B

Bayesian Maximum Entropy

Junyu He[1,2] and George Christakos[3]
[1]Ocean Academy, Zhejiang University, Zhoushan, China
[2]Ocean College, Zhejiang University, Zhoushan, China
[3]Department of Geography, San Diego State University, San Diego, CA, USA

Definition

The Bayesian maximum entropy (BME) theory of spatiotemporal geostatistics is concerned with the modeling and estimation/mapping of natural attributes (such as physical variables, ecologic processes, health parameters, and social indicators varying in the chronotopologic domain) in a chronotopologic domain (both the space-time and the spacetime are considered here). BME has introduced significant advances in the field of spatiotemporal geostatistics including an improved knowledge-theoretic modeling framework under condition of in situ uncertainty that is free of the restrictive assumptions of classical geostatistics (normality, linearity etc.); it can integrate a variety of core-general and case-specific knowledge bases (including physical laws, scientific theories, theoretical models, empirical evidence, soft data, and auxiliary information).

Introduction

The introduction of *Bayesian maximum entropy* (*BME*; Christakos 1990) provided a new way to address some shortcomings of previous notions and techniques (e.g., purely spatial considerations, spatial statistical regression) by means of a more broadly designed epistemic framework for geostatistical analysis and mapping. For example, Kriging, a linear estimator, is a statistical error minimization technique that relies on the Gaussian assumption; it focuses on low-order moments, and it cannot process important types of uncertain information and core knowledge (Olea 1999). BME is a knowledge-based approach, which, based on a theoretical support of considerable generality, can integrate the content of different information sources to achieve a realistic representation of natural phenomena which generate predictions of improved accuracy (Christakos and Serre 2000; Lee et al. 2008). BME provides the foundations to integrate knowledge bases of various kinds, both site-specific bases (such as additional ground measurements, satellite observations, and empirical charts) and core bases (including physical laws, basic principles, and theoretical models) (Kolovos et al. 2012). This is a particularly attractive feature in contemporary data-driven analytical environments, because BME enables the generation and integration of valuable information from core knowledge bases that might be otherwise left unused (Kolovos et al. 2002). Most prominently, BME extends one's ability to use measured data beyond single-value observations (Christakos and Li 1998; Christakos 2000). Specifically with respect to uncertainty, BME discriminates between two types of data, namely, the hard data and the soft data. Data whose observations are considered to be exact in the scope of an analysis are termed hard data. In contrast, soft data represent observed data that might carry nontrivial uncertainty in the context of a study. For example, soft data can be the result of routine sources of uncertainty such as the inaccuracy of measuring devices, modeling uncertainties, and human error. Alternatively, soft data can be derived as a result of inadequate knowledge, revised assessments, intuition estimates, and a host of other similar mechanisms that may provide uncertain informative content (Kolovos et al. 2016). BME is also free of additional limitations and restrictions that burden other geostatistical methods; for example, the BME analysis is independent of the data distribution (e.g., non-Gaussian distributions can be automatically incorporated), whereas Kriging is unable to accurately handle heavy tailed data (Christakos et al. 2001). In addition, the BME theoretical foundation enables both univariate and multivariate

B. S. Daya Sagar et al. (eds.), *Encyclopedia of Mathematical Geosciences*, Encyclopedia of Earth Sciences Series,
https://doi.org/10.1007/978-3-030-26050-7_50-1

predictions (Hristopulos and Christakos 2001). Per the detailed method comparison in Christakos (2000), in the course of the decades since its introduction, BME has been providing equally or more accurate prediction patterns than Kriging (e.g., Lee et al. 2008; Bogaert et al. 2009; Lee et al. 2010; Christakos 2010; Li et al. 2013c; Gao et al. 2014; Shi et al. 2015a).

Since its inception, BME has demonstrated its robust ability for spatiotemporal data analysis and modeling through a wide spectrum of applications in Earth and environmental sciences, public health, ecology, remote sensing, energy, resources science, and real estate research. A sample of the relevant literature includes Christakos et al. (2004), Heywood et al. (2006), Law et al. (2006), Brus et al. (2008), Savelieva et al. (2010), Kolovos et al. (2010), Yu et al. (2011), Angulo et al. (2013), Li et al. (2013a, b), Shi et al. (2015b), Sun et al. (2015), Yang and Christakos (2015), Zagouras et al. (2015), Hu et al. (2016), Tang et al. (2016), Kou et al. (2016), Hayunga and Kolovos (2016), He et al. (2019a), Zhang and Yang (2019a, b), and Wei et al. (2020). The BME methodology will be briefly reviewed next.

BME Methodology

The Knowledge Bases (KB)

At the beginning of BME (and of any scientific inquiry, as a matter of fact), knowledge bases is the available *knowledge* about the phenomenon of interest. Broadly speaking, the knowledge considered in real-world applications comes from a variety of sources, including all kinds of valid information available at a given moment and obtained by the competent scientist using effectively a scientific procedure. BME distinguishes between two major *knowledge bases* (*KB*):

(i) *Core* or *general G*-KB, which is background knowledge and justified beliefs relative to the phenomenon of interest. The *G* may include physical laws, scientific theories and models, structured patterns and assumptions, analytic and synthetic statements, and previous experience with similar phenomena independent of any case-specific data.

(ii) *Site-specific* or *specificatory S*-KB, which is targeted knowledge about the specific situation that may include both external or demonstrative evidence (actual measurements, perceptual or data of sense, etc.) and internal or inductive evidence (inferring one thing from another thing, empirical propositions, expertise with similar situations).

The synthesis of the *G*-KB and the *S*-KB provides the *total* KB. More specifically, the *G*-KB is expressed in terms of a series of suitable functions g_α of the attribute realization $\boldsymbol{\chi}_{map}$ at the point set $\boldsymbol{p}_{map}$ of interest, i.e., $g_\alpha(\boldsymbol{\chi}_{map})$, $\alpha = 0, 1, \ldots, N_c$ (by convention $g_0 = 1$). The g_α-functions account for a variety of core knowledge sources including statistical moments, empirical relationships, and physical laws. There are two conditions that guide the choice of the g_α-functions: (*a*) their stochastic expectations $\overline{g}_a = \int d\boldsymbol{\chi}_{map}\, g_a\, f_G$ must be calculable at the point set $\boldsymbol{p}_{map}$ (which includes the sampled points and any unsampled points of interest), where f_G is a probability density function (PDF) constructed on the basis of the *G*-KB, and (*b*) the resulting equations must be solvable with respect to the unknown parameters of the g_α functions (these parameters are functions of $\boldsymbol{p}_{map}$). Examples of *G*-KB formulated in terms of the g_α-functions include (Christakos 2000; Wu et al. 2021) $\int d\boldsymbol{\chi}_{map}\, g_\alpha\, O_\alpha[f_G] = 0$, where the g_α have forms associated with algebraic or differential operators O_α with respect to f_G (which includes derivatives of the attribute *PDF* with respect to the coordinates).

For practical purposes, the following *S*-KB distinction is usually adopted by BME: (*a*) *hard* data $\boldsymbol{\chi}_h$, which is the *S*-KB obtained from real-time observation devices with the probability of the attribute value in each datum being 1 (which is why it is also called *certain* data), and (*b*) *soft* data $\boldsymbol{\chi}_s$, which is the *S*-KB that includes non-negligible uncertainty (imprecision or vagueness) and is stochastically expressed as $\int_I d\Phi(\boldsymbol{\chi}_s) f_S(\Phi) \in [0, 1]$, where f_S is a PDF constructed on the basis of the available *S*-KB and Φ and I are a specified empirical model (a wide variety of choices, depending on the real-world application) and interval of values, respectively (Wu et al. 2021).

Compared to the mainstream geostatistical approaches, BME not only follows a data-driven workflow but also considers general knowledge to describe a natural process or phenomena, including physical laws, basic principles, scientific theories, etc. In this sense, BME is a *theory-based data-driven* approach. Using this kind of knowledge may steer easier or more precise chronotopologic data analysis in the following cases: (*i*) an inadequate number of data might be available for conclusive analysis, and (*ii*) uncertainty in data can bias results and might even lead to flawed analyses if unacknowledged.

The Chronotopologic Domain

A *chronotopologic domain* is a term used to describe either a space-time domain (i.e., space and time are considered separately) or a spacetime domain (i.e., space and time are considered as an integrated whole). Chronotopologic domain

Bayesian Maximum Entropy, Table 1 Chronotopologic point sets to be considered in *BME*

Entire point set	Vector $\boldsymbol{p}_{map} = [\boldsymbol{p}_1 \cdots \boldsymbol{p}_n]^{\mathrm{T}}$, which includes the nodes of the chronotopologic mapping grid that covers the entire domain of interest
Hard data point subset	Vector $\boldsymbol{p}_h = [\boldsymbol{p}_{h1} \cdots \boldsymbol{p}_{hn_h}]^{\mathrm{T}} \subset \boldsymbol{p}_{map}$, which includes the points where exact data is available
Soft data point subset	Vector $\boldsymbol{p}_s = [\boldsymbol{p}_{s1} \cdots \boldsymbol{p}_{sn_s}]^{\mathrm{T}} \subset \boldsymbol{p}_{map}$, which includes the points where uncertain data is available
Data point subset	Vector $\boldsymbol{p}_d = [\boldsymbol{p}_{d1} \cdots \boldsymbol{p}_{dn_d}]^{\mathrm{T}} \subset \boldsymbol{p}_h \cup \boldsymbol{p}_s$, i.e., the union of $\boldsymbol{p}_h$ and $\boldsymbol{p}_s$.
Unsampled point subset	Vector $\boldsymbol{p}_k = [\boldsymbol{p}_{k1} \cdots \boldsymbol{p}_{kn_k}]^{\mathrm{T}}$, where no empirical evidence exists but attribute estimates (interpolations, extrapolations, predictions) are sought
Point subset link	Relationship holds between the set and the two main point subsets leading to the entire point set: $\boldsymbol{p}_{map} = \boldsymbol{p}_d \cup \boldsymbol{p}_k$.

points are denoted by the vector $\boldsymbol{p}$, which includes both spatial and temporal coordinates. In Table 1, a distinction is presented between sets of points $\boldsymbol{p}$ within the chronotopologic domain.

The Spatiotemporal Random Field Model

The *spatiotemporal random field* theory (*STRF*, Christakos 1991, 2017) studies natural attributes and phenomena that evolve in the chronotopologic continuum. An S/TRF $X(\boldsymbol{p})$ is a collection of complementary field realizations associated with the values of an attribute field at points of the chronotopologic domain, where the j^{th}-realization at $\boldsymbol{p}_{map}$ is written as $\boldsymbol{\chi}_{map}^{(j)} = \left[\chi_{d1}^{(j)} \cdots \chi_{dn_d}^{(j)} \chi_{k1}^{(j)} \cdots \chi_{kn_k}^{(j)}\right]^{\mathrm{T}} (j = 1, 2, \cdots)$. In light of the notation in Table 1, it is commonly written that $\boldsymbol{\chi}_{map} = [\boldsymbol{\chi}_d \; \boldsymbol{\chi}_k]$, where $\boldsymbol{\chi}_{data} = [\chi_{d1} \cdots \chi_{dn_d}]^{\mathrm{T}}$ denotes attribute measurements or observations at points $\boldsymbol{p}_i$ and $\boldsymbol{\chi}_k = [\chi_{k1} \cdots \chi_{kn_k}]^{\mathrm{T}}$ denotes the unknown attribute values at points $\boldsymbol{p}_k (k \neq i)$.

In stochastic terms, $X(\boldsymbol{p})$ is fully described by the infinite sequence of probability density functions (PDF) $f_X(\boldsymbol{\chi}_{map})$ covering all domain points. The mean S/TRF function $\overline{X}(\boldsymbol{p})$ represents trends and systematic structures that exist in a larger scale, the covariance function $c_X(\boldsymbol{p}, \Delta\boldsymbol{p}) = \overline{\left[X(\boldsymbol{p}) - \overline{X}(\boldsymbol{p})\right]\left[X(\boldsymbol{p} + \Delta\boldsymbol{p}) - \overline{X}(\boldsymbol{p} + \Delta\boldsymbol{p})\right]}$ describes chronotopologic dependencies between pairs of points $\boldsymbol{p}$ and $\boldsymbol{p} + \Delta\boldsymbol{p}$, and the trivariogram function $\zeta_X(\boldsymbol{p}, \Delta\boldsymbol{p}, \Delta\boldsymbol{p}') = \overline{\left[X(\boldsymbol{p}) - \overline{X(\boldsymbol{p})}\right]\left[X(\boldsymbol{p} + \Delta\boldsymbol{p}) - \overline{X(\boldsymbol{p} + \Delta\boldsymbol{p})}\right]\left[X(\boldsymbol{p} + \Delta\boldsymbol{p}') - \overline{X(\boldsymbol{p} + \Delta\boldsymbol{p}')}\right]}$ describes chronotopologic dependencies between triplets of points $\boldsymbol{p}$, $\boldsymbol{p} + \Delta\boldsymbol{p}$, and $\boldsymbol{p} + \Delta\boldsymbol{p}'$, where $\Delta\boldsymbol{p}$ and $\Delta\boldsymbol{p}'$ are chronotopologic lags.

The BME Workflow

The basic BME stages are outlined in Fig. 1. Table 2 describes the BME methodology in terms of the g_α $(a = 1, \ldots, N_c)$ and Φ functions and the PDF f_G and f_S.

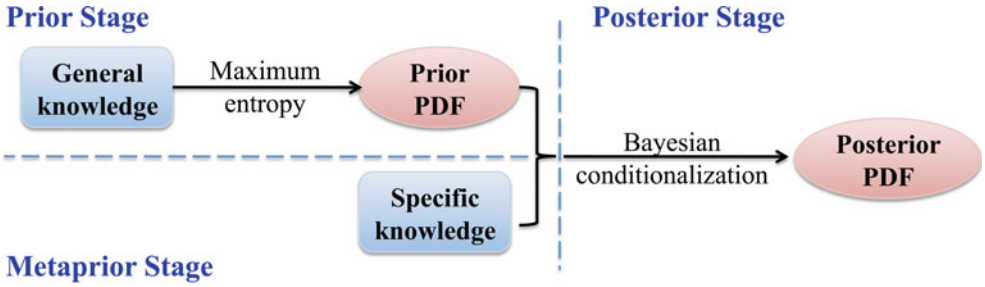
In more words, at the prior (*G-based*) stage, there is an inverse relation between prior information and prior probability: the more vague and general a theory is, the more alternatives it includes (and it is, hence, more probable) but also the less informative it is. In BME, maximization of information input serves as acquiring prior knowledge that can be filtrated, integrated, and optimized by specified rules and is a pivotal step to describe general attribute characteristics.

Thus, the prior PDF (*G*-based) f_G is defined by the calculated μ_α and the selected g_a, i.e., $f_G\left(\boldsymbol{\chi}_{map}\right) = e^{\sum_{\alpha=0}^{N_c} \mu_\alpha g_\alpha}$. The Lagrange multipliers μ_α $(\alpha = 1, \ldots, N_C)$ calculated in the prior stage by information maximization quantify the significance of each g_a-function, whereas the μ_0 is a normalization parameter. The $\boldsymbol{\mu}_a$ are themselves functions of $\boldsymbol{p}_{map}$. Ideally, the more detailed G is, the thinner f_G becomes by excluding irrelevant outcomes, hence leading to a more informative model. Examples of the $(g_\alpha, \overline{g_\alpha})$ pairs include theoretical models of the attribute mean, covariance, variogram, trivariogram, higher-order chronotopologic moments, and equations expressing physical laws.

At the metaprior (*S*-based) stage, the *S-KB* is collected and organized in appropriate quantitative forms that can be explicitly incorporated into the BME formalism. This includes data cleaning, transformations, and evaluation, identification, and discrimination into hard and soft data, as appropriate for the analysis. Commonly encountered soft data forms are listed in Table 3.

In particular, the interval case in Table 3 shows that interval data values χ_i lie within a known interval I_i with lower and upper limits l_i and u_i, respectively; the probabilistic 1 case displays probabilistic data for which the cumulative distribution functions (CDF) $F_S(\zeta_i)$ are known; the probabilistic 2 case refers to threshold soft data, where the probabilities ϑ_i for certain threshold values I_i are known, and the functional case displays soft data, where the values can be derived from a given function h.

The posterior (*K*-based) stage is a synthesis stage in which the posterior probability is related to the *G*-based probability by means of a conditional probability knowledge-processing rule that draws inferences consistent with the available knowledge body and scientific reasoning. The *physical Bayesian conditionalization* rule is applied into the prior PDF f_G to produce the posterior PDF f_K and bring the model



Bayesian Maximum Entropy, Fig. 1 Workflow of the BME analysis

Bayesian Maximum Entropy, Table 2 Main *BME* stages

Stage	Formulation
Prior (*G*-based)	$\max_{\mu_\alpha}\left[\overline{\text{Info}_G(\boldsymbol{x}_{map})} = -\overline{\log e^{\sum_{\alpha=0}^{N_c}\mu_\alpha g_\alpha}}\right]$ s.t. $\int d\boldsymbol{\chi}_{map}\; g_\alpha\; e^{\sum_{\alpha=0}^{N_c}\mu_\alpha g_\alpha} = \overline{g_\alpha}$
Metaprior (*S*-based)	$\boldsymbol{\chi}_d = \{\boldsymbol{\chi}_h, \boldsymbol{\chi}_s\}, \Phi, f_S$
Posterior (*K*-based)	$f_K(\boldsymbol{\chi}_k) = \frac{1}{\int d\Phi\; f_S(\Phi) f_G(\boldsymbol{\chi}_h,\;\Phi)} \int d\Phi f_S(\Phi) f_G\left(\boldsymbol{\chi}_{map}\right)$

Bayesian Maximum Entropy, Table 3 Examples of soft data formulations

Type of S-KB	*Formulation* $\boldsymbol{\chi}_s = \left[\chi_{s1}\cdots\chi_{sn_s}\right]^T\ (i = 1,\cdots,n_s)$
Interval	$\chi_{si} \in I_i = [l_i,\;\; u_i]$
Probabilistic 1	$P_S[\chi_i \le \zeta_i] = F_S(\zeta_i)$
Probabilistic 2	$P_S[\chi_i \in I_i] = \vartheta_i$
Functional	$P_S[h(\chi_i) \le \zeta_i] = F_S(\zeta_i,\;\; h)$

description closer to natural truth. According to the BME perspective, the *G-KB* introduces scientific consistency into the *S-KB*, and the *S-KB* introduces empirical conditioning into *G-KB* in light of the specific site considered. Interestingly, the above conclusion may be linked to the verifiability criterion used in sciences to constrain theoretical constructions by means of empirical evidence. With regard to material object claims, this criterion is known as "phenomenalism," and with regard to the theoretical claims of science, as "operationalism." Some examples of formulations corresponding to the interval, probabilistic, and functional soft data are presented in Table 4, where I, R, and $I(\zeta)$ are definition domains that depend on the soft data.

BME Chronotopologic Estimation and Uncertainty Assessment

The obtained BME posterior PDF $f_K(\boldsymbol{\chi}_k)$ at points $\boldsymbol{p}_k$ provides complete statistical description of the estimated (interpolated, extrapolated, predicted) attribute. In particular, the BME *mode* estimate $\widehat{\chi}_{k,mode}$ is such that

Bayesian Maximum Entropy, Table 4 Examples of posterior PDF formulations

Type of S-KB	*Formulation* $f_K(\boldsymbol{\chi}_k)$
Interval	$\frac{1}{\int_{I_S} d\boldsymbol{\chi}_s f_G(\boldsymbol{\chi}_d)} \int_{I_S} d\boldsymbol{\chi}_s f_G(\chi_k, \boldsymbol{\chi}_d)$
Probabilistic 1	$\frac{1}{\int_{I_S} dF_G(\boldsymbol{\chi}_s) f_G(\boldsymbol{\chi}_d)} \int_{I_S} dF_G(\boldsymbol{\chi}_s) f_G(\chi_k, \boldsymbol{\chi}_s)$
Functional	$\frac{1}{\int_R dF_G(\zeta,h) \int_{I(\zeta)} d\boldsymbol{\chi}_d f_G(\boldsymbol{\chi}_d)} \int_R dF_G(\zeta,h) \int_{I(\zeta)} d\boldsymbol{\chi}_s f_G(\chi_k, \boldsymbol{\chi}_d)$

$$f_K\left(\widehat{\chi}_{k,mode}\right) = \max_{\chi_k} f_K(\chi_k);$$

the BME *mean* estimate $\widehat{\chi}_{k,mean}$ is given by

$$\widehat{\chi}_{k,mean} = \overline{X(\boldsymbol{p}_k)|\boldsymbol{\chi}_d} = \int d\boldsymbol{\chi}_k f_K(\boldsymbol{\chi}_k)\boldsymbol{\chi}_k;$$

and the BME *median* estimate $\widehat{\chi}_{k,med}$ is defined as

$$\widehat{\chi}_{k,med} = F_G^{-1}(0.5,\;\; \boldsymbol{p}_k): \quad \int_{\boldsymbol{\chi}_k \le \widehat{\chi}_{k,med}} d\boldsymbol{\chi}_k f_K(\boldsymbol{\chi}_k) = \frac{1}{2}.$$

The determination of chronotopologic BME estimation uncertainty depends on the shape of the posterior PDF (symmetric *vs*. asymmetric, single maximum *vs*. multiple maxima, etc.; Wu et al. 2021). In particular, in the case of a symmetric attribute PDF, the estimation error variance is

$$e_X^2(\boldsymbol{p}_k) = \overline{\left[X(\boldsymbol{p}_k) - \widehat{X}(\boldsymbol{p}_k)\right]^2} = \int d\chi_k \left(\chi_k - \widehat{\chi}_{k,\text{mean}}\right)^2 f_K(\chi_k),$$

which is a widely used formula.

Interestingly, consider the special case where the following conditions are met concurrently: (*i*) the *G*–KB comprises only the attribute mean trend and covariance, that is, the first two stochastic moments of the attribute, and (*ii*) the *S*–KB contains only hard data. In this case, the BME equations reduce to Kriging, thus rendering Kriging a special case of chronotopologic BME estimation. In particular, the BME posterior PDF f_K is then fully defined by its mean and

variance, and these measures correspond exactly to the Kriging prediction and prediction error, respectively.

BME Features and Applications

BME Versatility

Due to *BME*'s versatility, its basic equations possess significant generalization power on the basis of theory and data. As a consequence, the following essential *BME* features are of considerable interest as follows (Christakos 1998, 2000):

(i) Well-known statistical interpolation techniques are derived as *special cases* of the chronotopologic BME theory. In the case, e.g., where the knowledge conditions $G = \{\overline{X}(\boldsymbol{p}), \quad c_X(\boldsymbol{p}, \Delta\boldsymbol{p})\}$ and $S = \boldsymbol{\chi}_h$ are met concurrently, the BME equations reduce to those of Kriging, thus rendering Kriging a special case of BME estimation (f_K is then determined by its mean and variance, which correspond to the Kriging estimation and estimation error variance, respectively).
(ii) *BME* is a *nonlinear* chronotopologic estimator, in general.
(iii) It incorporates *multi-sourced* knowledge bases (core and site-specific).
(iv) It can process *higher-order* chronotopologic statistics associated with *non-Gaussian* data distributions, like skewness and kurtosis functions.
(v) It has *extrapolation* (or *prediction*) abilities due to its incorporation of physical laws and other forms of core knowledge.
(vi) It handles *functional*, *vector,* and *multipoint* chronotopologic estimation.
(vii) It is *computationally* efficient, particularly in estimation cases where closed-form analytical expressions are available.

BME borrows strength from the Bayes rule and information theory, making it a versatile approach. On the one hand, the maximum entropy rule enables filtering out redundant knowledge and having useful knowledge left, which can also lead to more realistic chronotopologic estimations, and, on the other hand, the Bayes rule allows us to bring prior information (including not only statistical characteristic of the studied attribute but also physical laws, expert judgment, life experience, and skills) into considerations, which may save time/costs and lead to precise inference, even in the case that entire lack of data (Shoemaker et al. 1999; Christakos and Hristopulos 1998; Serre and Christakos 2002; Kolovos et al. 2002; McCarthy and Masters 2005; Kuhnert et al. 2010; Choy et al. 2009; Martin et al. 2012; Lang and Christakos 2019).

Another strength of the BME framework is the rigorous handling of soft data, i.e., uncertain data, such as historical records, secondary information, expert opinions, and beliefs. Therefore, BME can make maximal use of multi-sourced data and also provides a port to combine various mathematical models (Money et al. 2011; Hu et al. 2016; Li et al. 2013a; Reyes and Serre 2014; He and Christakos 2018). This can significantly broaden the BME application range. Sometimes, soft data can be created by simple statistical methods for making up missing data in a time series case (Bayat et al. 2013; Lee et al. 2008). In other cases, the observed data may be vague, so sourcing expert advice, empirical handling of similar situations, intuition, and professional knowledge can be used to construct soft data (Serre et al. 2003; Puangthongthub et al. 2007, Yu and Wang 2013). Notice that compared to estimation without soft data, many studies have shown that including soft data can generate more accurate estimations (Christakos et al. 2001; D'Or et al. 2001; Shi et al. 2015b; Hayunga and Kolovos 2016).

The Emergence of Stochastic Logic in Geostatistics

The vast majority of the phenomena encountered in nature obey physical laws, which, in the case of classical physics, can be described in terms amenable to efficient *material causation* of logic theory. This is why BME's structure allows it to formally incorporate physical laws in a systematic and internally consistent manner. Mathematical logic is used because of the underlying deductive reasoning (if the premises are true, the conclusions are always true); on the other hand, when statistical inference is used (e.g., in statistical regression), because of the underlying inductive reasoning, the premises may be true and the conclusions wrong. Accordingly, BME has been extended in a stochastic logic context (Christakos 2002).

Indeed, beyond the conditional probability $P_G[\boldsymbol{\chi}_k|\boldsymbol{\chi}_d]$ used in the posterior stage of the original BME approach, other useful knowledge updating mechanisms can be expressed as the probability of *logical conditionals*, like the implication probability $P_G[\boldsymbol{\chi}_d \rightarrow \boldsymbol{\chi}_k]$, and the equivalence probability $P_G[\boldsymbol{\chi}_d \leftrightarrow \boldsymbol{\chi}_k]$, which are probabilistic expressions of the platitude that valid arguments are necessarily truth preserving. Otherwise said, these stochastic logic conditionals reveal that the relation of probability and deducibility involves the body of knowledge one actually has or the premises one actually accepts. Despite their conceptual differences, the above posterior probabilities are formally linked. For example, the standard (statistical) conditional probability is directly related to the implication probability as.

$[P_G[\boldsymbol{\chi}_k|\boldsymbol{\chi}_d] - 1]P_G[\boldsymbol{\chi}_d] = P_G[\boldsymbol{\chi}_d \rightarrow \boldsymbol{\chi}_k] - 1$. Concluding, stochastic logic becomes a promising research avenue, when one realizes that the Bayes rule and the standard (statistical)

conditional probability is not the only way to derive a posterior mapping probability in geostatistics.

BME in Software

There are several software libraries that can implement the BME theoretical developments, including the following: (*i*) BME library (BMElib) (Christakos et al. 2002), (*ii*) the Spatiotemporal Epistemic Knowledge Synthesis Graphical User Interface (SEKS-GUI) (Yu et al. 2007), and (*iii*) the Space-Time Analysis Rendering with BME (STAR-BME) (Yu et al. 2012, 2016). The first two tools are Matlab-based compilations, while the third one was developed on Python. Each tool features a guided analysis through a sequence of screens across the different analysis tasks that start at importing data and extend all the way to visualization or results and exporting the analysis output.

Applications of BME in the Real World

Let us recall the basic function of BME approach, i.e., use known knowledge to explore the unknown natural process in space-time domain. Therefore, BME can be applied in a variety of disciplines (He and Kolovos 2018), as follows:

- The chronotopologic distribution of $PM_{2.5}$, ozone, NO_X concentrations, etc., was characterized and mapped for studying the human expose purpose by using BME or its combinations with land use regression models in the field of atmosphere (Christakos and Kolovos 1999; Bogaert et al. 2009; Nol et al. 2010; Yu and Wang 2010; Beckerman et al. 2013; Reyes and Serre 2014; Adam-Poupart et al. 2014; He and Christakos 2018; He et al. 2019a).
- For the lithosphere, BME was employed to study the heavy metal concentrations, electrical conductivity, moisture, organic matter in soil, and soil categories (D'Or and Bogaert 2003; Douaik et al. 2005; Brus et al. 2008; Modis et al. 2013; Gao et al. 2014; Sun et al. 2015).
- For the ecosphere, BME not only can study the natural beings (such as mercury concentration in fish tissue, young cork oak plantation, leaf area index, pedestrian mortality rates, etc., Money et al. 2011; Sedda et al. 2011; Li et al. 2013a; Fox et al. 2015) but also can depict the transmission and S/ST characteristics of a disease (such as black death, influenza, hand-foot-mouth, dengue fever, syphilis, hemorrhagic fever with renal syndrome, etc., Christakos et al. 2005; Law et al. 2006; Choi et al. 2008; Cao et al. 2010; Angulo et al. 2013; Yu et al. 2014; He et al. 2017, 2018, 2019b).
- Regarding the hydrosphere, groundwater contamination, level, and hydraulic conductivity were investigated by BME (Vyas et al. 2004; Messier et al. 2012, 2014, 2015; Yu and Chu 2010). Similarly, the inland water (such as the fecal contamination) was also characterized through BME (Coulliette et al. 2009; Money et al. 2009). Besides, the evolution of precipitation was studied in different areas with BME (Mahbub et al. 2010; Bayat et al. 2014; Zhang et al. 2016). With the largest area on the Earth, the oceans play a vital role in our planet's environmental cycle. Although the BME studies on oceans are not as much as on the other part of the Earth, but it has great potential to show its versatile ability using the satellite remote sensing data on data fusion, spatiotemporal mapping, etc., see Li et al. (2013b), Tang et al. (2015), Shi et al. (2015b), and He et al. (2020a, 2020b).

Conclusions

BME provides a set of powerful tools supported by a rigorous theoretical framework, which are used to represent, estimate, and map natural attributes at unsampled locations under conditions of in situ uncertainty. BME displays a number of compelling characteristics on both theoretical and practical grounds, which make it stand out among mainstream geostatistics methods. On the theoretical front, BME makes no restrictive assumptions when it comes to properties like normality, linearity, and content-independency and is capable of ingesting information beyond the commonly used low-order moments and hard data. On the practical front, BME is built with the ability to integrate a variety of knowledge bases, including core knowledge bases (such as physical laws, theoretical models, and expert knowledge) and site-specific knowledge bases (such as empirical evidence, hard data, soft data, and secondary information).

Acknowledgement We would like to thank Dr. Alexander Kolovos for his valuable input to this work.

Bibliography

Adam-Poupart A, Brand A, Fournier M, Jerrett M, Smargiassi A (2014) Spatiotemporal modeling of ozone levels in Quebec (Canada): a comparison of kriging, Land-Use Regression (LUR), and combined Bayesian maximum entropy-LUR approaches. Environ Health Perspect 122:970–976

Angulo J, Yu H-L, Langousis A, Kolovos A, Wang J, Madrid AE, Christakos G (2013) Spatiotemporal infectious disease modeling: a BME-SIR approach. PLoS One 8:e72168

Bayat B, Zahraie B, Taghavi F, Nasseri M (2013) Evaluation of spatial and spatiotemporal estimation methods in simulation of precipitation variability patterns. Theor Appl Climatol 113:429–444

Bayat B, Nasseri M, Naser G (2014) Improving Bayesian maximum entropy and ordinary Kriging methods for estimating precipitations

in a large watershed: a new cluster-based approach. Can J Earth Sci 51:43–55
Beckerman BS, Jerrett M, Serre M, Martin RV, Lee SJ, van Donkelaar A, Ross Z, Su J, Burnett RT (2013) A hybrid approach to estimating national scale spatiotemporal variability of PM2.5 in the contiguous United States. Environ Sci Technol 47:7233–7241
Bogaert P, Christakos G, Jerrett M, Yu H-L (2009) Spatiotemporal modelling of ozone distribution in the State of California. Atmos Environ 43:2471–2480
Brus D, Bogaert P, Heuvelink G (2008) Bayesian maximum entropy prediction of soil categories using a traditional soil map as soft information. Eur J Soil Sci 59:166–177
Cao C, Xu M, Chang C, Xue Y, Zhong S, Fang L, Cao W, Zhang H, Gao M, He Q, Zhao J, Chen W, Zheng S, Li X (2010) Risk analysis for the highly pathogenic avian influenza in Mainland China using meta-modeling. Chin Sci Bull 55:4168–4178
Choi KM, Yu HL, Wilson ML (2008) Spatiotemporal statistical analysis of influenza mortality risk in the State of California during the period 1997–2001. Stoch Env Res Risk A 22:S15–S25
Choy SL, O'Leary R, Mengersen K (2009) Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology 90:265–277
Christakos G (1990) A Bayesian/maximum-entropy view to the spatial estimation problem. Math Geol 22:763–777
Christakos G (1991) On certain classes of spatiotemporal random-fields with applications to space-time data-processing. IEEE Trans Syst Man Cybern 21:861–875
Christakos G (1998) Spatiotemporal information systems in soil and environmental sciences. Geoderma 85(2–3):141–179
Christakos G (2000) Modern spatiotemporal geostatistics. Oxford University Press, New York
Christakos G (2002) On a deductive logic-based spatiotemporal random field theory. Theory Probab Math Stat 66:54–65
Christakos G (2010) Integrative problem-solving in a time of decadence. Springer, New York
Christakos G (2017) Spatiotemporal random fields: theory and applications. Elsevier, Netherlands: Amsterdam
Christakos G, Hristopulos DT (1998) Spatiotemporal environmental health modelling. Kluwer Academic Publishers, Boston
Christakos G, Kolovos A (1999) A study of the spatiotemporal health impacts of ozone exposure. J Expo Anal Environ Epidemiol 9: 322–335
Christakos G, Li X (1998) Bayesian maximum entropy analysis and mapping: a farewell to kriging estimators? Math Geol 30:435–462
Christakos G, Serre ML (2000) BME analysis of spatiotemporal particulate matter distributions in North Carolina. Atmos Environ 34: 3393–3406
Christakos G, Serre ML, Kovitz JL (2001) BME representation of particulate matter distributions in the state of California on the basis of uncertain measurements. J Geophys Res Atmos 106:9717–9731
Christakos G, Bogaert P, Serre M (2002) Temporal GIS: advanced functions for field-based applications. Springer Science & Business Media, German: Berlin
Christakos G, Kolovos A, Serre ML, Vukovich F (2004) Total ozone mapping by integrating databases from remote sensing instruments and empirical models. IEEE Trans Geosci Remote Sens 42: 991–1008
Christakos G, Olea RA, Serre ML, Wang LL, Yu HL (2005) Interdisciplinary public health reasoning and epidemic modelling: the case of black death. Springer, Berlin/Heidelberg
Coulliette AD, Money ES, Serre ML, Noble RT (2009) Space/time analysis of fecal pollution and rainfall in an Eastern North Carolina Estuary. Environ Sci Technol 43:3728–3735
D'Or D, Bogaert P (2003) Continuous-valued map reconstruction with the Bayesian maximum entropy. Geoderma 112:169–178
D'Or D, Bogaert P, Christakos G (2001) Application of the BME approach to soil texture mapping. Stoch Env Res Risk A 15:87–100
Douaik A, Van Meirvenne M, Toth T (2005) Soil salinity mapping using spatio-temporal kriging and Bayesian maximum entropy with interval soft data. Geoderma 128:234–248
Fox L, Serre ML, Lippmann SJ, Rodriguez DA, Bangdiwala SI et al (2015) Spatiotemporal approaches to analyzing pedestrian fatalities: the case of Cali, Colombia. Traffic Inj Prev 16:571–577
Gao S, Zhu Z, Liu S, Jin R, Yang G, Tan L (2014) Estimating the spatial distribution of soil moisture based on Bayesian maximum entropy method with auxiliary data from remote sensing. Int J Appl Earth Obs Geoinf 32:54–66
Hayunga DK, Kolovos A (2016) Geostatistical space–time mapping of house prices using Bayesian maximum entropy. Int J Geogr Inf Sci. https://doi.org/10.1080/13658816.2016.1165820
He J, Christakos G (2018) Space-time PM2.5 mapping in the severe haze region of Jing-Jin-Ji (China) using a synthetic approach. Environ Pollut 240:319–329
He J, Kolovos A (2018) Bayesian maximum entropy approach and its application: a review. Stoch Env Res Risk A 32(4):859–877
He J, Christakos G, Zhang W, Wang Y (2017) A space-time study of hemorrhagic fever with Renal Syndrome (HFRS) and its climatic associations in Heilongjiang Province, China. Front Appl Math Stat 3:16
He J, Christakos G, Wu J, Cazelles B, Qian Q, Mu D, Wang Y, Yin W, Zhang W (2018) Spatiotemporal variation of the association between climate dynamics and HFRS outbreaks in Eastearn China during 2005–2016 and its geographic determinants. PLoS Negl Trop Dis 12(6):e0006554
He J, Christakos G, Jankowski P (2019a) Comparative performance of the LUR, ANN and BME techniques in the multi-scale spatiotemporal mapping of PM2.5 concentrations in North China. IEEE J Sel Topics Appl Earth Observ Remote Sens 12(6):1734–1747
He J, Christakos G, Wu J, Jankowski P, Langousis A, Wang Y, Yin W, Zhang W (2019b) Probabilistic logic analysis of the highly heterogeneous spatiotemporal HFRS incidence distribution in Heilongjiang Province (China) during 2005–2013. PLoS Negl Trop Dis 13(1): e0007091
He J, Chen Y, Wu J, Stow D, Christakos G (2020a) Space-time chlorophyll-a retrieval in optically complex waters that accounts for remote sensing and modeling uncertainties and improves remote estimation accuracy. Water Res 171:1–17
He M, He J, Christakos G (2020b) Space-time mapping of sea surface salinity in western pacific ocean using contingogram modelling. Stoch Env Res Risk A 34:355–368
Heywood B, Brierley A, Gull S (2006) A quantified Bayesian maximum entropy estimate of Antarctic krill abundance across the Scotia Sea and in small-scale management units from the CCAMLR-2000 survey. Ccamlr Sci 13:97–116
Hristopulos DT, Christakos G (2001) Practical calculation of non-Gaussian multivariate moments in spatiotemporal Bayesian maximum entropy analysis. Math Geol 33(5):543–568
Hu JG, Zhou J, Zhou GM, Luo YQ, Xu XJ, Li PH, Liang JY (2016) Improving estimations of spatial distribution of soil respiration using the Bayesian maximum entropy algorithm and soil temperature as auxiliary data. PLoS One 11(1):e0146589
Kolovos A, Christakos G, Serre ML, Miller CT (2002) Computational BME solution of a stochastic advection-reaction equation in the light of site-specific information. Water Resour Res 38(12):1318–1334

Kolovos A, Skupin A, Jerrett M, Christakos G (2010) Multi-perspective analysis and spatiotemporal mapping of air pollution monitoring data. Environ Sci Technol 44:6738–6744

Kolovos A, Angulo JM, Modis K, Papantonopoulos G, Wang J-F, Christakos G (2012) Model-driven development of covariances for spatiotemporal environmental health assessment. Environ Monit Assess. https://doi.org/10.1007/s10661-012-2593-1

Kolovos A, Smith LM, Schwab-McCoy A, Gengler S, Yu H-L (2016) Emerging patterns in multi-sourced data modeling uncertainty. Spat Stat 18A:300–317. https://doi.org/10.1016/j.spasta.2016.05.005

Kou X, Jiang L, Bo Y, Yan S, Linna Chai L (2016) Estimation of land surface temperature through blending MODIS and AMSR-E data with the Bayesian maximum entropy method. Remote Sens 2016(8):105. https://doi.org/10.3390/rs8020105

Kuhnert PM, Martin TG, Griffiths SP (2010) A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol Lett 13:900–914

Lang Y, Christakos G (2019) Ocean pollution assessment by integrating physical law and site-specific data. Environmetrics 30(3):e2547

Law DCG, Bernstein KT, Serre ML, Schumacher CM, Leone PA, Zenilman JM, Miller WC, Rompalo AM (2006) Modeling a syphilis outbreak through space and time using the Bayesian maximum entropy approach. Ann Epidemiol 16:797–804

Lee S-J, Balling R, Gober P (2008) Bayesian maximum entropy mapping and the soft data problem in urban climate research. Ann Assoc Am Geogr 98:309–322

Lee S-J, Wentz EA, Gober P (2010) Space–time forecasting using soft geostatistics: a case study in forecasting municipal water demand for Phoenix, Arizona. Stoch Env Res Risk A 24:283–295

Li AH, Bo YC, Chen L (2013a) Bayesian maximum entropy data fusion of field-observed leaf area index (LAI) and landsat enhanced thematic mapper plus-derived LAI. Int J Remote Sens 34:227–246

Li AH, Bo YC, Zhu YX, Guo P, Bi J, He YQ (2013b) Blending multi-resolution satellite sea surface temperature (SST) products using Bayesian maximum entropy method. Remote Sens Environ 135:52–63

Li X, Li P, Zhu H (2013c) Coal seam surface modeling and updating with multi-source data integration using Bayesian geostatistics. Eng Geol 164:208–221

Mahbub P, Ayoko GA, Goonetilleke A, Egodawatta P, Kokot S (2010) Impacts of traffic and rainfall characteristics on heavy metals build-up and wash-off from urban roads. Environ Sci Technol 44:8904–8910

Martin TG, Burgman MA, Fidler F, Kuhnert PM, Low-Choy S, Mcbride M, Mengersen K (2012) Eliciting expert knowledge in conservation science. Conserv Biol 26:29–38

McCarthy MA, Masters P (2005) Profiting from prior information in Bayesian analyses of ecological data. J Appl Ecol 42:1012–1019

Messier KP, Akita Y, Serre ML (2012) Integrating address geocoding, land use regression, and spatiotemporal geostatistical estimation for groundwater tetrachloroethylene. Environ Sci Technol 46(5):2772–2780

Messier KP, Kane E, Bolich R, Serre ML (2014) Nitrate variability in groundwater of North Carolina using monitoring and private well data models. Environ Sci Technol 48:10804–10812

Messier KP, Campbell T, Bradley PJ, Serret ML (2015) Estimation of groundwater radon in North Carolina using land use regression and Bayesian maximum entropy. Environ Sci Technol 49:9817–9825

Modis K, Vatalis KI, Sachanidis C (2013) Spatiotemporal risk assessment of soil pollution in a lignite mining region using a Bayesian maximum entropy (BME) approach. Int J Coal Geol 112:173–179

Money ES, Carter GP, Serre ML (2009) Modern space/time geostatistics using river distances: data integration of turbidity and E.coli measurements to assess fecal contamination along the Raritan river in New Jersey. Environ Sci Technol 43:3736–3742

Money ES, Sackett DK, Aday DD, Serre ML (2011) Using river distance and existing hydrography data can improve the geostatistical estimation of fish tissue mercury at unsampled locations. Environ Sci Technol 45:7746–7753

Nol L, Heuvelink GBM, Veldkamp A, de Vries W, Kros J (2010) Uncertainty propagation analysis of an N_2O emission model at the plot and landscape scale. Geoderma 159:9–23

Olea RA (1999) Geostatistics. Kluwer Acadamic Publication, Boston

Puangthongthub S, Wangwongwatana S, Kamens RM, Serre ML (2007) Modeling the space/time distribution of particulate matter in Thailand and optimizing its monitoring network. Atmos Environ 41:7788–7805

Reyes JM, Serre ML (2014) An LUR/BME framework to estimate PM2.5 explained by on road mobile and stationary sources. Environ Sci Technol 48:1736–1744

Savelieva E, Utkin S Kazakov S, Demyanov V (2010) Modeling spatial uncertainty for locally uncertain data. In: geoENV VII – Geostatistics for environmental applications, pp 295–306.

Sedda L, Atkinson PM, Filigheddu MR, Cotzia G, Dettori S (2011) Spatio-temporal analysis of tree height in a young cork oak plantation. Int J Geogr Inf Sci 25:1083–1096

Serre ML, Christakos G (2002) BME-based hydrogeologic parameter estimation in groundwater flow modelling. Acta Univ Carol Geol 46:566–570

Serre ML, Kolovos A, Christakos G, Modis K (2003) An application of the holistochastic human exposure methodology to naturally occurring arsenic in Bangladesh drinking water. Risk Anal 23:515–528

Shi T, Yang X, Christakos G, Wang J, Liu L (2015a) Spatiotemporal interpolation of rainfall by combining BME theory and satellite rainfall estimates. Atmos 6:1307–1326

Shi Y, Zhou X, Yang X, Shi L, Ma S (2015b) Merging satellite ocean color data with Bayesian maximum entropy method. IEEE J Sel Topics Appl Earth Observ Remote Sens 8:3294–3304

Shoemaker JS, Painter IS, Weir BS (1999) Bayesian statistics in genetics – a guide for the uninitiated. Trends Genet 15:354–358

Sun XL, Wu YJ, Lou YL, Wang HL, Zhang C, Zhao YG, Zhang GL (2015) Updating digital soil maps with new data: a case study of soil organic matter in Jiangsu, China. Eur J Soil Sci 66:1012–1022

Tang SL, Yang XF, Dong D, Li ZW (2015) Merging daily sea surface temperature data from multiple satellites using a Bayesian maximum entropy method. Front Earth Sci 9:722–731

Tang Q, Bo Y, Zhu Y (2016) Spatiotemporal fusion of multiple- satellite aerosol optical depth (AOD) products using Bayesian maximum entropy method. J Geophys Res Atmos 121:4034–4048. https://doi.org/10.1002/2015JD024571

Vyas VM, Tong SN, Uchrin C, Georgopoulos PG, Carter GR (2004) Geostatistical estimation of horizontal hydraulic conductivity for the Kirkwood-Cohansey aquifer. J Am Water Resour Assoc 40:187–195

Wei X, Chang N, Bai K (2020) A comparative assessment of multisensor data merging and fusion algorithms for high-resolution surface reflectance data. IEEE J Sel Topics Appl Earth Observ Remote Sens 13:4044–4059. https://doi.org/10.1109/JSTARS.2020.3008746

Wu JP, He J, Christakos G (2021) Quantitative analysis and modeling of earth and environmental data: space-time and spacetime data considerations. Elsevier, New York

Yang Y, Christakos G (2015) Spatiotemporal characterization of ambient PM2.5 concentrations in Shandong Province (China). Environ Sci Technol 49:13431–13438

Yu HL, Chu HJ (2010) Understanding space-time patterns of groundwater system by empirical orthogonal functions: a case study in the Choshui River alluvial fan, Taiwan. J Hydrol 381:239–247

Yu HL, Wang CH (2010) Retrospective prediction of intraurban spatiotemporal distribution of PM2.5 in Taipei. Atmos Environ 44:3053–3065

Yu HL, Wang CH (2013) Quantile-based Bayesian maximum entropy approach for spatiotemporal modeling of ambient air quality levels. Environ Sci Technol 47:1416–1424

Yu HL, Kolovos A, Christakos G, Chen JC, Warmerdam S, Dev B (2007) Interactive spatiotemporal modelling of health systems: the SEKS-GUI framework. Stoch Env Res Risk A 21:555–572

Yu HL, Wang CH, Liu MC, Kuo YM (2011) Estimation of fine particulate matter in Taipei using landuse regression and bayesian maximum entropy methods. Int J Environ Res Public Health 8(6): 2153–2169. https://doi.org/10.3390/ijerph8062153

Yu H-L, Ku S-J, Kolovos A (2012) Advanced space-time predictive analysis with STAR-BME. In: Proceedings of the 20th international conference on advances in geographic information systems. ACM, pp 593–596

Yu H-L, Angulo JM, Chen M-H, Wu J, Christakos G (2014) An online spatiotemporal prediction model for dengue fever epidemic in Kaohsiung (Taiwan). Biom J 56(3):428–440. https://doi.org/10.1002/bimj.201200270

Yu HL, Ku SC, Kolovos A (2016) A GIS tool for spatiotemporal modeling under a knowledge synthesis framework. Stoch Env Res Risk A 30:665–679

Zagouras A, Kolovos A, Coimbra CFM (2015) Objective framework for optimal distribution of solar irradiance monitoring networks. Renew Energy 80:153–165. https://doi.org/10.1016/j.renene.2015.01.046

Zhang C, Yang Y (2019a) Can the spatial prediction of soil organic matter be improved by incorporating multiple regression confidence intervals as soft data into BME method? Catena 178:322–334

Zhang C, Yang Y (2019b) Improving the spatial prediction of soil Zn by converting outliers into soft data for BME method. Stoch Env Res Risk A 33:855–864

Zhang FS, Yang ZT, Zhong SB, Huang QY (2016) Exploring mean annual precipitation values (2003–2012) in a specific area (36 degrees N-43 degrees N, 113 degrees E-120 degrees E) using meteorological, elevational, and the nearest distance to coastline variables. Adv Meteorol 2016:2107908. https://doi.org/10.1155/2016/2107908, 13 p