II. INTRODUCTION TO SPACE/TIME RANDOM FIELD MODELLING IN THE LIGHT OF UNCERTAIN PHYSICAL KNOWLEDGE

Most natural variables assume values at points within a spatiotemporal continuum. Each point in this continuum is represented by its space and time coordinates $\boldsymbol{p} = (\boldsymbol{s}, t)$. Randomness of a natural variable manifests itself as an ensemble of possible realizations over the space/time continuum. This leads to the concept of the spatiotemporal random field (S/TRF) $X(\boldsymbol{p})$, which is a collection of realizations (possibilities) of the spatiotemporal distribution of the natural variable. The value taken by the random field $X(\boldsymbol{p})$ at a given space/time point $\boldsymbol{p}_i$ is called a random variable and is noted as $x_i = X(\boldsymbol{p}_i)$. Herein, we will use capital English letter, e.g. $X(\boldsymbol{p})$, to denote random fields, small English letters x_1, x_2, etc. to denote random variables, and small Greek letters χ_1, χ_2, etc. to denote their realizations. For convenience we will let the vector of random variables $\boldsymbol{x} = [x_1 \ldots x_m]^T$ denote the random field $X(\boldsymbol{p})$ at points $\boldsymbol{p}_i$ (i=1,...m), and the vector $\boldsymbol{\chi} = [\chi_1 \ldots \chi_m]^T$ denote its realization. In spatiotemporal mapping we are interested in estimating the values of $X(\boldsymbol{p})$ at point $\boldsymbol{p}_k$ given physical data at space/time points $\boldsymbol{p}_i$ (i=1,...m). We let $\boldsymbol{x}_{\text{map}} = [x_1, \ldots, x_m, x_k]^T$ denote $X(\boldsymbol{p})$ at points $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, m, k)$, and the subsets of these values at points $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, m)$, where (possibly uncertain) physical knowledge is available, is noted as $\mathbf{x}_{\text{data}} = [x_1, \ldots, x_m]^T$. Having stated the framework in which STRF will be used, we will need some foundation for their mathematical interpretation. In this chapter we will give some short mathematical preliminaries on probability space, random variables and random fields. We will then introduce the concept of uncertain physical knowledge, and we will revisit traditional STRF modelling in the light of uncertain physical knowledge.

2.1. Introduction to Space/Time Random Fields Modelling

2.1.1. The Space/Time Random Field

The formal description of probability space and random variable are based on set-theoretic notions, and some useful mathematical preliminaries are provided in Appendix 2.1. In short a random variable x may acquire any one value from a distribution of values. The distribution of values χ that the random variable x may take is described for a given probability space (Ω, F, P) by the cumulative distribution function (cdf)

$$F_x(\chi) = P[x \le \chi]$$

The derivative of the cdf $F_x(\chi)$ with respect to χ, when it exists, is called the probability density function (pdf) of the random variable x, namely

$$f_x(\chi) = \frac{\partial F_x(\chi)}{\partial \chi} \tag{2.1}$$

As explained in more details in Appendix 2.1, the probability measure P on a measurable space (Ω, F), where Ω is the sample space and F a σ-algebra of subsets of Ω, is a function $P: F \to [0,1]$ satisfying Kolmogorov's axioms, i.e.

(a) $P(\Omega) = 1$

(b) $0 \le P(A_i) \le 1$ for all sets $A_i \in F$

(c) if $A_i \cap A_j = \varnothing$ (i≠j), then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

Any function possessing these properties is a valid cdf, and a random variable is said to be completely characterized if its cdf or pdf is known.

A random field describes a natural process which takes values at points within a space/time continuum. Space/time is viewed as a continuous spatial arrangement combined with a temporal order of events. This union of space and time is defined in terms of their Cartesian product. Spatial coordinates $\boldsymbol{s} = (s_1, ..., s_n)$ are elements of the Euclidean space $\Re^n$, i.e. $\boldsymbol{s} \in S \subset \Re^n$, and coordinate t is along the time axis $T \subset \Re^1$. The space/time coordinates $\boldsymbol{p} = (\boldsymbol{s}, t)$ are defined on the Cartesian product $S \times T$, i.e. $\boldsymbol{p} = (\boldsymbol{s}, t) \in S \times T$. A formal description of random field arises from set-theoretic notions, and the readers is referred to Christakos; 1992 for a detailed treatment. Conceptually a spatiotemporal random field $X(\boldsymbol{p})$ may be defined as a parameterized collection of random variables

$$X(\boldsymbol{p}) = \{X_{\boldsymbol{p}}\}_{\boldsymbol{p} \in \Re^n \times [0,\infty)} \tag{2.2}$$

where the random variables $X_{\boldsymbol{p}}$ are defined on a probability space, and the space/time point $\boldsymbol{p}$ takes values in the space/time continuum $S \times T$. For mapping purposes it is sufficient to consider the random field $X(\boldsymbol{p})$ as being the family of random variables $\{x_1, x_2, ...\}$ at some specified points $\boldsymbol{p}_1, \boldsymbol{p}_2, ...$ (usually the data and estimation points). In this latter view the random field $X(\boldsymbol{p})$ is completely described by the join (or multivariate) cdf

$$F_{\boldsymbol{x}}(\chi_1, \chi_2, \ldots, \chi_m) = P[x_1 \le \chi_1, x_2 \le \chi_2, \ldots, x_m \le \chi_m] \tag{2.3}$$

In order to be valid, the multivariate cdf must satisfy Kolmogorov's axioms, and the derivative of the cdf with respect to $\boldsymbol{\chi} = [\chi_1, \chi_2, \ldots, \chi_m]^T$, when it exists, is called the multivariate pdf of the random vector $\boldsymbol{x} = [x_1, x_2, ..., x_m]^T$, namely

$$f_{\boldsymbol{x}}(\chi_1, \chi_2, \ldots, \chi_m) = \frac{\partial^n F_{\boldsymbol{x}}(\chi_1, \chi_2, \ldots, \chi_m)}{\partial \chi_1 \partial \chi_2, \ldots \partial \chi_m} \tag{2.4}$$

A random field would be completely characterized by its multivariate cdf or pdf if known, however in most cases this probability law is not known a priori. In this case the random field may be incompletely characterized in terms of ensemble functions of the general form

$$\overline{\Phi[X(\boldsymbol{p}_1),\ldots X(\boldsymbol{p}_m)]} = \int d\chi_1 \ldots \int d\chi_m \, \Phi[\chi_1,\ldots,\chi_m] \, f_x(\chi_1,\ldots,\chi_m;\boldsymbol{p}_1,\ldots,\boldsymbol{p}_m) \tag{2.5}$$

where $f_x(\chi_1,\ldots,\chi_m;\boldsymbol{p}_1,\ldots,\boldsymbol{p}_m)$ is the multivariate pdf for the S/TRF $X(\boldsymbol{p})$ taken at points $\boldsymbol{p}_1,\ldots,\boldsymbol{p}_m$., $\overline{[\ldots]}$ denotes the operation of stochastic expectation and $\Phi[\cdot]$ is a function of $\chi_1,\ldots,\chi_m$. Statistical moments of order p are obtained by setting the functional $\Phi[\cdot]$ to a polynomial in $\chi_1,\ldots,\chi_m$ of order p. Certain moments can often be estimated from the data or physical models, and they provide information regarding the properties of the pdf. Commonly used moments are the moments of order one and two, namely the mean and covariance. The mean is given by

$$m_x(\boldsymbol{p}) = \overline{X(\boldsymbol{p})} = \int \chi \;\; f_x(\chi;\boldsymbol{p}) \;\; d\chi \tag{2.6}$$

where $f_x(\chi;\boldsymbol{p})$ represents the (univariate) pdf for the random field at point $\boldsymbol{p}$. The (centered) space-time covariance function is given by

$$\begin{aligned} c_x(\boldsymbol{p},\boldsymbol{p}') &= \overline{(X(\boldsymbol{p}) - m_x(\boldsymbol{p}))(X(\boldsymbol{p}') - m_x(\boldsymbol{p}'))} \\ &= \int\int (\chi - m_x(\boldsymbol{p}))(\chi' - m_x(\boldsymbol{p}')) \;\; f_x(\chi,\chi';\boldsymbol{p},\boldsymbol{p}') \;\; d\chi d\chi' \end{aligned} \tag{2.7}$$

where $f_x(\chi,\chi';\boldsymbol{p},\boldsymbol{p}')$ represents the bivariate pdf of the random field taken between points $\boldsymbol{p}$ and $\boldsymbol{p}'$. The covariance function provides information about the spatial distribution of the variable and the fluctuation correlations, but not its strength. The strength of the fluctuations of the variable at a given point, or local variance of the fluctuations, is given by the value of the covariance at zero lag (i.e. when $\boldsymbol{p}' = \boldsymbol{p}$) as

$$\sigma_x^2(\boldsymbol{p}) = c_x(\boldsymbol{p},\boldsymbol{p}) \tag{2.8}$$

Then the *correlation function* gives information about the strength of the fluctuation correlations by normalizing the covariance with the local variances as follows

$$\rho_x(\boldsymbol{p},\boldsymbol{p}') = c_x(\boldsymbol{p},\boldsymbol{p}')/\sigma_x(\boldsymbol{p})\sigma_x(\boldsymbol{p}'). \tag{2.9}$$

We now consider briefly some symmetry properties of S/TRFs. If all multivariate pdf's are invariant for all translations in space and time the S/TRF is called *statistically homogeneous-stationary* in the *strict sense*. Hence, strict homogeneity-stationary requires that

$$F_X[\chi_1(\boldsymbol{s}_1,t_1),\ldots,\chi_N(\boldsymbol{s}_N,t_N)] = F_X[\chi_1(\boldsymbol{s}_1+\boldsymbol{r},t_1+\tau),\ldots,\chi_N(\boldsymbol{s}_N+\boldsymbol{r},t_N+\tau)] \tag{2.10}$$

for all $N = 1,2,\ldots$, and for all spatial lags $\boldsymbol{r}$ and temporal lags τ, where F_X denotes the joint probability distribution. On the other hand, if the first two moments are invariant for all translations in space and time, the S/TRF is called *statistically homogeneous-stationary* in the *wide sense* (also, *second order* or *weakly* homogeneous). While strict homogeneity-stationary implies weak homogeneity-stationary, the converse is not always true. Thus, weak homogeneity-stationary requires only that the S/TRF $X(\boldsymbol{s},t)$ have a constant mean,

$$m_X(\boldsymbol{p}) = m_X; \tag{2.11}$$

and that the covariance be a function only of the spatial lag vector $\boldsymbol{r} = \boldsymbol{s} - \boldsymbol{s}'$ and temporal lag $\tau = t - t'$ between the two points, i.e.,

$$C_x(\boldsymbol{p};\boldsymbol{p}') = C_x(\boldsymbol{s}-\boldsymbol{s}';t-t') = C_x(\boldsymbol{r},\tau), \tag{2.12}$$

If a S/TRF is homogeneous-stationary in the wide sense, it can be further classified as *isotropic* (in the wide sense) if the covariance function depends only on the spatial lag distance; e.g.,

$$C_X(\boldsymbol{p};\boldsymbol{p}') = C_X(r;t-t'),\ \ r=|\boldsymbol{r}|. \tag{2.13}$$

A S/TRF that is not isotropic will display anisotropic spatial dependence.

If the second order ensemble moments are equal to the second order sample moments then the S/TRF is called *statistically ergodic* in the wide sense. While the ensemble moments are calculated in terms of the corresponding pdf's, the sample moments are calculated on the basis of a single realization which is usually available in practice. This will become important later, when we will discuss the subject of covariance estimation of a S/TRF from experimental data.

A S/TRF will be homogeneous-stationary only in the absence of any spatial or temporal trend. In the presence of such trends, the mean and covariance will be dependent on the location and the time of measurements; also the covariance will not decay to zero with increase in distance and time. However it is possible to model such fields as the composition of a homogeneous random field and some space/time trend. Methods exists to estimate the space time trend (Bogaert; 19xx). Therefore, since this work focuses on uncertain physical knowledge, we will without loss of generality assume that the space/time trend is know, and we will concentrate on the homogeneous-stationary random field component which accounts for all variability of the natural variable. The theory presented

herein will offer foundations which may allow for extension to the case of intrinsic random fields (i.e. fields that are non-homogeneous-non stationary, Christakos; 1992).

2.1.2. A Review of Some Space-Time Covariance Models

The covariance function describe important characteristics of a S/TRF, such as its variability (or randomness), its correlation in space and time, its continuity, etc. For instance a S/TRF is continuous in the mean square sense at the point $\boldsymbol{p}_0$, if and only if its covariance $c_x(\boldsymbol{p},\boldsymbol{p}')$ is continuous at the point $\boldsymbol{p}=\boldsymbol{p}'=\boldsymbol{p}_0$. It is therefore useful to present some general properties of covariance function and to review some commonly used covariance models.

Based on Definition of the covariance function, it can be shown that covariance functions must satisfy the following properties:

$$c_x(\boldsymbol{p},\boldsymbol{p}')=c_x(\boldsymbol{p}',\boldsymbol{p}). \tag{2.14}$$

which means that, in the case of homogeneous-stationary random fields, we must have $c_x(\boldsymbol{r},\tau)=c_x(-\boldsymbol{r},-\tau)$. Another property, known as the Schwartz inequality, provides an upper bound for the covariance between any two points in space-time in terms of the local variances as follow

$$|c_x(\boldsymbol{p},\boldsymbol{p}')|\le \sigma_x(\boldsymbol{p})\sigma_x(\boldsymbol{p}'). \tag{2.15}$$

which, in the case of homogeneous-stationary random fields, leads to $c_x(\boldsymbol{r},\tau)\le\sigma_x^2$ for all spatial lags and temporal lags. An additional property of covariance function is that the following bilinear form is non-negative for any number N of space-time points, and arbitrary numbers $q_i,\ i=1,\ldots,N$,

$$\sum_{i=1}^{N}\sum_{j=1}^{N} c_x(\boldsymbol{p}_i,\boldsymbol{p}_j)\,q_i\,q_j \geq 0 \qquad (2.16)$$

which follows from the fact that the quantity $\overline{\left|\sum_{i=1}^{N}[X(\boldsymbol{p}_i)-\overline{X(\boldsymbol{p}_i)}]q_i\right|^2}$ is non-negative.

A continuous function that satisfies properties (2.14)-(2.16) above is called a *non-negative definite function*. Hence, the covariance functions of mean square continuous S/TRFs are non-negative definite functions. The converse can also be shown, namely that every non-negative definite function is a covariance function of a mean square continuous random field. A function that verifies the non-negative definite properties is said to be permissible, and is a candidate to model a random field.

There exist many covariance models, but they are not all good candidates to model natural variables. For environmental and health processes, a common choice is to use separable space-time models. These are covariance models of the general form

$$c_x(\boldsymbol{s},t;\boldsymbol{s}',t') = c_{x(1)}(\boldsymbol{s},\boldsymbol{s}')c_{x(2)}(t,t'), \qquad (2.17)$$

where $c_{x(1)}(\boldsymbol{s},\boldsymbol{s}')$ is a purely spatial and $c_{x(2)}(t,t')$ a purely temporal covariance function. The space-time covariance $c_x(\boldsymbol{s},t;\boldsymbol{s}',t')$ is permissible as long as its components $c_{x(1)}(\boldsymbol{s},\boldsymbol{s}')$ and $c_{x(2)}(t,t')$ are valid covariance functions. The assumption of separability is convenient in the estimation of the covariance model from experimental data, since the spatial and temporal dependencies can be investigated separately. In the light of the separability assumption we present some covariance functions that are routinely used to model homogeneous-stationary random fields in environmental and health sciences.

0 - The **nugget** space-time model, defined by the nugget variance c_0:

$$c_x(\boldsymbol{r},\tau) = c_0\,\delta_{\boldsymbol{r}=\boldsymbol{0}}\,\delta_{\tau=0} \tag{2.18}$$

where $\delta_{(Z)}$ is 1 if z is true, 0 otherwise.

1 - The **spherical** space-time model, defined by the positive variance contribution or *sill* value c_0, the spatial range parameter a_r, and the temporal range parameter a_t:

$$c_x(\boldsymbol{r},\tau) = c_0\left[1-1.5\frac{||\boldsymbol{r}||}{a_r}+0.5\frac{||\boldsymbol{r}||^3}{a_r^3}\right]\delta_{||\boldsymbol{r}||<a_r}\left[1-1.5\frac{|\tau|}{a_r}+0.5\frac{|\tau|^3}{a_r^3}\right]\delta_{|\tau|<a_\tau} \tag{2.19}$$

2 - The **exponential** space-time model, defined by the positive variance contribution or *sill* value c_0, the spatial range parameter a_r, and the temporal range parameter a_t:

$$c_x(\boldsymbol{r},\tau) = c_0\exp(-\frac{|\boldsymbol{r}|}{a_r})\exp(-\frac{|\tau|}{a_t}) \tag{2.20}$$

3 - The **gaussian** space-time model, defined by the positive variance contribution or *sill* value c_0, the spatial range parameter a_r, and the temporal range parameter a_t:

$$c_x(\boldsymbol{r},\tau) = c_0\exp(-\frac{||\boldsymbol{r}||^2}{a_r^2})\exp(-\frac{\tau^2}{a_t^2}) \tag{2.21}$$

In practice we may use one of the models presented, or any positive linear combination of these models. In that case the covariance is said to have *nested* structures. One common example is the nugget effect, which seldom occurs alone (pure nugget effect), but rather in combination with models 1 to 3.

2.1.2 Choosing a Covariance Model from Experimental Data

In spatiotemporal mapping it is usually necessary to know what is the covariance of a S/TRF, and the traditional approach is to chose a covariance model from experimental data.

In the case of isotropic homogeneous-stationary random fields the covariance $c_x(r,\tau)$ is only a function of the spatial lag $r = |\boldsymbol{s} - \boldsymbol{s}'|$ and the temporal lag $\tau = |t - t'|$. Assuming the random field is stochastically *ergodic* in the wide sense, it is possible to estimate the covariance from observed (or experimental) values as follow

$$c_x(r,\tau) \approx \frac{1}{N(r,\tau)} \sum_{i=1}^{N(r,\tau)} (X_{-(r,\tau),i}\, X_{+(r,\tau),i}) - m_{-(r,\tau)} m_{+(r,\tau)}, \qquad (2.22)$$

where $N(r,\tau)$ is the number of pairs of observed values $(X_{-(r,\tau),i},\, X_{+(r,\tau),i})$ separated by the spatial and temporal lags r and t, $m_{-(r,\tau)}$ is the mean of the $X_{-(r,\tau),i}$ values and $m_{+(r,\tau)}$ is the mean of the $X_{+(r,\tau),i}$ values. We note that the observed values correspond actually to only one *realization* of the random field. The calculated value is therefore only a sample average, i.e. one calculated from sample values, but through the ergodicity assumption, this value is believed to be equal to an ensemble average, i.e. an average that would be calculated using different realizations.

Using Eq. (2.22) one obtains estimated values for the covariance corresponding to different spatial and temporal lags. However these are only estimated values polluted by some estimation error, and there is no guaranty that they verify the properties of a permissible covariance function. The solution routinely adopted to this problem is to fit a valid covariance model to the experimental values. This method insures that the model obtained is permissible, and it is said to be parametric in the sense that it determines the covariance through its *parameters* (e.g., sill and range parameters in Eq. 2.18 - 2.21)

Choosing a covariance model from experimental data is often the first step in traditional spatiotemporal analysis and mapping. Once the covariance model has been selected, one may proceed to generate stochastic simulations in order to obtain more insight about the S/TRF.

2.1.4. Stochastic Simulation of S/TRF

Stochastic simulation is the process of generating alternative, equally probable, realizations of the space/time random field. The (usually gridded) realizations represent possible images of the space/time distribution of the attribute value over the space time continuum. Each realization, also called a "stochastic image", reflects the properties that have been imposed on the S/TRF. There exist many simulation techniques which may incorporate different properties of the S/TRF.

Typically the S/TRF model is constrained by the sole covariance model, which may be obtained by fitting parameters to experimental data, as described earlier. Simulations are useful for many purposes. They may be used to provide visual insights about the characteristics corresponding to given properties of a random field. Consider for example the realization shown in Fig. 2.1 for a random field in a two dimensional spatial domain, obtained by imposing the exponential covariance model of Eq. 2.20, with variance c_0=1 and a_r=1. This picture represent a possible realization of the random field on a 20 by 20 grid. Many such realizations may be produced, so that their ensemble average yield a covariance in agreement with the model imposed. If in addition to a known covariance model, some observed values for the random field are known, then it is possible to draw realization that will honor the observed values. Such simulations are said to be conditional (to the observed value).

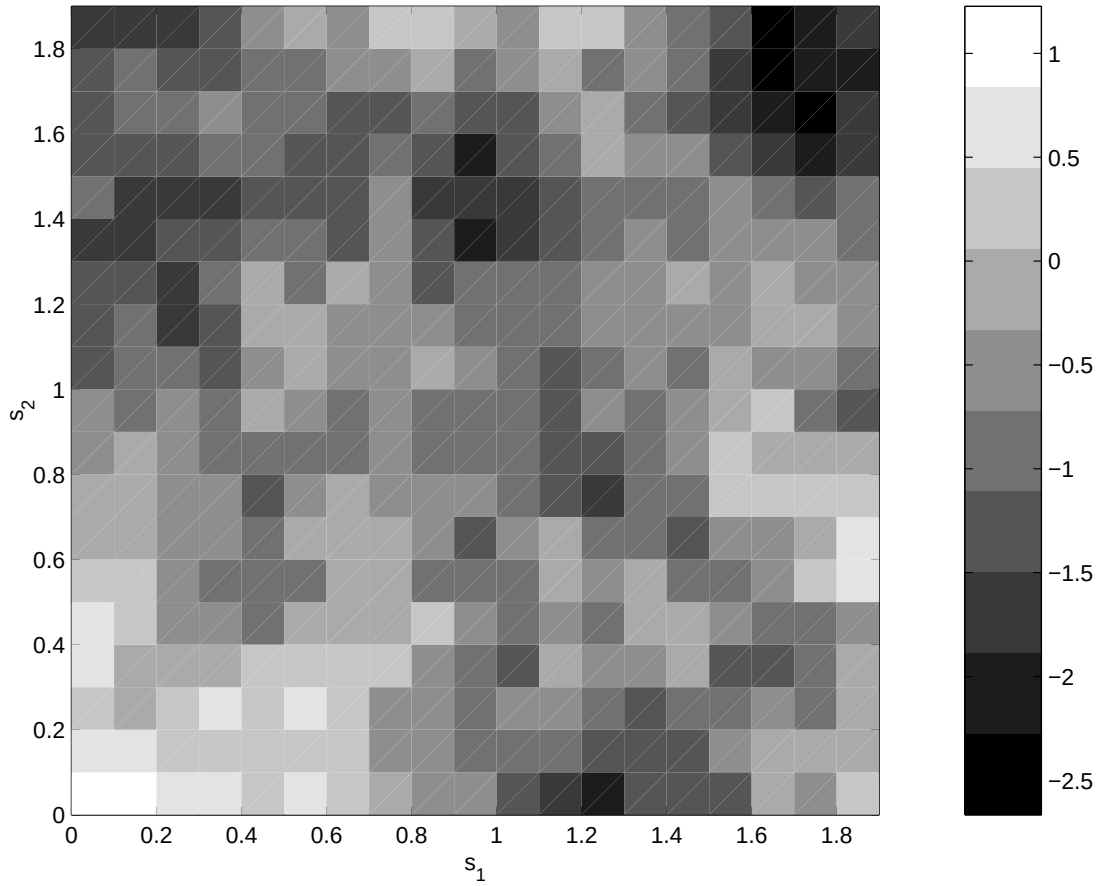

Figure 2.1: A simulated realization of random field with exponential covariance, c_0=1, a_r=1.

One simulation technique that will be used later in this work is the lower-upper triangular technique. The LU simulation technique produces Gaussian S/TRF corresponding to an imposed covariance model. Let $c_x(\boldsymbol{p}_i, \boldsymbol{p}_j)$ be the covariance function of a zero mean S/TRF $X(\boldsymbol{p})$. Assume that we seek simulations at m (usually gridded) points. The application of the LU approach to this kind of problem consist of three steps given in Table 2.1 (from Christakos; 1992, p 327).

TABLE 2.1: The LU simulation technique

Step 1: Calculate the $m \times m$ covariance matrix corresponding to the covariance model $c_x(\boldsymbol{p}_i, \boldsymbol{p}_j)$ at the m space/time points $\boldsymbol{p}_i$, $i = 1, \ldots, m$

$$\boldsymbol{C}_x = \begin{bmatrix} c_x(\boldsymbol{p}_1, \boldsymbol{p}_1) & \cdots & c_x(\boldsymbol{p}_1, \boldsymbol{p}_m) \\ \cdots & & \cdots \\ c_x(\boldsymbol{p}_m, \boldsymbol{p}_1) & \cdots & c_x(\boldsymbol{p}_m, \boldsymbol{p}_m) \end{bmatrix}$$

Step 2: Decompose the $m \times m$ non negative-definite positive covariance matrix $\boldsymbol{C}_x$ into a product of a lower triangular matrix $\boldsymbol{L}$ and an upper triangular matrix $\boldsymbol{U}$ by means of the Cholesky algorithm; that is $\boldsymbol{C}_x = \boldsymbol{LU}$, where $\boldsymbol{U} = \boldsymbol{L}^T$.

Step 3: Suppose that $\boldsymbol{v}$ is a vector of m independent standard Gaussian random variables, and define the vector $\boldsymbol{x} = \boldsymbol{Lv}$. The vector $\boldsymbol{x} = [X(\boldsymbol{p}_1), X(\boldsymbol{p}_2), \ldots, X(\boldsymbol{p}_m)]$ provides the simulated values of a zero mean S/TRF with covariance $E[(\boldsymbol{LV})(\boldsymbol{LV})^T] = \boldsymbol{L}E[\boldsymbol{VV}^T]\boldsymbol{U} = \boldsymbol{LIU} = \boldsymbol{C}_x$, as required.

Stochatic simulations will be useful in this work when comparing different estimation methods. In this context several realizations of a S/TRF are generated, leading to pictures similar to Fig. 2.1. The simulated values are interpreted as the "true" values taken by the S/TRF. A subset of the true values are selected to serve as the observed data, and based on those, the value of the random field is estimated (or, more exactly, "re-estimated") at the remaining grid points. Then the estimated values may be compared with the "true" values to yield estimation errors. This technique allows to compare the estimation errors between different estimation methods, and is very useful when developing novel approaches to the spatiotemporal estimation problem.

2.2. Defining the Physical Knowledge Bases Available in Spatiotemporal Mapping

When we talk about spatiotemporal mapping we are talking about two components: (a) an organized body of physical knowledge (ontological component), and (b) a distinctive methodology for processing that knowledge (epistemological component). In order to develop a sound and flexible methodology to process knowledge (the latter epistemological component), we need to first define rigorously what is the physical knowledge available for spatiotemporal mapping (the former ontological component). The methods in classical geostatistics (e.g., kriging estimators) were primarily developed to use hard data (i.e., sets of exact measurements) which does not include important sources of physical knowledge such as physical laws, empirical models, higher-order space/time moments, and uncertain information. Here we define physical knowledge in a way that includes these additional sources of information, and we will in subsequent chapters present the epistemological component to process this physical knowledge.

The total physical knowledge K available regarding a natural process may be assumed to consist of two kinds of knowledge: general knowledge G and specificatory (or case-specific) knowledge S, so that $K = G \cup S$. General knowledge G denotes the background knowledge and justified beliefs, and it may include laws of science, structured patterns and assumptions, statistical moments, analytic or synthetic statements, etc.. The knowledge is considered "general" in the sense that it is vague enough to characterize a large class of fields or situations. For example, the mean and covariance of the S/TRF $X(\boldsymbol{p})$ are considered part of G since they characterize several random fields. In geostatistics and statistical geography, a commonly available general knowledge base consists of the space/time statistical moments (means, covariances, etc.) of the S/TRF $X(\boldsymbol{p})$. Indeed, the moments are considered part of G since they characterize several random fields. This important knowledge base can be mathematically expressed as

$$G: \int d\boldsymbol{\chi}_{\text{map}}\, g_\alpha(\boldsymbol{\chi}_{\text{map}})\, f_G(\boldsymbol{\chi}_{\text{map}}) = \overline{g_\alpha}, \qquad \alpha = 1,..,N_\alpha, \tag{2.23}$$

where $\boldsymbol{\chi}_{\text{map}}^T = [\boldsymbol{\chi}_{\text{data}}^T\ \chi_k]$ includes all the data and estimation points, $f_G(\boldsymbol{\chi}_{\text{map}})$ is the corresponding probability density function (pdf), and the g_α are properly chosen functions so that their expectations $\overline{g_\alpha}$ provide the space/time statistical moments of interest (by convention $g_0 = 1$ and $\overline{g}_0 = 1$, i.e.. g_0 is the normalization constraint). It is worth noting that Eqs. (2.23) offer a flexible setup that allows to process various types of general knowledge. For example, knowledge of the statistical moment of order q of $X(\boldsymbol{p})$ at a point $\boldsymbol{p}_i$ is obtained by simply letting $g_i(\chi_i) = \chi_i^q$. The covariance and any other higher-order space/time moment are obtained using the appropriate g_α functions. Also, in stochastic hydrology and statistical turbulence a physical law expressed as a stochastic partial differential equation governing $X(\boldsymbol{p})$ is part of G and can be, also, incorporated in Eq. (2.23).

On the other hand, knowledge S is knowledge that applies to the specific case. For mapping purposes S will denote physical data $\boldsymbol{\chi}_{\text{data}}$ that were obtained at points $\boldsymbol{p}_i$ ($i = 1,...,m$) for the specific mapping situation at hand. For example, measurements of a contaminant concentration at a hazardous waste site are considered part of S for the specific site. The specificatory knowledge S associated with physical data at points $\boldsymbol{p}_i$ ($i = 1,...,m$) may be divided into two groups, as follows

$$S: \boldsymbol{\chi}_{\text{data}} = [\boldsymbol{\chi}_{\text{hard}}^T\ \boldsymbol{\chi}_{\text{soft}}^T]^T = [\chi_1 \cdots \chi_m]^T, \tag{2.24}$$

where:

(a) Hard data $\boldsymbol{\chi}_{\text{hard}} = [\chi_1 \ldots \chi_{m_h}]^T$ at points $\boldsymbol{p}_i$ $(i = 1, \ldots, m_h)$ are exact measurements of the natural process. In other words, hard data means that the random vector $\boldsymbol{x}_{\text{hard}} = [x_1 \ldots x_{m_h}]^T$ takes the measured values $\boldsymbol{\chi}_{\text{hard}}$ with probability one, i.e.

$$P[\boldsymbol{x}_{\text{hard}} = \boldsymbol{\chi}_{\text{hard}}] = 1 \tag{2-25}$$

(b) Soft data $\boldsymbol{\chi}_{\text{soft}} = [\chi_{m_h+1} \ldots \chi_m]^T$ at points $\boldsymbol{p}_i$ $(i = m_h + 1, \ldots, m)$ may include uncertain knowledge, experience, intuition, etc. about the random vector $\boldsymbol{x}_{\text{soft}} = [x_{m_h+1} \ldots x_m]^T$. In this work we consider soft data of the interval and probabilistic types. In particular, when the unmeasured exact values χ_i $(i = m_h + 1, \ldots, m)$ lie with probability one within known intervals of the form $I_i = [l_i, u_i]$, the specificatory knowledge is of the interval type, i.e.,

$$\{\chi_i \in I_i = [l_i, u_i],\ i = m_h + 1, \ldots, m\}. \tag{2.26}$$

If the available evidence allows us to derive the specificatory probability density function (pdf) $f_S(\boldsymbol{\xi})$ of the random vector $\boldsymbol{x}_{\text{soft}}$, the soft data is said to be of the probabilistic type. In this case, the uncertain knowledge about the unmeasured values $\boldsymbol{\chi}_{\text{soft}}$ is expressed as

$$\boldsymbol{\chi}_{\text{soft}}:\ P_S(\boldsymbol{x}_{\text{soft}} \le \boldsymbol{\xi}) = \int_{-\infty}^{\xi} d\boldsymbol{\chi}_{\text{soft}}\ f_S(\boldsymbol{\chi}_{\text{soft}}), \tag{2.27}$$

where the subscript S simply denotes that the knowledge base used to derive these probabilities was specificatory.

The ontological classification chosen distinguishes between general knowledge and specificatory knowledge. This distinction allows to incorporate important sources of

information at different stages and therefore adds an element of flexibility to the spatiotemporal mapping approach. For example physical laws or empirical models may be incorporated as general knowledge, while uncertain measurements are incorporated as specificatory knowledge.

Uncertain measurements (or soft data) is an important source of information, but traditional kriging methods lack the theoretical underpinnings and practical flexibility to account for that type of information. Similarly the stochastic simulation methods presented earlier are set to generate primarily values of hard data (i.e. exact values of the random field on the gridded domain), and it is necessary to extend their capability to generate soft data. This is considered next.

2.3. Stochastic Simulation of S/TRF in the Light of Uncertain Physical Knowledge

In order to test and compare different spatiotemporal mapping methods in the context of uncertain physical knowledge, it is useful to generate several realizations of the S/TRF $X(\boldsymbol{p})$ containing both hard and soft data. We will consider soft data of the probabilistic type, since interval type is a special case that is obtained from probabilistic type by setting the soft pdf to a uniform distribution. Whereas a stochastic simulation is traditionally a realization of exact values taken by the random field at the location of a (usually gridded) set of points (see Fig. 2.1), in the case of uncertain physical knowledge, a stochastic realization will consist of a set of exact values $\boldsymbol{\chi}_{\text{hard}}^{(\ell)} = [\chi_1 \ldots \chi_{m_h}]^{T,(\ell)}$ at points $\boldsymbol{p}_i$ $(i = 1,\ldots,m_h)$, and a realization for the soft pdf $f_S^{(\ell)}(\boldsymbol{\chi}_{\text{soft}})$ describing the values $\boldsymbol{\chi}_{\text{soft}} = [\chi_{m_h+1} \ldots \chi_m]^T$ at points $\boldsymbol{p}_i$ $(i = m_h+1,\ldots,m)$, where the superscript $\ell = 1,\ldots,L$ denotes the specific field-realization. Consider for example the configuration of Fig. 2.2 showing the location of 8 hard data points represented with triangles, and 80 soft data

points represented with circles: A stochastic simulation in this case consists of 8 hard data values and of a realization for the soft pdf $f_S^{(\ell)}(\boldsymbol{\chi}_{\text{soft}})$ for the 80 soft data points representing the uncertain information at those points.

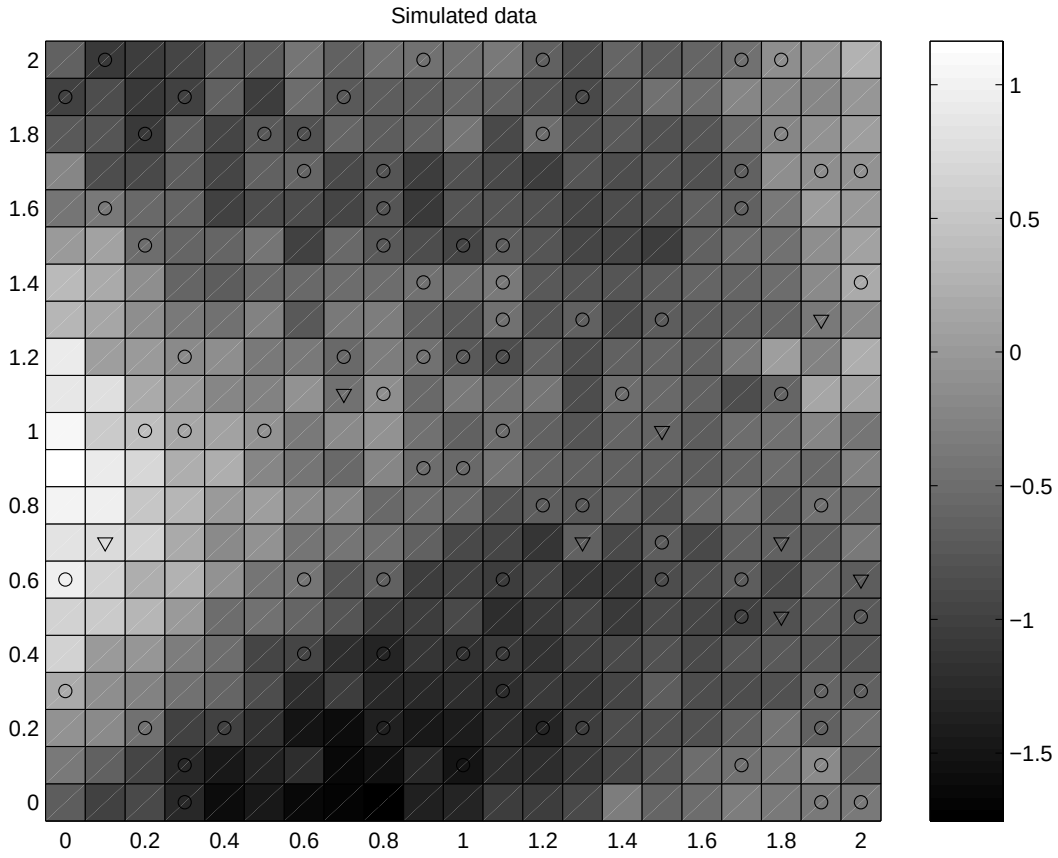

Figure 2.2: Configuration of 8 hard data and 80 soft data points distributed on a 21 by 21 grid. the hard data are shown with triangles while the soft data are shown with circles.

Without loss of generality, consider only one soft data point so that we are seeking the realizations of an univariate soft pdf, $f_S^{(\ell)}(\chi_{\text{soft}})$, $\ell = 1,\ldots,L$. Using the LU-simulation method we may generate realizations of exact values at both the hard data points, the $\boldsymbol{\chi}_{\text{hard}}^{(\ell)}$, and at the soft data point, denoted as $\chi_{\text{soft}}^{(\ell),*}$. The problem is that we do not want the values $\chi_{\text{soft}}^{(\ell),*}$, rather we want to generate realization of uncertain information about $\chi_{\text{soft}}^{(\ell),*}$, namely the soft pdfs $f_S^{(\ell)}(\chi_{\text{soft}})$. This is done by considering that the exact value $\chi_{\text{soft}}^{(\ell),*}$ is not known, instead what is known is an uncertain measurement $\psi_\ell = \chi_{\text{soft}}^{(\ell),*} - \upsilon_\ell$, where υ_ℓ is

a measurement error. The measurement error is a random variable v, and its distribution is named the *generator pdf* $f_v(\upsilon)$. Consider for example the generator pdf in Fig. 2.3, this generator pdf represents the distribution of measurement errors which pollute the (uncertain) measured values ψ_ℓ. Considering that only the measured values ψ_ℓ and the generator pdfs $f_v(\upsilon)$ are known (the $\chi_{\text{soft}}^{(\ell),*}$ are ignored), we may obtain the soft pdfs $f_S^{(\ell)}(\chi_{\text{soft}})$ as follow: Since the random variables x_{soft} and v are related by $v = x_{\text{soft}} - \psi$, where ψ is a known (fixed) value, we find $f_S^{(\ell)}(\chi_{\text{soft}}) = f_v(\chi_{\text{soft}} - \psi_\ell)$, $\ell = 1,\ldots,L$.

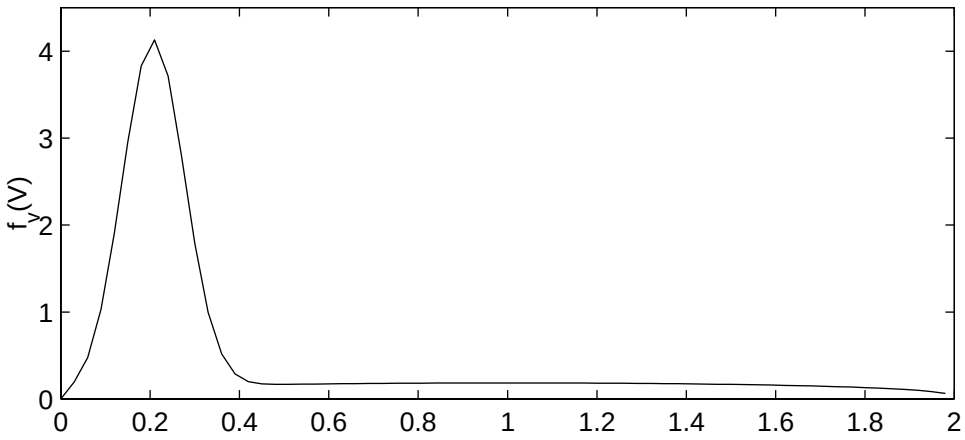

Figure 2.3: Probability distribution function $f_v(\vartheta)$ used to generate the soft pdf realizations

As a way of summary, the steps involved in generating stochastic simulations of a random field containing both hard and soft data of pdf type are outlined in Table 2.2. This technique generates realizations of the soft pdf at one soft data point. Note that this technique is easily extended to more than one soft data point by generating independent pdfs for each point, in which case the multi-point soft pdf is just written as the product of the (independent) pdfs for each soft data point.

TABLE 2.2: A technique for generating random field realizations

Step 1: Using LU decomposition $\boldsymbol{\chi}_{\text{hard}}^{(\ell)}$ and $\chi_{\text{hard}}^{(\ell)}$ values are generated at the hard and soft data points for each realization $\ell = 1,...,L$, assuming zero mean and a covariance function $c_x(\boldsymbol{p}, \boldsymbol{p}')$.

Step 2: The υ_ℓ-values ($\ell = 1,...,L$) are generated for the random variable v with pdf $f_v(\upsilon)$.

Step 3: For each realization, the value $\psi_\ell = \chi_{\text{soft}}^{(\ell)} - \upsilon_\ell$ is calculated; then, only the ψ_ℓ and $f_u(\upsilon)$ values are retained (the $\chi_{\text{soft}}^{(\ell)}$-value is not needed).

Step 4: Since the random variables x_{soft} and v are related by $v = x_{\text{soft}} - \psi$, where ψ is a known (fixed) value, we find $f_S^{(\ell)}(\chi_{\text{soft}}) = f_v(\chi_{\text{soft}} - \psi_\ell)$, $\ell = 1,...,L$.

For illustration purposes consider the 8 hard data and 80 soft data points of Fig. 2.2. Applying the technique described herein we generate one realization of the random field containing both hard data and soft data of the pdf type using the generator pdf shown with dotted line in Fig 2.4a. The histogram of values $\vartheta_i = (\chi_{s,i} - \psi_{s,i})$, i=1,...,80 generated using the generator pdf $f_v(\vartheta)$, shown in Fig. 2.4a in plain line, represents the measurement error polluting each of the soft data point measurement. The resulting soft pdf $f_{S,i}^{(\ell)}(\chi_{s,i})$ obtained at a typical soft data point is shown in Fig 2.4b.

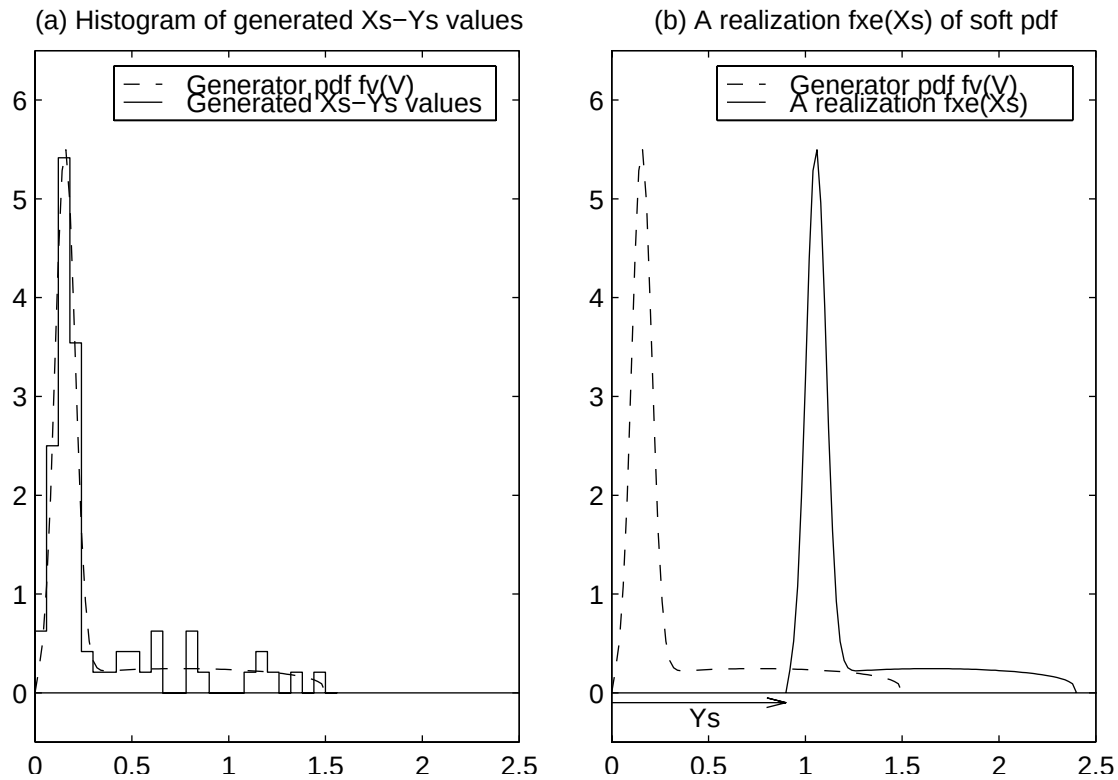

Figure 2.4.: (a) Histogram of values $\vartheta_i=(\chi_{s,i}-\psi_{s,i})$, i=1,...,80 generated using the generator pdf $f_v(\vartheta)$, (b) a realization of the soft pdf $f_{S,i}^{(\ell)}(\chi_{s,i})$ obtained for a given Ys.

2.4. A Review of Some Kriging Methods.

The kriging methods of classical geostatistics were primarily developed to account for exact measurements, and they lack the flexibility to process important sources of physical knowledge, such as uncertain information. A survey of all the flavors of kriging methods is not possible in the limited space of this work, therefore we will present some selected forms of kriging methods: namely the Simple Kriging (SK), the Indicator Kriging (IK), and the Simple Kriging with Measurement Error (SKME) approaches. As explained earlier, in spatiotemporal mapping we are interested in estimating the values taken by a S/TRF $X(\boldsymbol{p})$ at point $\boldsymbol{p}_k$, denoted as χ_k, given physical data at space/time points $\boldsymbol{p}_i$ (i=1,...m). We let $\boldsymbol{\chi}_{\text{hard}}=[\chi_1...\chi_{m_h}]^T$ denote the exact measurements (hard data) at points $\boldsymbol{p}_i$ $(i=1,...,m_h)$, and $\boldsymbol{\chi}_{\text{soft}}=[\chi_{m_h+1}...\chi_m]^T$ denote the soft information at points $\boldsymbol{p}_i$

($i = m_h + 1, \ldots, m$). The soft data may be of either interval type (Eq. 2-26), or probabilistic type (Eq. 2-27).

2.4.1. Simple Kriging

In the kriging methods the estimator is a linear combination of observed values. The estimator is constructed such that it is unbiased (i.e. the expected value of the estimator is equal to the expected value of the vlaue to estimate), and such that it minimizes the mean square error. Thus, kriging method are referred to as Best Linear Unbiased Estimation (BLUE) methods. The Simple Kriging method is presented first since it is the most simple (though it's name) of the kriging methods.

In the Simple Kriging method, the space/time random field $X(\boldsymbol{p})$ is assumed to have a known mean $m_x(\boldsymbol{p}) = \overline{X(\boldsymbol{p})}$, as well as a known covariance function $c_x(\boldsymbol{p}, \boldsymbol{p}')$, usually obtained from fitting to experimental data. The SK does not provide any direct mechanism for the incorporation of soft data, hence only the hard data $\boldsymbol{\chi}_{\text{hard}}$ may be considered in the estimation. As mentioned earlier, the estimator is given as a linear combination of the hard data, i.e.,

$$x_k^* = \lambda_o + \boldsymbol{\lambda}^T \boldsymbol{x}_{\text{hard}} \tag{2.28}$$

where x_k^* is the estimator random variable and λ_o and $\boldsymbol{\lambda}^T = [\lambda_1, \ldots, \lambda_{m_h}]$ are the weights of the linear combination. The unbiasness condition imposes that $\overline{x_k^*} = \overline{x_k}$. Taking the expected value of Eq. 2.28 and substituting $\overline{x_k^*} = \overline{x_k}$, we get $\lambda_o = \overline{x_k} - \boldsymbol{\lambda}^T \bar{\boldsymbol{x}}_{\text{hard}}$. Incorporating the expression for λ_o in Eq. 2.28, and defining the terms $m_k := \overline{x_k}$ and $\boldsymbol{m}_{\text{hard}} = [m_1, \ldots, m_{m_h}] := \bar{\boldsymbol{x}}_{\text{hard}}$, we can rewrite the (now unbiased) estimator as

$$x_k^* - m_k = \boldsymbol{\lambda}^T (\boldsymbol{x}_{\text{hard}} - \boldsymbol{m}_{\text{hard}}) \tag{2.29}$$

The estimation error random variable is defined as

$$e_k = x_k - x_k^* \tag{2.30}$$

Since the estimator is unbiased, we have $\bar{e}_k = 0$. The variance of the estimation error (also called the mean square error) is equal to $\sigma_e^2 = \overline{(e_k - \bar{e}_k)^2} = \overline{e_k^2}$. Substituting for the expressions of Eq. 2.30 and 2.29, and after some manipulations, we get

$$\sigma_e^2 = C_{k,k} - 2\boldsymbol{C}_{k,h}\boldsymbol{\lambda} + \boldsymbol{\lambda}^T\boldsymbol{C}_{h,h}\boldsymbol{\lambda} \tag{2.31}$$

where $C_{k,k} = c_x(\boldsymbol{p}_k, \boldsymbol{p}_k)$, $\boldsymbol{C}_{k,h} = [c_x(\boldsymbol{p}_k, \boldsymbol{p}_1) \quad \ldots \quad c_x(\boldsymbol{p}_k, \boldsymbol{p}_{m_h})]$ and

$$\boldsymbol{C}_{h,h} = \begin{bmatrix} c_x(\boldsymbol{p}_1, \boldsymbol{p}_1) & \ldots & c_x(\boldsymbol{p}_1, \boldsymbol{p}_{m_h}) \\ \ldots & & \ldots \\ c_x(\boldsymbol{p}_{m_h}, \boldsymbol{p}_1) & \ldots & c_x(\boldsymbol{p}_{m_h}, \boldsymbol{p}_{m_h}) \end{bmatrix}.$$

The mean square error is minimized by solving the system of linear equations

$$\frac{\partial \sigma_e^2}{\partial \lambda_i} = 0 \ , \ i = 1, \ldots, m_h \tag{2.32}$$

Substituting for the expression of Eq. 2.31 and taking the derivative we can write the system of linear equations as $\boldsymbol{C}_{h,h}\boldsymbol{\lambda} = \boldsymbol{C}_{k,h}$, which yield the expression for the weights coefficients $\boldsymbol{\lambda} = \boldsymbol{C}_{h,h}^{-1}\boldsymbol{C}_{h,k}$, or $\boldsymbol{\lambda}^T = \boldsymbol{C}_{k,h}\boldsymbol{C}_{h,h}^{-1}$. Substituting in Eq. 2.29 and 2.31 we get the celebrated SK estimator $x_k^* = \boldsymbol{C}_{k,h}\,\boldsymbol{C}_{h,h}^{-1}\,\boldsymbol{x}_{\text{hard}}$, which leads to the estimated value

$$\chi_k^* = \boldsymbol{C}_{k,h}\,\boldsymbol{C}_{h,h}^{-1}\,\boldsymbol{\chi}_{\text{hard}} \tag{2.33}$$

and the variance of x_k^*, or error variance, given by

$$\sigma_e^2 = C_{k,k} - \boldsymbol{C}_{k,h}\,\boldsymbol{C}_{h,h}^{-1}\,\boldsymbol{C}_{h,k} \tag{2.34}$$

One aspect of interest to us in examining this derivation is that the kriging methods do not identify between the ontological component (what is the knowledge base used?) and epistemological component (how is the knowledge base processed?). Therefore SK lacks any formal mechanism to extend the method beyond it's original capabilities. Examining what is the ontological component, we find that the general knowledge used includes the mean and covariance of the random field, and the specificatory knowledge consist of only hard data. Thus Simple Kriging cannot take advantage of additional information, such as physical laws or soft data. Effort have been made to nonetheless incorporate information of different kind in a kriging method, and one such attempt is the Indicator Kriging method, presented next.

2.4.2. Indicator Kriging (IK)

Indicator Kriging (IK) allows to estimate the cumulative probability distribution (cdf) $F(\chi_k)$ of the random variable to estimate x_k (A. Journel; 1989), given data information (m) at m data points. The cdf $F(\chi_k)$ is defined as the following probability

$$F(\chi_k) = P(X_k < \chi_k | (m)) \tag{2.35}$$

Introducing the indicator value $\mathrm{I}(\chi, x_i)$ for the random variable x_i , i=1,..,m, at threshold value χ:

$$I(\chi, x_i) = \begin{cases} 1, & if\ \chi \le x_i \\ 0, & if\ \chi > x_i \end{cases}, \tag{2.36}$$

the cdf for x_k is expressed as (Journel; 1989)

$$F(\chi_k) = [I(\chi, x_k)]^* = \sum_{i=1}^{m} \lambda_i(\chi)\ I(\chi, x_i). \tag{2.37}$$

The m weights $\lambda_i(\chi)$ are given by the ordinary kriging system of equations:

$$[\lambda_1,..,\lambda_m,\mu] = [C_I(\chi)_{k1},..,C_I(\chi)_{km},1] \begin{bmatrix} C_I(\chi)_{1,1} & .. & C_I(\chi)_{1,m} & 1 \\ .. & .. & .. & .. \\ C_I(\chi)_{m,1} & .. & C_I(\chi)_{m,m} & 1 \\ 1 & .. & 1 & 0 \end{bmatrix}^{-1} \tag{2.38}$$

where $C_I(\chi)$is the covariance matrix for the vector of indicator random values, with elements given by

$$C_I(\chi)_{i,j} = \overline{I(\chi, x_i) I(\chi, x_j)} - \overline{I(\chi, x_i)}\ \overline{I(\chi, x_j)},\ \text{i,j=k,1,..,m} \tag{2.39}$$

Under the assumption that the random variables x_i, i=k,1,..,m is multivariate normal with covariance matrix **C**, the element $C_I(\chi)_{i,j}$ of the indicator covariance is directly calculated using bivariate gaussian distribution for x_i and x_j.

With the cumulative distribution calculated using the Indicator Kriging method we then calculate an estimate for the unknown value x_k. The estimate that we calculate from the cumulative distribution is the median of the distribution, which we may obtain in the following way:

- Select a set of increasing values for χ, say χ_n, n=1,..,N
- For each χ_n calculate the probabilities F(χ_n) for the random variable x_k using Indicator Kriging

- Select the maximum value χ_n with a probability below 0.5
- Interpolate linearly between the selected χ_n and χ_{n+1} in order to get the value corresponding to a probability of 0.5. This gives the median of the distribution which we use as an estimate for the unknown value x_k.

It is noted that the nature of the data at the m data points was left unspecified. The specificatory information at the data points is coded by means of the indicator function (Eq. 2.36). Examining Eq. 2.36 we note that it is possible to code soft data of the interval type, as long as the lower and upper bounds of the soft interval coincide with the chosen values for the thresholds. Hard data may of course be accounted for as well, though this is done at the expense of some loss of information. Hence, the IK method suffers from a lack of accuracy when predicting the estimated value, but it provides nonetheless a candidate with which to compare novel and more accurate methods, as will be presented latter in this work.

2.4.3. Simple Kriging with Measurement Error

A modification to the Simple Kriging method that was developed in this work and found to perform surprisingly well within the realm of its limitation was dubbed the Simple Kriging with Measurement Error (SKME) method. This method is presented in two parts; first a kriging approach accounting for some measurement error is presented, then a formulation is proposed to write hard and soft data in terms of measurement errors. This method is still of course bound to the usual limitations of the kriging family, but it serves to enlighten our desire to move forward to the next chapter, where a novel approach to spatiotemporal mapping is presented.

Let us define a set of exact measurements $\boldsymbol{\chi}_{\text{hard}} = [\chi_1 \ldots \chi_{m_h}]^T$ at points $\boldsymbol{p}_i$ ($i = 1,\ldots,m_h$), and a set of uncertain measurements $\boldsymbol{\psi}_{\text{soft}} = [\psi_{m_h+1} \ldots \psi_m]^T$ at points $\boldsymbol{p}_i$ ($i = m_h+1,\ldots,m$). The uncertain measurements are "polluted" by the measurement errors $\boldsymbol{\upsilon} = [\upsilon_{m_h+1} \ldots \upsilon_m]^T$, i.e.

$$\boldsymbol{\psi}_{\text{soft}} = \boldsymbol{\chi}^*_{\text{soft}} - \boldsymbol{\upsilon} \tag{2.40}$$

where $\boldsymbol{\chi}^*_{\text{soft}}$ are exact but unknown values at points $\boldsymbol{p}_i$ ($i = m_h+1,\ldots,m$). We assume that the mean and covariance of the random field with values $\boldsymbol{\chi}_{\text{hard}}$ and $\boldsymbol{\chi}^*_{\text{soft}}$ has a known mean $m_x(\boldsymbol{p})$ and covariance function $c_x(\boldsymbol{p},\boldsymbol{p}')$, and that the measurement errors are zero mean, independent random variables and that their mean and variance is known, i.e. we assume that it is true that

$$\begin{cases} \overline{v_i} = 0, & i = m_h+1,\ldots,m \\ \overline{v_i v_j} = 0, & i = m_h+1,\ldots,m;\ j = m_h+1,\ldots,m \\ \overline{v_i^2} = \sigma_v^2, & i = m_h+1,\ldots,m \\ C_{x_i,v_j} = \overline{(x_i - \overline{x}_i)(v_j - \overline{v}_j)} = 0, & i = m_h+1,\ldots,m;\ j = m_h+1,\ldots,m \end{cases} \tag{2.41}$$

Along the lines of the SK method, we write the estimator as a linear combination of the exact and uncertain measurements

$$x^*_k = \lambda_o + \boldsymbol{\lambda}^{\mathrm{T}}_{\text{hard}} \boldsymbol{x}_{\text{hard}} + \boldsymbol{\lambda}^{\mathrm{T}}_{\text{soft}} \boldsymbol{y}_{\text{soft}} \tag{2.42}$$

Assuming without loss of generality that the mean of the random field is zero, applying the unbiasness condition and minimizing the mean square error in a way similar to that presented for the Simple Kriging method, we obtain without any complications the estimated value

$$\chi_k^* = \boldsymbol{C}_{k,hs}\,\boldsymbol{C}_{hs,hs(\sigma_v)}^{-1}[\boldsymbol{\chi}_{\text{hard}}^T\ \boldsymbol{\psi}_{\text{soft}}]^T \tag{2.43}$$

with corresponding error variance

$$\sigma_e^2 = C_{k,k} - \boldsymbol{C}_{k,hs}\,\boldsymbol{C}_{hs,hs(\sigma_v)}^{-1}\,\boldsymbol{C}_{hs,k} \tag{2.44}$$

where $\boldsymbol{C}_{hs,hs(\sigma_v)}$ is obtained by adding σ_v^2 to the diagonal elements $i = m_h + 1,\ldots,m$ (corresponding to uncertain measurements) of the covariance matrix

$$\boldsymbol{C}_{hs,hs} = \begin{bmatrix} c_x(\boldsymbol{p}_1,\boldsymbol{p}_1) & \ldots & c_x(\boldsymbol{p}_1,\boldsymbol{p}_m) \\ \ldots & & \ldots \\ c_x(\boldsymbol{p}_m,\boldsymbol{p}_1) & \ldots & c_x(\boldsymbol{p}_m,\boldsymbol{p}_m) \end{bmatrix}.$$

Now that we have defined a kriging approach accounting for uncertain data by means of a measurement error, let us see how soft data of probabilistic type may be modelled using the concept of measurement error. At points $i = m_h + 1,\ldots,m$ we define the uncertain measurement as the mean with respect to soft pdfs at the soft data point, i.e.

$$\boldsymbol{\psi}_{\text{soft}} = \overline{\boldsymbol{x}_S} \tag{2.45}$$

where $\overline{\boldsymbol{x}_S}$ has components $\overline{x_{S,i}} = \int d\chi_{\text{i}}\,\chi_{\text{i}}\,f_S(\chi_{\text{i}}),\ \ i = m_h + 1,\ldots,m,$ and we let the measurement error random variables be $v_i = x_i - \psi_i,\ \ i = m_h + 1,\ldots,m.$ Since the ψ_i are known, it follow immediately that the measurement errors have zero mean and a variance of

$$\sigma_{v,i}^2 = \int d\chi_{\text{i}}\,(\chi_{\text{i}} - \overline{x_{S,i}})^2\,f_S(\chi_{\text{i}}),\ \ i = m_h + 1,\ldots,m \tag{2.46}$$

If the assumptions of 3.41 hold, the estimated value and its variance are given by Eq. 3.43 and 3.44, with $\boldsymbol{\psi}_{\text{soft}}$ given by Eq. 3.45 and the measurement error variances given by Eq. 3.46.

In the case of soft data of the probabilistic data, the expressions simplify further. Let $\boldsymbol{l} = [l_{m_h+1} \ldots l_m]^T$ and $\boldsymbol{u} = [u_{m_h+1} \ldots u_m]^T$ be the lower and upper bounds of the soft data at points $\boldsymbol{p}_i,\ i = m_h + 1, \ldots, m.$ Then the uncertain measurements are $\boldsymbol{\psi}_{\text{soft}} = (\boldsymbol{u} + \boldsymbol{l})/2,$ and the measurement error variances are given by $\sigma_{v,i}^2 = (u_i - l_i)^2/12,\ i = m_h + 1, \ldots, m.$