j=1}^{m,k} \lambda_{ij}^{-1}\mu_i + m_i$, where λ_{ij}^{-1} is the ij-th element of the inverse matrix $\boldsymbol{\lambda}^{-1}$,

or

$\mu_i = 0$ for $i = 1,2,\ldots,m,k$.

Now that we have calculated $\mu_i = 0$ for $i = 1,2,\ldots,m,k$, what we have left to calculate are the remaining Lagrange coefficient $\mu_{i,j}$, for i , $j = 1,2,\ldots,m,k$. We recall that the corresponding constraints are, for i , $j = 1,2,\ldots,m,k$,

$$c_{ij} = \int d\boldsymbol{\chi}_{\text{map}}\,(\chi_i - m_i)(\chi_j - m_j)\, f_G(\boldsymbol{\chi}_{\text{map}})$$

$$= (2\pi)^{-\frac{m+1}{2}} |\boldsymbol{\lambda}|^{+\frac{1}{2}} \int d\boldsymbol{\chi}_{\text{map}}\,(\chi_i - m_i)(\chi_j - m_j)\exp[-\tfrac{1}{2}(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})^T \boldsymbol{\lambda}\,(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})]$$

From well known results for multivariate normal distribution (see appendix C and D) we get

$$c_{ij} = \lambda_{ij}^{-1}$$

where again λ_{ij}^{-1} is the ij-th element of the inverse matrix $\boldsymbol{\lambda}^{-1}$, so that $\mathbf{C} = \boldsymbol{\lambda}^{-1}$, i.e. the prior pdf is multivariate gaussian

$$f_G(\boldsymbol{\chi}_{\text{map}}) = \frac{1}{(2\pi)^{\frac{m+1}{2}} |\boldsymbol{C}|^{\frac{1}{2}}} \exp[-\tfrac{1}{2}(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})^T \boldsymbol{C}^{-1}\,(\boldsymbol{\chi}_{\text{map}} - \mathbf{m}_{\text{map}})] \qquad (3.22)$$

3.2.3. Going beyond the Gaussian form: Incorporating physical laws

A mean and covariance are valuable general information characterizing a natural variable, and their incorporation leads to a multivariate Gaussian distribution of the prior pdf. An additional source of information that is part of the general knowledge are any physical laws governing the natural variable. Consider for example the three dimensional Darcy equation for groundwater flow in stochastic hydrology. This physical law is very informative when mapping groundwater flows or water table elevations in aquifers. Methods of classical geostatistics lack the theoretical underpinnings to incorporate that type of general knowledge. As a result the approach usually taken is to calculate the deterministic numerical solution of the Darcy equation for several realizations of the random porous media, and taking ensemble averages to analyze the stochastic behavior of groundwater flows -- the so-called Monte Carlo approach. But the high heterogeneity of porous media, the fine discretization necessary for three dimensional numerical approximations and the large size of real world natural systems lead to overwhelming requirements for computational memory and speed, even for modern days computers. On the other hand the physical law may directly be incorporated in the general knowledge in the BME approach, leading to a prior pdf that is non-Gaussian in general. In order to contrast the two approaches, I'll present first the classical approach for the Darcy equation by means of an implementation of the Space Transformations method which I developed and tested, and I'll then discuss how a physical law may on the other hand be directly incorporated as part of the general knowledge base.

<u>A classical example: The three-dimensional flow equation, and it's solution by means of Space Transformations</u>

The steady state flow equation for three dimensional porous media may be written as

$$\sum_{j=1}^{3}[\frac{\partial^2 \mathrm{H}_3(\boldsymbol{s})}{\partial \mathrm{s}_j^2} + \frac{\partial \ln \mathrm{K}_3(\boldsymbol{s})}{\partial \mathrm{s}_j}\frac{\partial \mathrm{H}_3(\boldsymbol{s})}{\partial \mathrm{s}_j}] = 0\,, \tag{3.23}$$

where $\mathrm{H}_3(\boldsymbol{s})$ is the hydraulic head and $\mathrm{K}_3(\boldsymbol{s})$ is the hydraulic conductivity. Using the numerical discretization of the Space Transformation method (see Appendix E), the three-dimensional hydraulic head is given by the expression

$$\mathrm{H}_3(\boldsymbol{s}) = -\frac{1}{2\pi}\,\Psi_1^3[\hat{\mathrm{J}}_{1,\boldsymbol{\theta}}(0)\,\tilde{\kappa}_{\boldsymbol{\theta}}(\boldsymbol{s}\cdot\boldsymbol{\theta})\exp\{-\int_0^{\boldsymbol{s}\cdot\boldsymbol{\theta}}\mathrm{du}\,\tilde{\kappa}_{\boldsymbol{\theta}}(\mathrm{u})\}]\ , \tag{3.24}$$

where

$$\hat{\mathrm{J}}_{1,\boldsymbol{\theta}}(0) = -\frac{\partial \hat{\mathrm{H}}_{1,\boldsymbol{\theta}}(\sigma)}{\partial \sigma}\bigg|_{\sigma=0} \tag{3.25}$$

is a function that depends on the direction vector and represents the unidimensional boundary conditions (BC). The solution (3.24) is based on an ST of the three-dimensional flow equation that leads to unidimensional flow equations in the direction of the ST lines. In the above, surface terms generated by the ST of the partial derivatives are neglected. The function $\tilde{\kappa}_{\boldsymbol{\theta}}(\cdot)$ is given by the ratio of two ST that involve both the unknown local three-dimensional hydraulic gradient and the hydraulic conductivity. In the case of a constant specific discharge vector $\boldsymbol{Q} = -\mathrm{K}_3(\boldsymbol{s})\nabla \mathrm{H}_3(\boldsymbol{s})$, it is possible to express the function $\tilde{\kappa}_{\boldsymbol{\theta}}(\cdot)$ purely in terms of the hydraulic conductivity ST as

$$\tilde{\kappa}_{\boldsymbol{\theta}}(\mathrm{p}) = \frac{T_3^1\left[\mathrm{K}_3^{-1}(\boldsymbol{s})\left(\boldsymbol{\theta}\cdot\nabla\ln\mathrm{K}_3(\boldsymbol{s})\right)\right]\ (p,\boldsymbol{\theta})}{T_3^1\left[\mathrm{K}_3^{-1}(\boldsymbol{s})\right]\ (\mathrm{p},\boldsymbol{\theta})}\,. \tag{3.26}$$

I implemented numerically this approach for the case of flow in a domain with random hydraulic conductivity (Hristopulos *et al.*; 1999). A correlated log-conductivity random field with unidimensional variability in the direction $\boldsymbol{e}_2$ was simulated with a Gaussian covariance

$$c_f(h) = \sigma_f^2 \exp[-\frac{h_2^2}{\ell^2}], \tag{3.27}$$

where the correlation length was assumed to be $\ell = 3$.. In Fig. 3.1 I show the longitudinal profile of the hydraulic head calculated using hydraulic conductivity mean $\overline{K} = 17.3$ and standard deviation $\sigma_K = 10$. The solution obtained is found to be in good agreement with that obtained from the Finite Difference Method (FDM), as seen in Fig. 3.1. This work showed that the Space Transformation method allows to obtain accurate solutions at a numerical cost that was competitive when compared to other numerical methods (such as the FDM). However the many realizations of the porous media necessary in the Monte Carlo approach makes this approach very expensive in practice. Hence in order to study the effect of porous media heterogeneity to the stochastic behavior of the hydraulic heads, it would be much more convenient to directly include the Darcy physical law into a mapping approach for hydraulic head. This is exactly what the BME approach offers, it allows to incorporate a physical law in the general knowledge, and processes that information by means of the general knowledge operator. Showing how the general knowledge operator may include a physical law is best done by means of a simple example. Hence this is done next for a simple physical law governing a temporal random field.

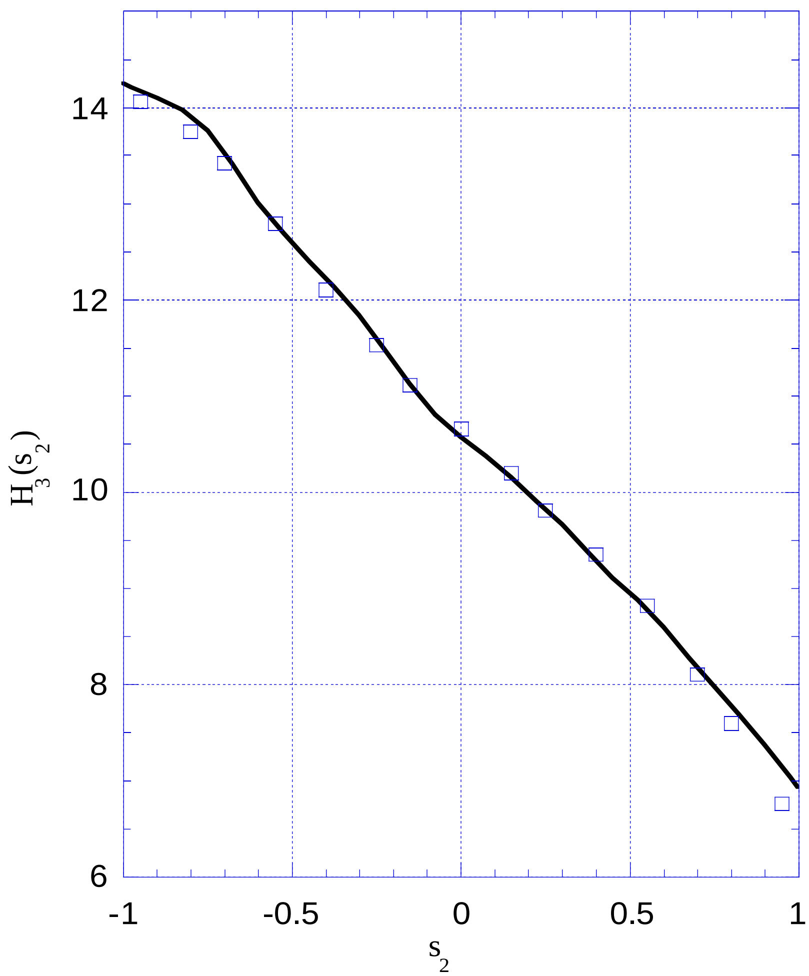

Figure 3.1: Hydraulic head profile versus the coordinate in the flow direction for a unidimensional, correlated, log-normal hydraulic conductivity random field with $\overline{K} = 17.3$ and $\sigma_K = 10$. The FDM solution (continuous line) and the ST-estimate (squares) are shown.

<u>Incorporating physical laws in the general knowledge operator: A small example</u>

A general framework that permits to incorporate physical laws expressed in various forms, i.e., by means of algebraic functions or differential equations, is given in Chapter VIII. However let's consider here a small example to illustrate the principles.

Consider a S/TRF $X(t)$ governed by the following physical law

$$\frac{dX(t)}{dt} + Y(t)X(t) = 0 \tag{3.27}$$

In this example $X(t)$ and $Y(t)$ are random fields depending on time t. The natural variable $X(t)$ may represent a pollutant concentration and $Y(t)$ some soil property. Let's assume that $Y(t)$ has a known mean $m_y(t)$ and $\sigma_y^2(t)$. The information we want to incorporate in the general knowledge operator are the mean and the variance of $Y(t)$ and some additional information derived from the physical law (3.27). Let us consider the multivariate pdf $f_G(\psi,\chi;t)$ for the random fields $Y(t)$ and $X(t)$ at time t. The general knowledge is incorporated by defining the functions $g_\alpha(\boldsymbol{\chi}_{\text{map}})$ of Eqs. (3.6). The knowledge of the mean $m_y(t)$ and variance $\sigma_y^2(t)$ of $Y(t)$ is obtained by selecting

$$g_1(\psi,\chi) = \psi; \qquad \bar{g}_1 = m_y(t), \tag{3.28}$$

$$g_2(\psi,\chi) = \psi^2; \qquad \bar{g}_2 = \sigma_y^2(t) + m_y^2(t). \tag{3.29}$$

As for the physical law, we may extract from it several constraints of the form of Eqs (3.6). Let's for instance take the expected value of Eq (3.27) to get

$$\overline{Y(t)X(t)} = -\frac{d\overline{X}(t)}{dt} = -\int d\chi\ \chi\ \frac{df_G(\psi,\chi;t)}{dt} \tag{3.30}$$

This information is incorporated by using the following constraint

$$g_3(\psi,\chi) = \psi\chi\,; \qquad \bar{g}_3 = -\frac{d\overline{X}(t)}{dt}. \tag{3.31}$$

Another useful information is provided by multiplying Eq. (3.27) by $X(t)$ and then taking the mean, so that $\overline{Y(t)X^2(t)} = -(1/2)\, d\overline{X^2}(t)/dt$. By integrating with respect to time we may write

$$\overline{X^2}(t) = \overline{X^2}(0) - 2\int_0^t du\,\overline{Y(u)X^2(u)}. \tag{3.32}$$

Eq. (3.32) is incorporated in the general knowledge by using the following constraint

$$g_4(\psi,\chi) = \chi^2 \qquad \bar{g}_4 = -2\int_0^t du\,\overline{Y(u)X^2(u)} \tag{3.33}$$

The constraints (3.28), (3.29), (3.31) and (3.33) are summarized in table 3.1. The constraints of table 3.1 are adequate for the purpose of including general knowledge about both the moments of $Y(t)$ and the physical law governing $X(t)$. The posterior pdf incorporating the general knowledge of table (3.1) is given by Eq (3.8) and can be expressed here as

$$f_G(\psi,\chi;t) = Z^{-1}\exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] \tag{3.34}$$

where $\boldsymbol{\mu}(t) = [\mu_1(t),\mu_2(t),\mu_3(t),\mu_4(t)]$ are the Lagrange coefficient, and $Y_G(\psi,\chi;\boldsymbol{\mu}(t))$ is given by

$$Y_G(\psi,\chi;\boldsymbol{\mu}(t)) = \mu_1(t)\,\psi + \mu_2(t)\,\psi^2 + \mu_3(t)\,\psi\chi + \mu_4(t)\,\chi^2 \tag{3.35}$$

The unknown coefficients $\boldsymbol{\mu}(t)$ are obtained by solving the system of *integrodifferential* equations obtained from Table 3.1, i.e.

$$\begin{cases} \int d\psi\, d\chi\, \psi\, Z^{-1} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] = m_y(t) \\ \int d\psi\, d\chi\, \psi^2\, Z^{-1} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] = \sigma_y^2(t) + m_y^2(t) \\ \int d\psi\, d\chi\, \psi\chi\, Z^{-1} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] = -\int d\psi\, d\chi\, \chi\, Z^{-1} \dfrac{dY_G(\psi,\chi;\boldsymbol{\mu}(t))}{dt} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] \\ \int d\psi\, d\chi\, \chi^2\, Z^{-1} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(t))] = \overline{X^2}(0) - 2\int_0^t du \int d\psi\, d\chi\, \psi\chi^2\, Z^{-1} \exp[Y_G(\psi,\chi;\boldsymbol{\mu}(u))] \end{cases} \tag{3.36}$$

This system of equations has a solution $\boldsymbol{\mu}(t)$ for a given set of initial conditions $\boldsymbol{\mu}(0)$, and $\dfrac{d\boldsymbol{\mu}(0)}{dt}$, which is usually obtained using numerical methods.

As demonstrated in this example, the general knowledge operator is a powerful epistemological principle of the BME method which allows to incorporate many sources of information in the spatiotemporal BME mapping solution. The general knowledge operator takes different forms depending on the type of general knowledge available, which combined with the specificatory knowledge operator, provides the posterior pdf (Eq. 3.1) of BME mapping. Using the BME posterior pdf it is possible to derive several estimators of the natural variable depending on the mapping situation at hand, as shown next.

TABLE 3.1: General knowledge constraints

α	g_α	$\bar{g}_\alpha$
1	$g_1(\psi,\chi) = \psi$	$\bar{g}_1 = m_y(t)$
2	$g_2(\psi,\chi) = \psi^2$	$\bar{g}_2 = \sigma_y^2(t) + m_y^2(t)$
3	$g_3(\psi,\chi) = \psi\chi$	$\bar{g}_3 = -\dfrac{d\overline{X}(t)}{dt}$
4	$g_4(\psi,\chi) = \chi^2$	$\bar{g}_4 = -2\int_0^t du \overline{Y(u)X^2(u)}$

3.3. BME Estimates

The BME method provides a full description of the map $X(\boldsymbol{p_k})$ by means of its posterior pdf $f_K(\boldsymbol{\chi}_{\mathrm{k}})$. For mapping purposes it is valuable to obtain estimates $\hat{X}(\boldsymbol{p_k})$ of $X(\boldsymbol{p_k})$. BME provides the flexibility to choose among several estimates, and the choice of the best suited estimate will depend on the mapping situation. Indeed, knowing the posterior pdf allows the flexibility to obtain several estimates, such as the mode, the mean, the median, or any percentile we want. The selection of the appropriate estimate depends on theoretical, computational and physical considerations regarding the mapping situation. Below we consider some of the estimates commonly used.

The *BME mode* estimate --which represents the most probable realization-- is clearly the value that maximizes the posterior pdf. In this case, each estimation point is considered independently, the posterior pdf $f_K(\chi_{\mathrm{k}})$ is univariate, and its mode is referred to as the single-point estimator $\hat{\chi}_{\mathrm{k}}$.

Another estimate of interest in single-point mapping is the *BME mean* estimate defined as $\overline{x}_{k|K} = \int d\chi_k \, \chi_k \, f_K(\chi_k)$, where the subscript K expresses dependence on the total knowledge. The mean estimate is suitable for mapping situations where one is interested in minimizing the mean square estimation error.

The estimation points are usually located on a regular grid that includes the points $\boldsymbol{p}_i$ $(i = 1,\ldots,m)$ where physical data $\boldsymbol{\chi}_{\mathrm{data}}$ are available. The estimated values are used to create spatiotemporal maps, which can be scientifically interpreted to provide a useful picture of reality. However, because of the inherent randomness of the natural process and physical data inaccuracies, it is essential to obtain an assessment of the uncertainty associated with the estimated values. This important issue is studied next.

3.4. Uncertainty Assessment

Often, a measure of the uncertainty is offered by the variance of the estimation error (see, e.g., the geostatistical estimation error variance; Olea, 1997; Bogaert and Christakos, 1997). BME can certainly provide this simple measure of accuracy --as typically done with traditional geostatistical methods. In addition, BME allows a more sophisticated and accurate assessment of the estimation error by means of the posterior pdf $f_K(\chi_k)$ derived on the basis of the total physical knowledge K. Using $f_K(\chi_k)$ one can calculate confidence intervals --which provide a more realistic assessment of estimation error than merely an error variance. The confidence interval concept can be generalized to lead to the BME confidence set, as will be discussed in Chapter 7. for the general case of multi-point BME mapping. Confidence sets provide an appropriate description of the uncertainty associated with the spatiotemporal mapping of a natural variable at several estimation points simultaneously, and can easily be extended to the case involving joint mapping of several natural processes.

A simple measure of estimation error in single-point mapping is the variance of the BME posterior pdf, i.e.,

$$\sigma_{k|K}^2 = \int d\chi_k (\chi_k - \bar{x}_{k|K})^2 f_K(\chi_k). \tag{3.37}$$

This quantity corresponds to the variance of the estimation error $e_k = x_k - \bar{x}_{k|K}$. The $\sigma_{k|K}^2$ is an accuracy measure typically reported by traditional estimation methods, and it provides an adequate assessment of estimation error when the shape of the posterior pdf is not very complicated. For a Gaussian posterior pdf, e.g., the probability that x_k lies in the interval $[\bar{x}_{k|K} - 1.96\,\sigma_{k|K},\ \bar{x}_{k|K} + 1.96\,\sigma_{k|K}]$ is 95%.

Another realistic assessment of the mapping error in the case of a posterior pdf with asymmetric shape may be achieved using confidence sets of smallest size given by the following proposition:

The confidence interval of smallest size is an interval written as $\mathrm{C}_\eta = [\hat{\boldsymbol{\chi}}_k - a, \hat{\boldsymbol{\chi}}_k + b]$ *where a and b are such that*

$$\mathrm{f}(\hat{\chi}_k - \mathrm{a} \,|\, \boldsymbol{\chi}_{data}) = \mathrm{f}(\hat{\chi}_k + \mathrm{b} \,|\, \boldsymbol{\chi}_{data}). \tag{3.38}$$

A proof of the above proposition will be given in Chapter 7 as a special case of BME confidence sets. For illustration purpose see the BME posterior pdf in Figure 3.2. The confidence intervals corresponding to different confidence levels are shown, as obtained from the above proposition.

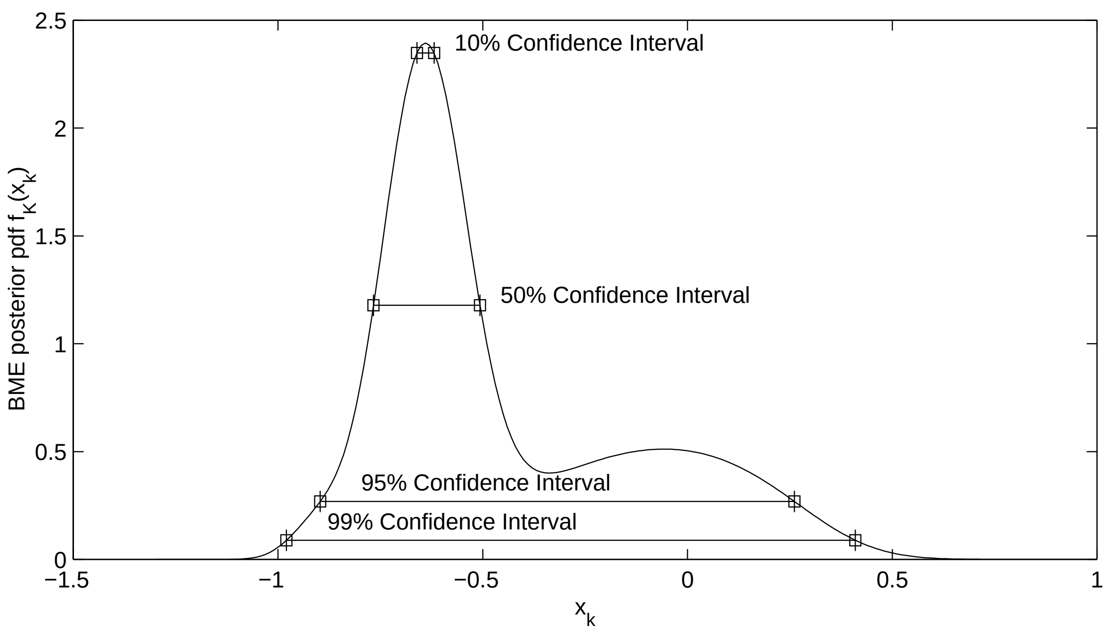

Figure 3.2: BME confidence intervals plotted on a typical BME posterior pdf.