Christakos, G., Angulo, J. M. and H.-L. Yu. 2011. Constructing space-time PDFs in geosciences. *Boletín Geológico y Minero,* 122 (4): 531-542
ISSN: 0366-0176

Constructing space-time pdfs in Geosciences

G. Christakos[(1)], J. M. Angulo[(2)] and H.-L. Yu[(3)]

(1) Dept of Geography, SDSU, San Diego, CA, USA; and College of the Environment & Natural Resources, Zhejiang University, Hangzhou, China.
PierianSpring@gmail.com
(2) Dept of Statistics and Operations Research, UG, Granada, Spain.
jmangulo@ugr.es
(3) Dept of Bioenvironmental Systems Engineering, NTU, Taipei, Taiwan.
hlyu@ntu.edu.tw

ABSTRACT

The focus of this work is the comparative analysis of techniques for constructing multivariate probability density function (Mv-pdf) models that can be used in a variety of geomathematics applications. The paper is concerned with formal and substantive model building methods. The former includes models that are speculative and analytically tractable, whereas the latter is based on substantive knowledge synthesis. More specifically, the present work focuses on the factoras and copulas techniques of Mv-pdf building, and their comparative analysis. It also discusses a substantive Mv-pdf building method that generates models on the basis of natural knowledge bases and takes into account the contentual and contextual domain of the in situ situation. The methods are compared in terms of a simulation study.

Key words: BME, copulas, factoras, multivariate, probability, spatiotemporal

Construcción de fdp's espacio-temporales en geociencias

RESUMEN

El objetivo de este trabajo es realizar un análisis comparativo de técnicas para la construcción de modelos de función de densidad de probabilidad multivariante (Mv-fdp) que pueden utilizarse en una variedad de aplicaciones geomatemáticas. El interés del artículo se centra en métodos formales y sustantivos para el desarrollo de modelos. Los primeros incluyen modelos que son especulativos y analíticamente tratables, mientras que los últimos se basan en la síntesis de conocimiento sustantivo. Más concretamente, el presente trabajo se enfoca a las técnicas basadas en fáctoras y cópulas para la construcción de Mv-fdp's, y su análisis comparativo. Se discute asimismo un método de construcción de Mv-fdp's sustantivo que genera modelos a partir de bases de conocimiento natural y tiene en cuenta el dominio conceptual y contextual de la situación in situ. Los métodos se comparan en términos de un estudio de simulación.

Palabras clave: BME, cópulas, espacio-temporal, fáctoras, multivariante, probabilidad

VERSIÓN ABREVIADA EN CASTELLANO

Perspectiva metodológica

Un campo aleatorio espacio-temporal (s/trf) $X_{\boldsymbol{p}},\ \boldsymbol{p}=(\boldsymbol{s},t)$, *representando la evolución de un atributo natural, queda completamente caracterizado en términos de la función de densidad de probabilidad multivariante (Mv-fdp),* $f_{X;\boldsymbol{p}}=f_{X;\{\boldsymbol{p}_i\}}$ *(Christakos, 1992). En este trabajo se discuten las técnicas basadas en fáctoras, cópulas y síntesis de conocimiento para la construcción de modelos de Mv-fdp's. Fáctoras y cópulas comparten la idea de expresar la Mv-fdp desconocida en términos de alguna función mejor conocida, la fáctora o la cópula, y las correspondientes fdp's univariantes (Uv-fdp's), que se suponen conocidas. Un enfoque diferente se basa en la síntesis sustantiva de bases de conocimiento núcleo e in situ (Christakos, 1990, 2000), donde el modelo Mv-fdp emerge de la solución de un sistema de ecuaciones involucrando varios puntos del espacio-tiempo.*

Métodos basados en fáctoras y cópulas

Bajo condiciones generales, una Mv-fdp, $f_{X;\boldsymbol{p}}$, *puede expresarse en términos de sus Uv-fdp's* $f_{X;\boldsymbol{p}_i}$ *y una función multivariante de* $\chi_{\boldsymbol{p}_i}\ (i=1,\ldots,k)$. *El hecho de que una Uv-fdp (supongamos, Gaussiana) pueda estar relacionada con una Mv-fdp de una clase diferente (i.e., no Gaussiana) puede causar serios problemas en muchas aplicaciones in situ. La cuestión crítica es cómo extender las Uv-fdp's que puedan estar disponibles en la práctica a una Mv-fdp que se ajuste al atributo geológico de interés.*

La clase de s/trf factorizables *(de orden* k*) se define como sigue (Christakos, 1986, 1989): Sea* $\theta(\chi_{\boldsymbol{p}}),\ \chi_{\boldsymbol{p}}=(\chi_{\boldsymbol{p}_1},\ldots,\chi_{\boldsymbol{p}_k})$, *una función multivariante de* $L_2(R^k,\prod_{i=1}^k f_{X;\boldsymbol{p}_i})$ *con* $r_k=\int \boldsymbol{d}\chi_{\boldsymbol{p}}\prod_{i=1}^k f_{X;\boldsymbol{p}_i}\theta^2(\chi_{\boldsymbol{p}})<\infty$ *y*

$$f_{X;\boldsymbol{p}}=f(\chi_{\boldsymbol{p}_1},\ldots,\chi_{\boldsymbol{p}_k})=[\textstyle\prod_{i=1}^k f_{X;\boldsymbol{p}_i}]\theta(\chi_{\boldsymbol{p}}). \qquad (1)$$

Sean $\varpi_{ji}(\chi_{pi})$ conjuntos completos de polinomios de grado $j_i = 0,1,\ldots$ en $L_2(R^1, f_{X;\,pi})$, ortonormales con respecto a $f_{X;\,pi}$. Se puede escribir la función fáctora como

$$\theta(\chi_{\boldsymbol{p}}) = [\prod_{i=1}^{k} \sum_{j_i=0}^{\infty}](\theta_{j_1 \ldots j_k} \prod_{i=1}^{k} \varpi_{j_i}(\chi_{\boldsymbol{p}_i})) = \theta_{X;\{\boldsymbol{p}_i\}}, \tag{2}$$

donde los coeficientes $\theta_{j_1 \ldots j_k}$ satisfacen la relación de completitud $[\prod_{i=1}^{k} \sum_{j_i=0}^{\infty}]\theta^2_{j_1 \ldots j_k} = \mathrm{r}_k$. Por consiguiente, la Mv-fdp (1) se reduce a la descomposición (Christakos, 1986; 1992)

$$f_{X;\boldsymbol{p}} = [\prod_{i=1}^{k} f_{X;\boldsymbol{p}_i}]\theta_{X;\{\boldsymbol{p}_i\}}, \tag{3}$$

donde las fáctoras ($\theta_{X;\boldsymbol{p}} = \theta_{X;\{\boldsymbol{p}_i\}}$) expresan interacciones entre funciones univariantes de $\chi_{\boldsymbol{p}_i}$.

Una forma de determinar las $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ es en términos de la Ec. (2). En el caso bivariante, la Ec. (3) puede simplificarse a

$$f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2} = f_{X;\boldsymbol{p}_1} f_{X;\boldsymbol{p}_2} \sum_{j=0}^{\infty} \theta_j \varpi_j(\chi_{\boldsymbol{p}_1}) \varpi_j(\chi_{\boldsymbol{p}_2}), \tag{4}$$

para todo $\boldsymbol{p}_1, \boldsymbol{p}_2$. En este caso, $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2} = \sum_{j=0}^{\infty} \theta_j \varpi_j(\chi_{\boldsymbol{p}_1}) \varpi_j(\chi_{\boldsymbol{p}_2})$.

Un paso clave en la Ec. (4) es calcular ϖ_j que sean ortogonales con respecto a una Uv-fdp. Esxisten varios métodos con este fin, donde los ϖ_j incluyen polinomios de Hermite, Laguerre, Laguerre Generalizados, Legendre, Gegenbauer, Jacobi y Stieltjes-Wigert.

Si X_p es un s/trf factorizable, y $\phi(.)$ una function estrictamente monótona, entonces el campo aleatorio $Z_{\boldsymbol{p}} = \phi(X_{\boldsymbol{p}})$ es también factorizable. Entre otras propiedades, las fáctoras generan estimadores de sistemas no lineales que pueden ser superiores a los del filtro de Kalman (Christakos. 1989, 1992).

En general, cualquier Mv-fdp continua puede escribirse en términos de la así llamada función cópula *multivariante (Sklar, 1959; Genest y Rivest, 1993; Nelsen, 1999),*

$$F_{X;\{\boldsymbol{p}_i\}} = C_X(F_{X;\boldsymbol{p}_1}, \ldots, F_{X;\boldsymbol{p}_k}) = P_X[F_{X;\boldsymbol{p}_1} \leq \upsilon_{\boldsymbol{p}_1}, \ldots, F_{X;\boldsymbol{p}_k} \leq \upsilon_{\boldsymbol{p}_k}] = C_{X;\{\boldsymbol{p}_i\}}(\upsilon_{\boldsymbol{p}_1}, \ldots, \upsilon_{\boldsymbol{p}_k}) \tag{5}$$

donde $F_{X;\boldsymbol{p}_i}$ son distribuciones de probabilidad condicionada univariantes (Uv-fdc's), y $\upsilon_{\boldsymbol{p}_i}$ son realizaciones de campos aleatorios uniformemente distribuidos $U_{\boldsymbol{p}_i} = F^{-1}_{X;\boldsymbol{p}_i} \sim \mathcal{U}(0,1)$, $i = 1,\ldots,k$. Cualquier Mv-fdc con soporte en $[0,1]^k$ y marginales uniformes ha sido llamada una cópula (Mikosch, 2006a). La correspondiente densidad de cópula *$\varsigma_{X;\{\boldsymbol{p}_i\}}$ se define por $\varsigma_{X;\{\boldsymbol{p}_i\}} \boldsymbol{d}\upsilon = \boldsymbol{d}C_{X;\{\boldsymbol{p}_i\}}$. La $f_{X,\boldsymbol{p}}$ puede reformularse en términos de sus Uv-fdp's y la densidad de cópula multivariante como*

$$f_{X;\boldsymbol{p}} = f_{X;\{\boldsymbol{p}_i\}} = [\prod_{i=1}^{k} f_{X;\boldsymbol{p}_i}]\varsigma_{X;\{\boldsymbol{p}_i\}}. \tag{6}$$

De nuevo, $\varsigma_{X;\{\boldsymbol{p}_i\}}$ expresa una cierta forma de interacción entre Uv-fdp's.

Bajo ciertas condiciones, la cópulas proporcionan descripciones paramétricas útiles de Mv-fdp's no Gaussianas (Scholzel y Friederichs, 2008). La cópulas son invariantes en escala en el sentido de que la cópula de $Z_{\boldsymbol{p}} = \phi(X_{\boldsymbol{p}})$ es igual a la cópula de X_p si $\phi(.)$ es una función estrictamente monótona. No existe un planteamiento general para construir la cópula más apropiada para un atributo, mientras que la elección de una familia de cópulas para un problema in situ a menudo se basa no en un razonamiento sustantivo, sino en mera conveniencia matemática (Mikosch, 2006a y b). Además, las cópulas no resuelven satisfactoriamente el problema de la dimensionalidad (Scholzel y Friederichs, 2008). Con relación a cuestiones interpretativas, hay muchos atributos que no son continuo-valuados, lo que significa que el teorema de la transformada integral (del que depende la tecnología de cópulas) no puede implementarse. Por su definición, las cópulas fracasan en estructuras dependientes del tiempo.

Cópulas pueden relacionarse directamente con fáctoras. Suponiendo, por ejemplo, un s/trf con Bv-fdc absolutamente continua $F_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ y la correspondiente Bv-fdp $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2} = f_{X;\boldsymbol{p}_1} f_{X;\boldsymbol{p}_2} \theta(\chi_{\boldsymbol{p}_1}, \chi_{\boldsymbol{p}_2})$, se tiene que $\theta(F^{-1}_{X;\boldsymbol{p}_1}(\upsilon_{\boldsymbol{p}_1}), F^{-1}_{X;\boldsymbol{p}_2}(\upsilon_{\boldsymbol{p}_2}))$ es también una Bv-fdp en $[0,1]^2$ con Uv-fdp's uniformes, i.e. una función cópula bivariante. Esto puede extenderse para una Bv-fdc $F_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}$ con marginales arbitrarias $F_{Z;\boldsymbol{p}_1}$ y $F_{Z;\boldsymbol{p}_2}$, mediante la transformación en términos de una función estrictamente monótona adecuada $\phi(\cdot)$, $\zeta_{\boldsymbol{p}_i} = \phi(\chi_{\boldsymbol{p}_i})$.

Construcción sustantiva-el método BME

Las consecuencias de usar una fdp errónea en situaciones reales pueden ser severas. El propósito de la modelización sustantiva es incorporar en la derivación de la fdp tanto conocimiento físico como sea posible (Christakos, 1992, 2010). Christakos (1990, 2000) presentó un marco de síntesis de conocimiento general para construir Mv-pdf's de manera que incorpora la base de conocimiento general o núcleo, G-KB (consistente en leyes naturales, modelos teóricos, teorías científicas, relaciones empíricas), y la de conocimiento especificatorio, S-KB (conocimiento específico de lugar como datos duros, información incierta, fuentes secundarias), de la situación in situ. Un enfoque bien conocido de síntesis de conocimiento para construir Mv-fdp's es el de Entropía Máxima Bayesiana (BME). La S-KB está disponible en $\boldsymbol{p}_D = (\boldsymbol{p}_H, \boldsymbol{p}_S)$, donde $\boldsymbol{p}_H$: $\boldsymbol{p}_i$ ($i = 1,2,\ldots,m_H$) y $\boldsymbol{p}_S$: $\boldsymbol{p}_i$ ($i = m_H + 1,\ldots,m$) son los conjuntos de puntos en que se dispone de mediciones duras (exactas) y datos blandos (inciertos), respectivamente; $\boldsymbol{p}_G = (\boldsymbol{p}_D, \boldsymbol{p}_k)$ es el conjunto de puntos en que está disponible la G-KB; y $\boldsymbol{p}_k = (p_{k_1}, \ldots, p_{k_\rho})$ es el conjunto de puntos de la Mv-fdp, i.e.,

$$L_2(R^k, \prod_{i=1}^{k} f_{X;\boldsymbol{p}_i}), \tag{7}$$

*donde **A** es un parámetro de normalización; **g**G es un vector cuyos elementos representan la G-KB; μ_G es un vector espacio-temporal cuyos elementos asignan pesos apropiados a los elementos de **g**G y son las soluciones del sistema de ecuaciones*

$$\int d\chi_G (\boldsymbol{g}_G - \overline{\boldsymbol{g}}_G) e^{\mu_G \boldsymbol{g}_G^T} = 0, \tag{8}$$

y ξ_s representa la S-KB disponible. La Ec. (7) introduce un proceso de integración de S *con* G *de una manera física y lógicamente consistente.*

Estudios de simulación

Dentro de un dominio espacio × tiempo de tamaño [0,1]2 × 3*, considerar un atributo geológico homogéneo X_p, $\boldsymbol{p} = (\boldsymbol{s}, t)$, donde $\boldsymbol{s} = (s_1, s_2)$, con media constante $\overline{X}_p = 5$ y función de covarianza espacio-temporal separable $c_X(r,t) = c_0 e^{-3r/a_r} e^{-3t^2/a_t^2}$ con $[c_0, a_r, a_t] = [1, 0.5, 1]$. Se consideran diez datos blandos (inciertos) de atributo, localizados como se muestra en la Fig. 1, y siguiendo distribuciones Gamma con diferentes parámetros, ver Fig. 2. Los tres métodos bajo discusión se usan para estimar la fdp del atributo en la localización espacio-temporal no observada (0.4, 0.5, 2), ver Fig. 1. El método de krigeado ordinario convencional (KO; Olea, 1999), suponiendo datos blandos “endurecidos” (i.e., el valor medio del dato blando se usa en cada localización), se aplica también para comparación.*

Estimación basada en fáctoras

Basándose en la suposición de Gaussianidad de las distribuciones univariantes, los ϖ_j son polinomios de Hermite normalizados de grado j. Como se ha discutido antes, ϕ es una transformación monótona creciente tal que $Z_p = \phi(X_p)$, donde $\phi^{-1}(Z_p)$ es una Gaussiana estándar.

Estimación basada en cópulas

Puesto que el campo aleatorio subyacente viene caracterizado por su matriz de covarianza R, la correspondiente cópula es Gaussiana. La estimación espacio-temporal de la fdp marginal en una localización desconocida P_k puede llevarse a cabo en términos de la técnica de krigeado indicador (Kazianka y Pilz, 2010a).

Estimación basada en BME

En este estudio de simulación, la función de covarianza espacio-temporal separable forma parte de la G*-KB, en cuyo caso la fdp basada en* G *es Gaussiana. La* S*-KB incluye los tres datos inciertos con distribución Gamma. El método BME se utiliza para actualizar la fdp a posteriori en cualquier localización del dominio de estudio.*

Discusión

Las diferentes fdp’s en P_k calculadas usando los cuatro métodos se muestran en la Fig. 3, y en la Tabla 1 se listan estadísticas resumen de las simulaciones. La fdp en el punto de estimación obtenida por el análisis mediante fáctoras (combinado con krigeado indicador) tiene la menor desviación estándar. Sin embargo, se obtienen valores de fdp negativos (Fig. 3). Por consiguiente, las estadísticas de estimación (basadas en una pseudo-fdp) deberían interpretarse con cautela.

Las medias de estimación espacio-temporal obtenidas por los métodos de cópula y BME son similares (Tabla 1). Puesto que dan cuenta de información blanda (incierta), las estimaciones basadas en cópulas y BME experimentan mayor incertidumbre que la estimación KO. El análisis BME proporciona una estimación espacio-temporal más informativa que el método de cópulas, en términos de menores desviaciones estándar. Adicionalmente, el marco operativo Bayesiano del análisis BME genera una fdp realista en el punto de estimación.

Conclusiones

Este trabajo ha presentado tres métodos espacio-temporales diferentes, a saber fáctoras, cópulas y BME, para representar realidades in situ encontradas en muchas aplicaciones geocientíficas (información incierta no Gaussiana, estimación no lineal y dominio compuesto espacio-temporal).

Mientras que las fáctoras involucran series infinitas que han de ser truncadas, muchas cópulas están disponibles en forma cerrada. Esto ocurre al coste de algunas suposiciones restrictivas, tales como baja dimensionalidad, marginales uniformes y la aplicabilidad del teorema de la transformada integral. Las fáctoras permiten derivar clases más ricas de fdp’s aprovechando la ϕ-propiedad y fórmulas de generalización. La forma functional de $\theta_{X;\{\boldsymbol{p}_i\}}$ se da explícitamente en términos de polinomios conocidos, mientras que la forma explícita de $\varsigma_{X;\{\boldsymbol{p}_i\}}$ es generalmente desconocida y ha de derivarse cada vez. Las simulaciones numéricas muestran que los modelos de fdp generados por los tres métodos difieren uno de otro. Entre otras razones, los tres métodos incorporan las KBs disponibles de manera diferente. La forma de los estimadores subyacentes también difiere entre ellos.

Methodological Perspective

Many natural attributes (geological, mining, ecological, and environmental) vary across space-time following a law of change (Shvidler, 1965; Olea, 1999; Douaik *et al.,* 2004; Parkin *et al.,* 2005). Let X_p, $\boldsymbol{p} = (s,t)$, denote a spatiotemporal random field (s/trf) that represents such a natural attribute. The s/trf is fully characterized in terms of the multivariate probability density function (Mv-pdf), $f_{X;\boldsymbol{p}} = f_{X;\{\boldsymbol{p}_i\}}$ (Christakos, 1992). This work discusses three distinct techniques for building Mv-pdf models of an s/trf X_p, namely factoras, copulas and knowledge synthesis. While the methodological perspective underlying factoras and copulas is essentially the same, that underlying knowledge synthesis is different. More specifically, in the case of factoras and copulas the main idea is to express an unknown function, the Mv-pdf, in terms of a (hopefully) better-known function (the factora or the copula) and the corresponding univariate pdf's (Uv-pdf's) that are assumed known. The factora concept can be traced in the Gaussian tetrachoric series expansion of Pearson (1901); whereas the copula technology has its origins in the multivariate probability analysis of Sklar (1959). Although factora is apparently an older concept than copula, both concepts share some common features. Another approach of Mv-pdf model building, which is very different than the factora and copula methods, is based on the substantive synthesis of core and in situ knowledge bases (Christakos, 1990, 2000). The Mv-pdf model emerges from the solution of a set of equations that involve various space-time points. The three methods above can be used in various fields of geosciences (e.g., rainfall assessment, watersheds, soil contamination, mining and climate).

Factora Mv-PDF Modelling

Under certain general conditions in theory, an Mv-pdf, $f_{X;\boldsymbol{p}}$, can be expressed in terms of its Uv-pdf $f_{X;\boldsymbol{p}_i}$ and a multivariate function of $\chi_{\boldsymbol{p}_i}$ $(i=1,\ldots,k)$. A common feature of an Mv-pdf (e.g., Gaussian, lognormal, or gamma) is that the corresponding Uv-pdf's are all of the same kind (i.e., Gaussian, lognormal, or gamma, respectively). The inverse, however, is often not true. I.e., a Uv-pdf (say, Gaussian) may be associated with an Mv-pdf of a different kind (i.e., non-Gaussian). These aspects can cause serious problems in many in situ applications in which one deals with non-Gaussian geological attributes, X_p, that have different kinds of Uv-pdf $f_{X,\boldsymbol{p}_i}$ (e.g., the $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ is non-Gaussian, whereas the $f_{X;\boldsymbol{p}_1}$ is Gaussian and the $f_{X;\boldsymbol{p}_2}$ is gamma). In such cases, the critical question is how to extend the Uv-pdf's that may be available in practice to an Mv-pdf that fits the geological attribute of interest. This kind of problems constitutes a prime reason for the development of the factora and copula representations, which build Mv-pdf's starting from the Uv-pdf's.

Pearson's original insight (Pearson, 1901) can be extended in a non-Gaussian random field context, leading to the class of *factorable* s/trf as follows (Christakos, 1986, 1989): Let $\theta(\chi_{\boldsymbol{p}})$, $\chi_{\boldsymbol{p}}=(\chi_{\boldsymbol{p}_1},\ldots,\chi_{\boldsymbol{p}_k})$, be a multivariate function of $L_2(R^k,\Pi_{i=1}^k f_{X;\boldsymbol{p}_i})$ with $r_k = \int d\chi_{\boldsymbol{p}}\Pi_{i=1}^k f_{X;\boldsymbol{p}_i}\theta^2(\chi_{\boldsymbol{p}})<\infty$, in which case one can write

$$f_{X;\boldsymbol{p}} = f(\chi_{\boldsymbol{p}_1},\ldots,\chi_{\boldsymbol{p}_k}) = [\Pi_{i=1}^k f_{X;\boldsymbol{p}_i}]\theta(\chi_{\boldsymbol{p}}). \tag{1}$$

Next, let $\varpi_{j_i}(\chi_{\boldsymbol{p}_i})$ be complete sets of polynomials of degree $j_i=0,1,\ldots$ in $L_2(R^1,f_{X;\boldsymbol{p}_i})$ that are orthonormal with respect to $f_{X;\boldsymbol{p}_i}$. Then, one can write the factora function as

$$\theta(\chi_{\boldsymbol{p}}) = [\Pi_{i=1}^k\Sigma_{j_i=0}^\infty](\theta_{j_1\ldots j_k}\Pi_{i=1}^k\varpi_{j_i}(\chi_{\boldsymbol{p}_i})) = \theta_{X;\{\boldsymbol{p}_i\}} \tag{2}$$

where the coefficients $\theta_{j_1\ldots j_k}$ satisfy the corresponding completeness relationship $[\Pi_{i=1}^k\Sigma_{j_i=0}^\infty]\theta_{j_1\ldots j_k}^2=r_k$ (assures that the series expansions converges). Accordingly, the Mv-pdf (1) reduces to (Christakos, 1986; 1992)

$$f_{X;\boldsymbol{p}} = [\Pi_{i=1}^k f_{X;\boldsymbol{p}_i}]\theta_{X;\{\boldsymbol{p}_i\}}. \tag{3}$$

Eq. (3) decomposes the modelling of the Mv-pdf, $f_{X,\boldsymbol{p}}$, into the product of the Uv-pdf's (non-uniform, in general), $f_{X;\boldsymbol{p}_i}$, and the factoras $(\theta_{X;\boldsymbol{p}}=\theta_{X;\{\boldsymbol{p}_i\}})$ that express interactions between univariate functions of $\chi_{\boldsymbol{p}_i}$. This is an advantage of the way factoras are defined over that of copulas. Also, the factoras may offer a measure of the deviation of the multivariate pdf from the product of the univariate pdf's. If $\theta_{X;\{\mathbf{p}_i\}}\equiv 1$, then $f_{X;\mathbf{p}}=[\Pi_{i=1}^k f_{X;\mathbf{p}_i}]$. Departure from 1 can be seen as deviation from independence. In fact, Eq. (3) can be written as $\theta_{X;\{\boldsymbol{p}_i\}}\equiv f_{X;\boldsymbol{p}}[\Pi_{i=1}^k f_{X;\boldsymbol{p}_i}]^{-1}$. Hence, $E_{f_{X;\boldsymbol{p}}}[\log\theta_{X;\{\boldsymbol{p}_i\}}]=$ $=E_{f_{X;\boldsymbol{p}}}\log\{f_{X;\boldsymbol{p}}[\Pi_{i=1}^k f_{X;\boldsymbol{p}_i}]^{-1}\}$. This can be interpreted as 'mutual information' between the $\chi_{\boldsymbol{p}_i}$'s (a k-dimensional extension of the usual concept of mutual information between two r.v.'s), a measure which quantifies departure from independence. (Note that similar quantities can be obtained in terms of Tsallis mutual information –dependence version–, replacing 'log' with the corresponding approximation used for Tsallis entropy). An s/trf X_p that satisfies Eq. (3) is called factorable (of order k).

Certain variations of the basic Eqs. (1)-(3) have been considered. For example, Kotz and Seeger (1991) reconsidered Eq. (1) for $k=2$ and replaced the symbol $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ with the symbol $\psi_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$, which was called the density weighting function (dwf). The $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ (or $\psi_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$) can be determined in several ways. One of these is in terms of Eq. (2). In the bivariate case, Eq. (3) can be reduced to the Bv-pdf expansion

$$f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2} = f_{X;\boldsymbol{p}_1}f_{X;\boldsymbol{p}_2}\Sigma_{j=0}^\infty\theta_j\varpi_j(\chi_{\boldsymbol{p}_1})\varpi_j(\chi_{\boldsymbol{p}_2}), \tag{4}$$

for all $\boldsymbol{p}_1$, $\boldsymbol{p}_2$. In this case, $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=\Sigma_{j=0}^\infty\theta_j\varpi_j(\chi_{\boldsymbol{p}_1})\varpi_j(\chi_{\boldsymbol{p}_2})$; e.g., $\theta_0=1$, $\theta_1=\rho_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ and $\theta_j\delta_{jj'}=\overline{\varpi_j(\chi_{\boldsymbol{p}_1})\varpi_{j'}(\chi_{\boldsymbol{p}_2})}$, with $\varpi_0(\chi_{\boldsymbol{p}_i})=1$ and $\varpi_1(\chi_{\boldsymbol{p}_i})=(\chi_{\boldsymbol{p}_i}-\overline{X_{\boldsymbol{p}_i}})\sigma_{\boldsymbol{p}_i}^{-1}$ for all space-time points. In Eq. (4) the marginals have the same form, and the same set of polynomials (and corresponding spaces) for $i=1$ and $i=2$ are considered. This equation is simpler than Eq. (2) particularized for $k=2$, since cross-terms (involving different polynomials) are discarded, that is, θ_{j_1,j_2} is assumed to be 0 for $j_1 \neq j_2$, and renamed just θ_j for $j_1=j_2$.

It is interesting that in Eq. (4) knowledge of lower-order statistics is linked to the first terms of the series, whereas that of higher-order statistics is linked to later terms of the series. Using 1st order (Hermite, Laguerre etc.) polynomial expansions, Eq. (4) yields $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=f_{X;\boldsymbol{p}_1}f_{X;\boldsymbol{p}_2}(1+\theta_1\chi_{\boldsymbol{p}_1}\chi_{\boldsymbol{p}_2})$, which is also what one obtains by performing a Taylor expansion of a Bv-pdf around a fixed value of $\theta_1=\theta_1^*$. An example is given in Sungur (1990). The approximation error depends on θ_1. Since usually $\theta_1^*=0$, the approximation is satisfactory around the independence case, and worsens as dependence increases. Another technique proposed in the literature is that in order to build $\psi_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ (essentially, $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$) one can use the following dwf representation (Long and Krzysztofowicz, 1995)

$$\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=\psi_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=1+\lambda\,\omega(F_{X;\boldsymbol{p}_1}(\chi_{\boldsymbol{p}_1}),F_{X;\boldsymbol{p}_2}(\chi_{\boldsymbol{p}_2})) \tag{5}$$

where ω is called the "covariance characteristic" and λ the "covariance scaler." Using Eq. (4) the ω can be expressed in terms of the expansion $\lambda\omega=\sum_{j=1}^{\infty}\theta_j\varpi_j(\chi_{\boldsymbol{p}_1})\varpi_j(\chi_{\boldsymbol{p}_2})$. Other authors (de la Peña *et al.,* 2006) have extended the work of Kotz and Seeger (1991) and Long and Krzysztofowicz (1995) in a multivariate setting so that Eq. (5) can be generalized as follows

$$\theta_{X;\{\boldsymbol{p}_i\}}=[1+U_{X;\{\boldsymbol{p}_i\}}] \tag{6}$$

where $U_{X;\{\boldsymbol{p}_i\}}=\sum_{c=2}^{k}\ \sum_{1\le i_1<..<i_c\le i_k}s_{i_1..i_c}(\chi_{\boldsymbol{p}_1}..\chi_{\boldsymbol{p}_c})$, and the $s_{i_1...i_c}$ satisfy three general conditions (integrability, degeneracy, and positive definiteness).

A key step in Eq. (4) is to calculate ϖ_j that are orthogonal with respect to a univariate pdf. There exist several methods for this purpose, where the ϖ_j include Hermite, Laguerre, Generalized Laguerre, Legendre, Gegenbauer, Jacobi, and Stieltjes-Wigert polynomials (e.g., the Gaussian, Gamma, or Poisson Uv-pdf is associated with Hermite, Laguerre, or Charlier polynomials). In Eq. (4) one needs to define factoras $\theta_{X;\{\boldsymbol{p}_i\}}$ with the prescribed mathematical properties and associated complete sets of orthogonal polynomials (the difficulty increases with $k>2$). For this purpose, a widely applicable method is based on the formula, $\varpi_j(\chi_{\boldsymbol{p}_i})=f_{X;\boldsymbol{p}_i}^{-1}\frac{d^j}{d\chi_{\boldsymbol{p}_i}^j}[\upsilon(\chi_{\boldsymbol{p}_i})^j f_{X;\boldsymbol{p}_i}]$, where $\upsilon(\chi_{\boldsymbol{p}_i})$ is a function that satisfies specific conditions (Christakos, 1992). This formula has been used to find the $\varpi_j(\chi)$ for a wide range of continuous functions $f_X(\chi)$, including the Gaussian, exponential, and Pearson (Type I). For illustration, if $f_{X;\boldsymbol{p}_i}=\frac{1}{\sqrt{2\pi}}e^{\chi_{\boldsymbol{p}_i}^2/2}$ $(-\infty\le\chi_{\boldsymbol{p}_i}\le\infty)$, the $\varpi_j=H_{a(j)}$ are Hermite polynomials, and the bivariate factora is $\theta_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=\sum_{j=0}^{\infty}\rho_X^{a(j)}H_{a(j)}(\chi_{\boldsymbol{p}_1})H_{a(j)}(\chi_{\boldsymbol{p}_2})$, where $\theta_j=\rho_X^{a(j)}$ and ρ_X is a correlation coefficient. For $a(j)=j$, the $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ is a bivariate Gaussian density; but for $a(j)=2j$, the $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ is non-Gaussian (Christakos, 1992: 162-164). This is not surprising, since to a given Uv-pdf one may associate more than one Bv-pdf. Many other examples are found in the cited literature.

If X_p is a factorable s/trf field, and $\phi(\cdot)$ is a strictly monotonic function, the random field $Z_p=\phi(X_p)$ is also factorable. This means that starting from known classes of factorable fields X_p, new classes Z_p can be constructed using different kinds of $\phi(\cdot)$. The Bv-pdf is written as

$$f_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}=f_{Z;\boldsymbol{p}_1}f_{Z;\boldsymbol{p}_2}\sum_{j=0}^{\infty}\theta_j\varpi_j[\phi^{-1}(\zeta_{\boldsymbol{p}_1})]\varpi_j[\phi^{-1}(\zeta_{\boldsymbol{p}_2})]. \tag{7}$$

In other words, the structure of the factora is preserved under strictly monotonic transformation (as in the copula framework). In Eq. (7), the values of the parameters θ_j change, see also below since they have to be computed on the transformed attribute. Another interesting property of the factora model is that it satisfies the relationship

$$\int d\chi_{\boldsymbol{p}_1}(\chi_{\boldsymbol{p}_1}-\overline{X_{\boldsymbol{p}_1}})f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=c_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}\sigma_{X;\boldsymbol{p}_2}^{-2}(\chi_{\boldsymbol{p}_2}-\overline{X_{\boldsymbol{p}_2}})f_{X;\boldsymbol{p}_2}, \tag{8}$$

for all $\boldsymbol{p}_1$, $\boldsymbol{p}_2$. As a matter of fact, Eq. (8) is valid for s/trf classes other than factorable (Christakos, 2010). In the special case that $\overline{X_{\boldsymbol{p}_i}}=\overline{X}=\mu$ and $f_{X;\boldsymbol{p}_i}=f_X$ (for all $\boldsymbol{p}_i$), Eq. (8) reduces to a more tractable form,

$$\int d\chi_{\boldsymbol{p}_1}(\chi_{\boldsymbol{p}_1}-\mu)f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}=\rho_{X;\boldsymbol{h},\tau}(\chi_{\boldsymbol{p}_2}-\mu)f_X, \tag{9}$$

$\boldsymbol{h}=|\boldsymbol{s}_1-\boldsymbol{s}_2|$ and $\tau=|t_1-t_2|$. A direct consequence of (9) is,

$$\overline{X_{\boldsymbol{p}_1}X_{\boldsymbol{p}_2}^m}=\rho_{X;\boldsymbol{h},\tau}\overline{X_{\boldsymbol{p}_2}^{m+1}}-\mu(\rho_{X;\boldsymbol{h},\tau}-1)\overline{X_{\boldsymbol{p}_2}^m}, \tag{10}$$

i.e., a higher-order, two-point dependence is conveniently expressed in terms of one-point functions. Yet another interesting property of the factoras is that they generate estimators of nonlinear state-nonlinear measurement systems that can be superior to those of the Kalman filter (Christakos. 1989, 1992). For example, the Kalman filter estimates include only linear correlation, whereas the factora estimates include linear and nonlinear correlations; also, the Kalman filter is limited to the estimation of lower-moments (mean and variance), whereas the factora estimator can provide lower- and higher-order moments.

Copula Mv-PDF Modelling

In general, any continuous Mv-pdf can be written in terms of the so-called multivariate *copula* function (Sklar, 1959; Genest and Rivest, 1993; Nelsen, 1999),

$$F_{X;\{p_i\}} = C_X(F_{X;p_1},...,F_{X;p_k}) = P_X[F_{X;p_1} \le \upsilon_{p_1},...,F_{X;p_k} \le \upsilon_{p_k}] = C_{X;\{p_i\}}(\upsilon_{p_1},...,\upsilon_{p_k}), \quad (11)$$

where $F_{X;p_1}$ are univariate conditional probability distributions (Uv-cdf's), and υ_{p_1} are realizations of uniformly distributed random fields $U_{p_i} = F^{-1}_{X;p_i} \sim \mathcal{U}(0,1),\ i = 1,...,k$. Thus, a copula is an Mv-cdf with uniform marginals. Otherwise said, any Mv-cdf with support on $[0,1]^k$ and uniform marginals has been termed a copula (Mikosch, 2006a). The corresponding multivariate *copula density* $\varsigma_{X;\{p_i\}}$ is defined by $\varsigma_{X;\{p_i\}} d\upsilon = dC_{X;\{p_i\}}$ (assuming copula continuity and differentiability). Since continuous functions are assumed, the $f_{X,p}$ is reformulated in terms of its Uv-pdf's and the multivariate copula density as

$$f_{X;p} = f_{X;\{p_i\}} = [\Pi^k_{i=1} f_{X;p_i}]\varsigma_{X;\{p_i\}}. \quad (12)$$

Eq. (12) basically decomposes the Mv-pdf $(f_{X,p})$ into the product of the Uv-pdf's $(f_{X;p_i})$ and the multivariate copula density $(\varsigma_{X;\{p_i\}})$ that expresses a certain form of interaction between Uv-pdf's. It is argued that this decomposition may offer some modeling flexibility. Several assumptions about the shape of the copula density are made, including the elliptic, the Archimedian, the Marshal-Olkin, the Pareto, the Gaussian and the *t*-copulas. In other words, assumption about the shape of the Mv-pdf has been replaced by assumptions concerning the shapes of the Uv-pdf's and the copula density. The proponents of the copula approach seem to claim that a copula density can be assumed in a more rigorous and realistic manner than a Mv-pdf, although critics have argued that there is no logical reason for choosing one copula over the other. Instead, one purely makes such decision based on mathematical convenience. Also, it must be kept in mind that a suitable copula should be chosen as well as the corresponding Uv-pdf's. In other words, assuming a priori a Gaussian copula is like assuming Gaussian marginals without any theoretical reason or empirical evidence (Christakos, 1992).

As is the case with all technical apparatuses, the copula technology has its pros and cons. Basically, copula is yet another tool to estimate non-Gaussian Mv-pdf's, which is suitable for some applications, but not for some others (Joe, 2006). Eq. (12) makes it possible to separate modelling multivariate dependent models into two parts: fitting unidimensional marginal distributions, and fitting joint dependence across marginal cdf's. However, Eq. (12) may increase calibration errors due to the extra step in estimation. Under certain conditions, copulas yield useful parametric descriptions of non-Gaussian Mv-pdf (Scholzel and Friederichs, 2008). Copula technology has the flexibility for modelling a multivariate distribution given different marginal behaviors. This technology may help avoid misspecification of the marginal distributions and focus directly on the dependence structure. Copulas are scale-invariant in the sense that the copula of $Z_p = \phi(X_p)$ is equal to the copula of X_p if $\phi(\cdot)$ is a strictly monotonic function. This property is particularly relevant for financial research. Since copulas are simply joint probability distributions with uniform marginals, the above representations (5)-(6) can be also applied to them. On the other hand, one should keep in mind that the copula technology mainly applies to continuous-valued attributes so that the marginals are uniform according to the so-called probability integral transform theorem. No general approach exists to construct the most appropriate copula for an attribute, whereas the choice of a copula family for an in situ problem is often based not on substantive reasoning, but on mathematical convenience (Mikosch, 2006a and b). If construction methods are available for componentwise maxima, not unique approaches can be established for a set of attributes that are not all extremes. This is also the case of univariate analysis, where distribution functions are usually chosen on the basis of theoretical observations and goodness-of-fit criteria. Direct interpretation of the copula alone does not offer insight about the complete stochastic nature of the attribute and there is no dependence separately from the marginals. Also, copulas do not solve satisfactorily the dimensionality problem (Scholzel and Friederichs, 2008). Interpretive issues concerning the copulas' in situ applications emerge too. There are many real world attributes that are not continuous-valued but rather discrete- or mixed-valued (e.g., daily rainfall), which means that the integral transform theorem (on which the copula technology of continuous variables relies) cannot be implemented, since the $F_{X;p_i}$ are no longer uniformly distributed on the interval (0,1), thus giving rise to so-called unidentifiability issues (Genest and Nešlehová, 2007). In this respect, although copulas can be used in simulation and robustness studies, they have to be used with caution because some properties do not hold in the discrete case. Attracted by the possibility to select arbitrary marginals, analysts sometimes forget that a suitable copula should be chosen as well as marginals. As defined, copulas are "static" dependence measures, that is, they are only snapshots of components' marginal distributions at a specific time. It is then futile to try to obtain the marginal distributions of dependent time series models, so copulas fail in time dependent structures.

Copulas can be directly related to factoras. Let us start with a s/trf X_p having an absolutely continuous Bv-cdf $F_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}$ and the corresponding Bv-pdf $f_{X;\boldsymbol{p}_1,\boldsymbol{p}_2} = f_{X;\boldsymbol{p}_1} f_{X;\boldsymbol{p}_2} \theta(\chi_{\boldsymbol{p}_1},\chi_{\boldsymbol{p}_2})$. If we define

$$F_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_1},\zeta_{\boldsymbol{p}_2}) = F_{X;\boldsymbol{p}_1,\boldsymbol{p}_2}(F^{-1}_{X;\boldsymbol{p}_1}(F_{Z;\boldsymbol{p}_1}(\zeta_{\boldsymbol{p}_1})),F^{-1}_{X;\boldsymbol{p}_2}(F_{Z;\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_2}))), \quad (13)$$

the $F_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}$ is a Bv-cdf of $(\zeta_{\boldsymbol{p}_1},\zeta_{\boldsymbol{p}_2})$ with marginals (Uv-cdf's) $F_{Z;\boldsymbol{p}_1}$ and $F_{Z;\boldsymbol{p}_2}$. Let $F_{X;\boldsymbol{p}_1}$, $F_{X;\boldsymbol{p}_2}$, $F_{Z;\boldsymbol{p}_1}$ and $F_{Z;\boldsymbol{p}_2}$ be absolutely continuous functions with Uv-pdf's $f_{X;\boldsymbol{p}_1}$, $f_{X;\boldsymbol{p}_2}$, $f_{Z;\boldsymbol{p}_1}$ and $f_{Z;\boldsymbol{p}_2}$. Then, the Mv-pdf of the S/TRF Z_p is given by

$$f_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_1},\zeta_{\boldsymbol{p}_2}) = f_{Z;\boldsymbol{p}_1}(\zeta_{\boldsymbol{p}_1}) f_{Z;\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_2}) \theta(F^{-1}_{X;\boldsymbol{p}_1}(F_{Z;\boldsymbol{p}_1}(\zeta_{\boldsymbol{p}_1})),F^{-1}_{X;\boldsymbol{p}_2}(F_{Z;\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_2}))). \quad (14)$$

This implies that if $F_{X;\boldsymbol{p}_1\boldsymbol{p}_2}$ is a Bv-cdf with marginals $F_{X;\boldsymbol{p}_1}$ and $F_{X;\boldsymbol{p}_2}$ and factora $\theta(\chi_{\boldsymbol{p}_1},\chi_{\boldsymbol{p}_2})$, then the $\theta(F^{-1}_{X;\boldsymbol{p}_1}(\upsilon_{\boldsymbol{p}_1}),F^{-1}_{X;\boldsymbol{p}_2}(\upsilon_{\boldsymbol{p}_2}))$ is also a Bv-pdf in $[0,1]^2$ with uniform Uv-pdf's, i.e. a bivariate copula function $\varsigma(F^{-1}_{X;\boldsymbol{p}_1}(\upsilon_{\boldsymbol{p}_1}),F^{-1}_{X;\boldsymbol{p}_2}(\upsilon_{\boldsymbol{p}_2}))$.

Now consider a Bv-pdf $F_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2}$ with arbitrary marginals $F_{Z;\boldsymbol{p}_1}$ and $F_{Z;\boldsymbol{p}_2}$. We can apply a suitable strictly monotonic function $\phi(\cdot)$ such that $\zeta_{\boldsymbol{p}_i} = \phi(\chi_{\boldsymbol{p}_i})$. To ensure that the transformed attributes follow the desired marginals the simplest method may be to set up $\phi(\cdot) = F^{-1}_{Z;\boldsymbol{p}_i}(F_{X;\boldsymbol{p}_i}(\cdot))$ (under the hypothesis of continuous functions). Then, Eq. (8) leads to

$$f_{Z;\boldsymbol{p}_1,\boldsymbol{p}_2} = f_{Z;\boldsymbol{p}_1} f_{Z;\boldsymbol{p}_2} \Sigma_{j=0}^{\infty} \theta_j \varpi_j [F^{-1}_{X;\boldsymbol{p}_1}(F_{Z;\boldsymbol{p}_1}(\zeta_{\boldsymbol{p}_1}))] \varpi_j [F^{-1}_{X;\boldsymbol{p}_2}(F_{Z;\boldsymbol{p}_2}(\zeta_{\boldsymbol{p}_2}))], \quad (15)$$

where $\zeta_{\boldsymbol{p}_i} = \phi(\chi_{\boldsymbol{p}_i}) = F^{-1}_{Z;\boldsymbol{p}_i}(F_{X;\boldsymbol{p}_i}(\chi_{\boldsymbol{p}_i}))$ and $\zeta_{\boldsymbol{p}_i} = F^{-1}_{Z;\boldsymbol{p}_i}(F_{X;\boldsymbol{p}_i}(\chi_{\boldsymbol{p}_i}))$ imply $\chi_{\boldsymbol{p}_i} = \phi^{-1}(\zeta_{\boldsymbol{p}_i})$, and $\chi_{\boldsymbol{p}_i} = F^{-1}_{X;\boldsymbol{p}_i}(F_{Z;\boldsymbol{p}_i}(\zeta_{\boldsymbol{p}_i}))$, respectively; and the combination of the two yields $\phi^{-1}(\zeta_{\boldsymbol{p}_i}) = F^{-1}_{X;\boldsymbol{p}_i}(F_{Z;\boldsymbol{p}_i}(\zeta_{\boldsymbol{p}_i}))$. In other words, $\phi(\cdot)$, $F_{X;\boldsymbol{p}_i}$ and $F_{Z;\boldsymbol{p}_i}$ are linked: if we apply an arbitrary function $\phi(\cdot)$ without knowing the corresponding $F_{Z;\boldsymbol{p}_i}$, it is not possible to build a Bv-pdf with known marginals, whereas if we seek a Bv-pdf with specified marginals $F_{Z;\boldsymbol{p}_i}$, this is possible by letting $\phi(\cdot) = F^{-1}_{Z;\boldsymbol{p}_i}(F_{X;\boldsymbol{p}_i}(\cdot))$. Moreover, on the basis of the above analysis it follows that $\Sigma_{j=0}^{\infty} \theta_j \varpi_j [F^{-1}_{X;\boldsymbol{p}_1}(\upsilon_{\boldsymbol{p}_1})] \varpi_j [F^{-1}_{X;\boldsymbol{p}_2}(\upsilon_{\boldsymbol{p}_2})]$ is a Bv-pdf with uniform Uv-pdf's in $[0,1]^2$, namely, a copula density. By means of factoras, we can build Bv-pdf with equal non-uniform marginals and apply monotone transformations to change the marginals. However, in doing so we pass implicitly through copula representation.

Substantive Construction-The BME Method

Experience shows that the consequences of using the wrong pdf in real world situations can be severe, which is why substantive modelling incorporates in the pdf derivation as much physical knowledge as possible (Christakos, 1992, 2010). As such, the substantive approach of constructing Mv-pdf adopts a definite science-based viewpoint. A prime source of substantive knowledge is provided by natural laws and scientific theories (Savelieva *et al.,* 2005; Yu *et al.,* 2007a). This is core knowledge that should be used in the derivation of the Mv-pdf models, which is a definite advantage of substantive model construction (e.g., prior probability problems of the so-called objective and subjective Bayesian analyses can be avoided).

To illustrate the difference between a formal and a substantive approach consider the very simple situation occurring in real world problems that involves the physical condition $X_{\boldsymbol{p}_2} > X_{\boldsymbol{p}_1}$ (e.g., the event rainfall depth and the corresponding critical depth used in the depth-duration-frequency curves). In this case, the strategy is to write $X_{\boldsymbol{p}_2} = X_{\boldsymbol{p}_1} + \varepsilon$, where ε is a positive variable, and to model the joint distribution of $X_{\boldsymbol{p}_1}$ and ε. In this example, the physical bound is well-defined a priori. However, it points out that possibly complex and unknown physical relationships can occur in situ and need to be incorporated in pdf model building.

In light of the above considerations, Christakos (1990, 2000) presented a general knowledge synthesis framework for constructing Mv-pdf in a manner that incorporates the general or core knowledge base, G-KB (consisting of natural laws, theoretical models, scientific theories, empirical relationships) and specificatory knowledge base, S-KB (site-specific knowledge like hard data, uncertain information, secondary sources) of the in situ situation. A well-known knowledge synthesis approach is Bayesian maximum entropy (BME). The BME based construction of Mv-pdf is compactly expressed in Eqs (16)-(17) below. In real world applications the S-KB is available at the set of points $\boldsymbol{p}_D = (\boldsymbol{p}_H, \boldsymbol{p}_S)$, where $\boldsymbol{p}_H$: $\boldsymbol{p}_i$ $(i = 1,2,...,m_H)$ and $\boldsymbol{p}_S : \boldsymbol{p}_i$ $(i = m_H + 1,...,m)$ are the set of points where hard (exact) measurements and soft (uncertain) data are available, respectively; $\boldsymbol{p}_G = (\boldsymbol{p}_D, \boldsymbol{p}_{\boldsymbol{k}})$ is the set of points where G-KB is available; and $\boldsymbol{p}_{\boldsymbol{k}} = (p_{k_1}, .., p_{k_\rho})$ is the set of points of the Mv-pdf, i.e. (Yu *et al.*, 2007a),

$$f_{X;\boldsymbol{p}_k} = A^{-1} \int d\chi_S \xi_S e^{\mu_G \boldsymbol{g}_G^T}, \tag{16}$$

where A is a normalization parameter $\boldsymbol{g}_G$ is a vector with elements representing the G-KB (including the natural law); μ_G is a space-time vector with elements that assign proper weights to the elements of $\boldsymbol{g}_G$ and are the solutions of the system of equations

$$\int d\chi_G (\boldsymbol{g}_G - \overline{\boldsymbol{g}}_G) e^{\mu_G \boldsymbol{g}_G^T} = 0, \tag{17}$$

and ξ_S represents the S-KB available. Eq. (16) introduces a process of integrating S with G in a physically and logically consistent manner. Eq. (16) accounts for local and non-local attribute dependencies across space-time. The above are noticeable advantages of the knowledge synthesis method of pdf building.

Let us consider some analytical examples of Mv-pdf building using Eqs. (16)-(17). Assume that G-KB includes the theoretical mean $\overline{X_G} = (\overline{X_{\boldsymbol{p}_1}}, .., \overline{X_{\boldsymbol{p}_m}}, \overline{X_{\boldsymbol{p}_{k_1}}}, .., \overline{X_{\boldsymbol{p}_{k_\rho}}})$ and (centered) covariance $\boldsymbol{c}_{\boldsymbol{p}_G} = \overline{(X_G - \overline{X_G})(X_G - \overline{X_G})^T}$ models; and S-KB includes hard data at points $\boldsymbol{p}_H$ and soft data of the interval type at points $\boldsymbol{p}_S$. Then, the Mv-pdf (16) is (Christakos, 2000)

$$f_{X;\boldsymbol{p}_k} = A^{-1} \varphi(\chi_{\boldsymbol{k}}; B_{k|H}\chi_H, \boldsymbol{c}_{k|H}) \int_{\boldsymbol{l}-\boldsymbol{B}_{s|kH}\chi_{kH}}^{\boldsymbol{u}-\boldsymbol{B}_{s|kH}\chi_{kH}} d\chi_S \varphi(\chi_S; 0, \boldsymbol{c}_{S|kH}), \tag{18}$$

where $\chi_{kH} = (\chi_{\boldsymbol{k}}, \chi_H)$, $B_{k|H} = \boldsymbol{c}_{k,H}\boldsymbol{c}_{H,H}^{-1}$, $\boldsymbol{c}_{k|H} = \boldsymbol{c}_{k,k} - B_{k|H}\boldsymbol{c}_{H,k}$, $B_{S|kH} = \boldsymbol{c}_{S,kH}\boldsymbol{c}_{kH,kH}^{-1}$, $\boldsymbol{c}_{S|kH} = \boldsymbol{c}_{S,S} - B_{S|kH}\boldsymbol{c}_{kH,S}$, $\boldsymbol{l} = (l_{m_H+1},...,l_m)$, and $\boldsymbol{u} = (u_{m_H+1},...,u_m)$; the $\varphi(\chi; \overline{\boldsymbol{x}}, \boldsymbol{c})$ denotes a Gaussian distribution with mean vector $\overline{\boldsymbol{x}}$ and covariance matrix c; and $A = \int_{\boldsymbol{l}-\boldsymbol{B}_{s|H}\chi_H}^{\boldsymbol{u}-\boldsymbol{B}_{s|H}\chi_H} d\chi_S \varphi(\chi_S; \boldsymbol{0}, \boldsymbol{c}_{s|H})$. Next, assume that the G-KB is as above, whereas the S-KB includes hard data at points $\boldsymbol{p}_H$ and soft data of the probabilistic type at points $\boldsymbol{p}_S$. In this case, the Mv-pdf (16) reduces to

$$f_{X;\boldsymbol{p}_k} = A^{-1} \phi(\chi_{\boldsymbol{k}}; B_{k|H}\chi_H, \boldsymbol{c}_{k|H}) \int d\chi_S f_H(\chi_S) \phi(\chi_S; B_{s|kH}\chi_{kH}, \boldsymbol{c}_{S|kH}). \tag{19}$$

where $A = \int d\chi_S f_S(\chi_S) \phi(\chi_S; B_{s|H}\chi_H, \boldsymbol{c}_{s|H})$.

Simulation Studies

Within a space (two spatial dimensions) × time domain of size $[0, 1]^2$ x 3 consider a homogeneous geological attribute X_p, $\boldsymbol{p} = (\boldsymbol{s}, t)$, where $\boldsymbol{s} = (s_1, s_2)$. The X_p has a spatiotemporally constant mean of $\overline{X}_p = 5$ units and a space-time separable covariance function of the form $c_X(r,t) = c_0 e^{-3r/a_r} e^{-3t^2/a_t^2}$ with $[c_0, a_r, a_t] = [1, 0.5, 1]$ (in suitable units). Ten soft (uncertain) attribute data are distributed as shown in Fig. 1. These uncertain data follow Gamma distributions with different parameters, as displayed in Fig. 2. The three methods discussed in the preceding sections are used to estimate the attribute pdf at the unmonitored space-time location (0.4, 0.5, 2), see Fig. 1.

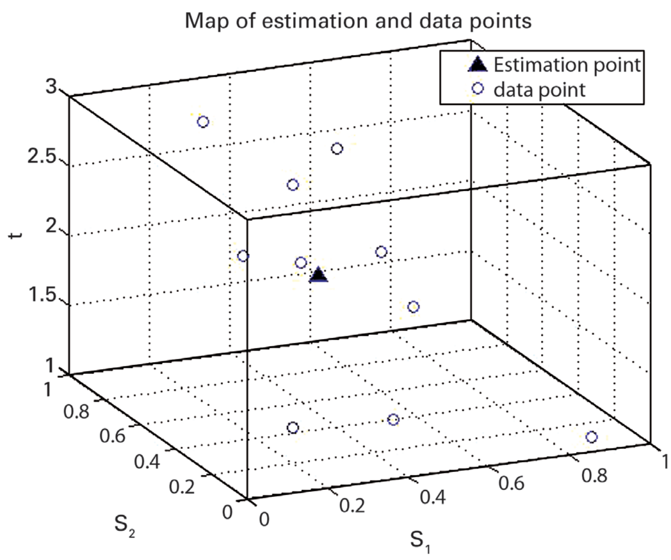

Figure 1. Space-time map of the data and estimation locations used in the simulation study.
Figura 1. Estadísticas de los fdp's marginales estimadas en la localización no observada.

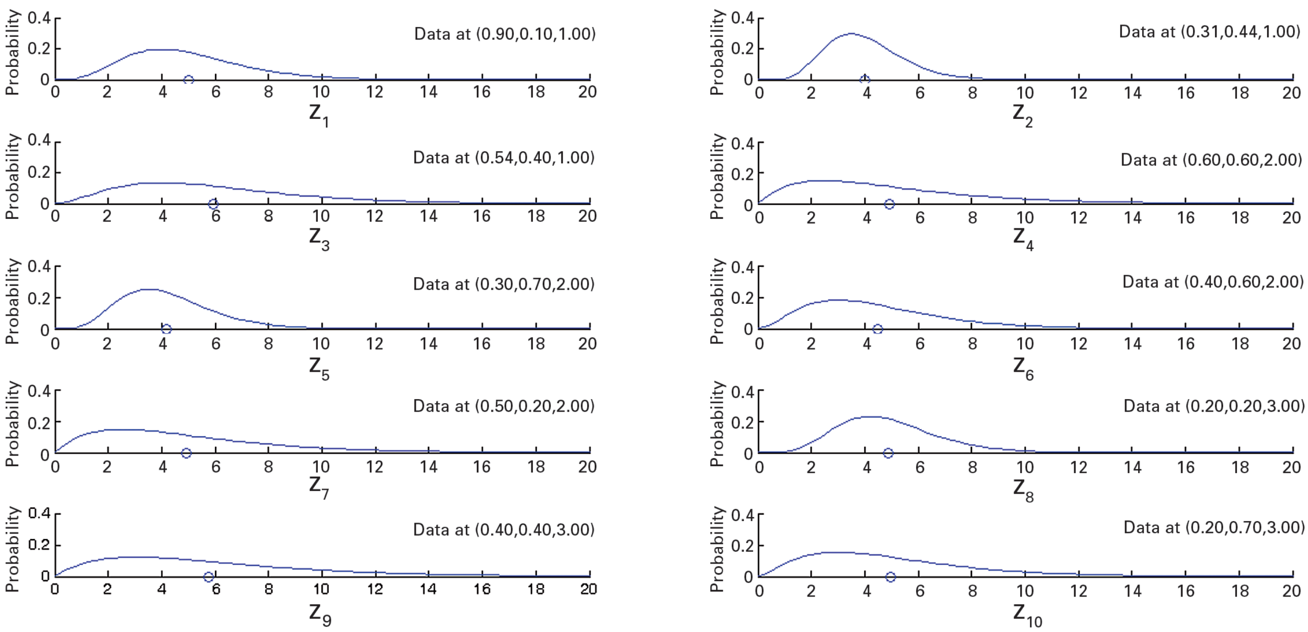

Figure 2. The space-time pdf's of the soft data in the simulation study.
Figura 2. fdp's espacio-temporales de los datos blandos en el estudio de simulación.

Factora-based spatiotemporal estimation

According to Eq. (7), the multivariate attribute distribution can be expressed in terms of a factora and the corresponding univariate distributions. In this simulation study, the Eq. (7) of the associated random field Z_s is rewritten as

$$f_{Z;p,p'} = f_{Z;p} f_{Z;p'} \Sigma_{j=0}^{\infty} \theta_j(h,\tau) \varpi_j[\phi^{-1}(z_p)] \varpi_j[\phi^{-1}(z_{p'})], \quad (20)$$

where the coefficients $\theta_j(h,\tau)$ are functions of the spatial and temporal distances between attribute locations; due to the Gaussian assumption of the univariate distributions, the ϖ_j are normalized Hermite polynomials of degree j. As discussed before, ϕ is a monotonic increasing transformation such that $Z_p = \phi(X_p)$ where $\phi^{-1}(Z_p)$ is standard Gaussian. The coefficients $\theta_j(h,\tau)$ generally express covariance functions among orthogonal polynomials of each degree, i.e. $\theta_j(h,\tau) = c_Z(\varpi_j[\phi^{-1}(z_p)], \varpi_j[\phi^{-1}(z_{p'})])$. In the case of normalized Hermite polynomials, $\theta_j(h,\tau) = [\rho_Z(h,\tau)]^j$ where $\rho_Z(h,\tau)$ denotes the corresponding correlation function. A special case of factora analysis is the spatial estimation technique of disjunctive kriging that involves orthogonal polynomials of the attribute rather than the original attribute itself (Olea, 1999). Indicator kriging (Emery, 2006) can be also used in the univariate analysis at location $\boldsymbol{s}_k$ as follows

$$F_Z(z_k) = \varphi_{0,z} + \Sigma_{j=1}^{\infty} \varphi_{j,z} [\Sigma_{i=1}^{n} \lambda_i \varpi_j(\phi^{-1}(z_{p_i}))], \quad (21)$$

where $\varphi_{j,z}$ are the factora coefficients of the associated indicator spatial random field with threshold z_k, and λ_i are kriging coefficients. As is the case with many indicator analysis results, $F_Z(z_k)$ is just a pseudo-cumulative distribution function (cdf) in the sense that $F_Z(z_k)$ may not be a no-decreasing function. In the present simulation study, only the first hundreds polynomials in Eq. (21) are used for the calculation of $F_Z(z_k)$. The truncation effect is relatively low in this case.

Copula-based spatiotemporal estimation

Copula analysis (Section III) also provides a way to specify separately the spatial dependence structure and the univariate marginal distribution of an Mv-pdf. In this case study, since the underlying random attribute field is characterized by its covariance matrix R, the corresponding copula is Gaussian and the Mv-pdf (12) reduces to

$$f_{Z;p_1,\ldots,p_n} = |R|^{-\frac{1}{2}} e^{-\frac{1}{2}u^T(R-I)u} \Pi_{i=1}^{n} \tfrac{1}{\sigma_i} \varphi(u_i), \quad (22)$$

where φ is the standard Gaussian pdf, $u_{p_i} = \Phi^{-1}(F_Z(z_{p_i}))$, Φ denotes the standard Gaussian cdf, and $F_Z(z_{pi})$ is the marginal cdf of Z_p at each $\boldsymbol{p}_i$, which is non-Gaussian, in general. Spatiotemporal estimation of the marginal pdf at an unknown location $\boldsymbol{p}_k$ can be made in terms of the indicator kriging technique (Kazianka and Pilz, 2010a), as follows

$$F_Z(z_k) = \sum_{i=1}^{n} \lambda_i \overline{I(Z_{p_i} < z_k)} = \sum_{i=1}^{n} \lambda_i F_Z(Z_{p_i} < z_k), \qquad (23)$$

where λ_i are indicator kriging coefficients with the corresponding semivariogram derived in terms of copulas as follows,

$$\gamma_{Z;p_i,p_j}(z_k, z_k) = C_\theta(F_{Z;p_i}(z_k), F_{Z;p_j}(z_k)) - F_{Z;p_i}(z_k) F_{Z;p_j}(z_k). \qquad (24)$$

In the case of Gaussian copulas, the copula functions can be estimated as follows

$$C_\theta(F_{Z;p_i}(z_k), F_{Z;p_j}(z_k)) = \Phi_{\rho(h,\tau)}(\Phi^{-1}(F_{Z;p_i}(z_k)), \Phi^{-1}(F_{Z;p_j}(z_k))), \qquad (25)$$

where $\Phi_{\rho(h,\tau)}$ denotes the standard bivariate Gaussian cdf with correlation function $\rho(h,\tau)$.

BME-based spatiotemporal estimation

In this simulation study, the space-time separable covariance function is part of the general knowledge base (G-KB) available concerning the attribute X_p, in which case the G-based pdf is Gaussian. The S-KB includes the three uncertain data with Gamma distributions. In the knowledge synthesis context, the operational BME method described in Section IV above was used to update the posterior pdf at any locations of the space-time study domain. According to BME, the problem of defining an appropriate Mv-pdf should be tackled by exploiting all sources of knowledge, and synthesizing them in a coherent framework (Vyas *et al.,* 2004; Akita *et al.,* 2007; Yu *et al.,* 2007b).

Discussion

In general, the implementation of the factora- and copula-based estimation techniques requires knowledge of the univariate pdf's at different times and locations across the entire study area. Since this information is not available, in most case studies (including the present one) the univariate pdf's are assumed to be Gaussian (which is also the maximum entropy pdf given the attribute mean and covariance). Hermite polynomials, Gaussian copulas, and Gaussian univariate attribute distributions are used by the factora, copula, and BME analysis, respectively. As mentioned earlier, for the technical details of these three methods the interested readers are referred to the relevant literature (e.g., Christakos, 1990, 1992, 2000; Bardossy and Li, 2008; Serinaldi, 2008, 2009; Kazianka and Pilz, 2010a, b; Song and Singh, 2010a, b) and references therein. In addition, the present simulation study used the mainstream ordinary kriging method (OK; Olea, 1999), which assumed "harden" soft data (i.e., the mean value of the soft datum was used at each location).

Although the true pdf at the space-time estimation point $\boldsymbol{p}_k$ is not available for direct comparison with the pdfs obtained by the above methods, some general observations were made under the given study conditions. The pdfs at $\boldsymbol{p}_k$ were calculated using the four methods. The pdfs are plotted in Fig. 3 and the summary statistics of the simulations are listed in Table 1. Interestingly, the methods lead to different pdfs at $\boldsymbol{p}_k$. As expected, in the case of OK the skewness is zero (reflecting the underlying Gaussian law). The pdf at the estimation point obtained by factora analysis (combined with indicator kriging) has the smallest std deviation. However, negative pdf values are obtained (Fig. 3) due to the indicator kriging feature to generate pseudo-pdf rather than a formal pdf (Emery, 2006), which could make estimation dubious. The negative values in the estimation pseudo-pdf, e.g., have the effect of changing the sign for z_p values approximately in the range [6.5, 8.2]. This issue also emerges in Table 1. Factora estimation statistics are affected by the small interval of negative values in the estimated pseudo-pdf: the mean is smaller since it is "attracted" down towards zero by the negative values; the std deviation is underestimated since the corresponding squares are computed negatively;

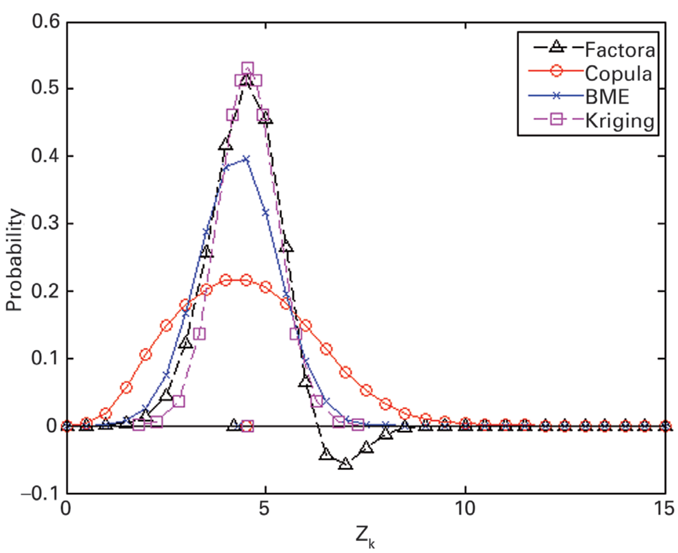

Figure 3: The marginal pdf's at the space-time estimation point obtained by the factora, copula, BME, and OK (ordinary kriging) methods (Z_k is the estimation value at the estimation location).
Figura 3. fdp's marginales en el punto de estimación espacio-temporal obtenidos por los métodos de fáctoras, copulas, BME, KO (Krigeado ordinario) (Z_k es el valor de estimación en la localización de la estimación).

Method	Mean	Std Deviation	Skewness
Ordinary Kriging	4.5509	0.7397	0.00
Factora	4.2066	0.3827	-64.3452
Copula	4.5067	1.7346	0.7111
BME	4.4032	0.7494	-1.5931

Table 1: Statistics of the estimated marginal pdf at the unknown location.
Tabla 1: Estadísticas de las fdp's marginales estimadas en la localización no observada.

and the skewness is also distorted at $\boldsymbol{p}_k$ (negative estimation skewness, although all soft data had positive skewness). Accordingly, estimation statistics (based on a pseudo-pdf) should be interpreted with caution.

One notices that the space-time estimation means obtained by the copula and BME methods are similar (about 4.5 and 4.4, respectively; Table 1). Since they account for soft (uncertain) information, the copula- and BME-based estimates experience higher uncertainty than the OK estimate that accounts only for hard data. BME analysis provides more informative space-time estimation than the copula method, in terms of smaller std deviations. In addition, the operational Bayesian framework of BME analysis generates a realistic pdf at the estimation point, unlike the pseudo-pdfs obtained by the factora and copula methods that used indicator kriging.

The space-time factoras, copulas and BME methods can be used, inter alia, to generate realistic non-Gaussian simulations of environmental processes (e.g., rainfall assessment, climate and watersheds) and to estimate extreme geohydrological values (e.g., droughts and hazards). On the other hand, under the mainstream statistics assumptions (normality, linearity, space-time separation etc.), geoscientific studies of the above kind can yield unrealistic results.

Conclusions

This work presented three different space-time methods to account for in situ realities encountered in many geoscientific applications (non-Gaussian uncertain information, nonlinear estimation and composite space-time domain). In a nutshell, underlying both the copula and factora technologies is the basic idea of replacing an unknown entity (original Mv-pdf) with another unknown entity (factora or copula), which is supposedly easier to infer from the available data and manipulate analytically. Whether this is actually a valid claim of practical significance depends on a number of technical and substantive issues, some of which were touched upon in this work.

In technical terms, a prime advantage of the copula technology is its analytical tractability, although this is mainly valid in low dimensions (2-4). While factoras involve infinite series that have to be truncated, many copulas are available in a closed-form. This comes at the cost of some restrictive assumptions, such as low dimensionality, uniform marginals, and the applicability of the integral transform theorem. Attempts to involve transforms of uniform marginals are rather *ad hoc* and can add considerable complexity to the process. Potential advantages of factoras include the elimination of restrictive requirements (uniform marginals, integral transform theorem, etc.), and the rich classes of pdfs derived by taking advantage of the Φ-property and generalization formulas. The functional form of $\theta_{X;\{\boldsymbol{p}_i\}}$ is explicitly given in terms of known polynomials, whereas the explicit form of $\varsigma_{X;\{\boldsymbol{p}_i\}}$ is generally unknown and needs to be derived every time. Numerical simulations show that the pdf models generated by the three methods differ from each other. There are certain reasons for this. For example, factoras and copulas use different pdf expansions. The methods use different spatiotemporal estimation techniques: factora analysis is combined with indicator kriging, whereas BME does not make such an assumption. Also, the three methods incorporate the available KBs in different ways. Last but not least, the form of the underlying estimators also differ from one method to the next. OK uses a linear estimator, whereas the BME space-time estimator can be nonlinear, in general.

Acknowledgements

José M. Angulo was partially supported by projects MTM2009-13250 of the SGPI, and P08-FQM-03834 of the Andalusian CICE, Spain.

References

Akita Y., Carter G., and Serre M. L. 2007. "Spatiotemporal non-attainment assessment of surface water tetrachloroethene in New Jersey." *Jour of Environmental Quality* 36(2): 508–520.

Bardossy, A. and Li, J., 2008. "Geostatistical interpolation using copulas." *Water Resources Research* 44(7): W07412, doi:10.1029/2007WR006115.

Christakos G. 1986. "Recursive estimation of nonlinear-state nonlinear-observation systems." *Research Rep. OF.86–29*, Kansas Geological Survey, Lawrence, KS.

Christakos G. 1989. "Optimal estimation of nonlinear-state nonlinear-observation systems." *Jour. of Optimization Theory and Application* 62: 29–48.

Christakos G. 1990. "A Bayesian/maximum-entropy view to the spatial estimation problem." *Mathematical Geology* 22(7): 763-776.

Christakos, G. 1992. *Random Field Models in Earth Sciences*. Academic Press, San Diego, CA.

Christakos, G. 2000. *Modern Spatiotemporal Geostatistics*. Oxford Univ. Press, New York, NY.

Christakos, G. 2010. *Integrative Problem-Solving in a Time of Decadence*. Springer, New York, NY.

de la Peña V.H., Ibragimov R. and Sharakhmetov S. 2006. "Characterizations of joint distributions, copulas, information, dependence and decoupling, with applications to time series." *IMS Lecture Notes–Monograph Series 2nd Lehmann Symposium–Optimality* 49: 183–209.

Douaik A., van Meirvenne M., Toth T. and Serre M. L. 2004. "Space-time mapping of soil salinity using probabilistic BME." *Stochastic Environmental Research and Risk Assessment* 18: 219–227.

Emery X. 2006. A disjunctive kriging program for assessing point-support conditional distributions. *Computers & Geosciences* 32(7): 965-983

Genest C. and Nešlehová J. 2007. "A primer on copulas for count data." *Astin Bulletin* 37(2): 475–515.

Genest C. and Rivest L.-P. 1993. "Statistical inference procedures for bivariate Archimedean copulas." *Jour. of American Statistical Association* 88: 1034–1043.

Joe H. 2006. "Discussion of 'Copulas: Tales and facts,' by Thomas Mikosch." *Extremes* 9: 37–41.

Kazianka H. and Pilz J. 2010a. "Copula-based geostatistical modeling of continuous and discrete data including covariates." *Stochastic Environmental Research and Risk Assessment* 24(5): 661-673.

Kazianka H. and Pilz J. 2010b. "Spatial interpolation using copula-based geostatistical models." In *GeoEnv VII*: *Geostatistics for Environmental Applications*: 307-319. Atkinson P.M.M. and Lloyd C.D.D. (eds), Springer, New York, NY.

Kotz S. and Seeger J.P. 1991. "A new approach to dependence in multivariate distributions." In *Advances in Probability Distributions*: 113-127, Dall'Aglio G., Kotz S. and Salinetti G. (eds.), Kluwer, Dordrecht, the Netherlands.

Kotz S., Balakrishnana N., and Johnson N.L. 2000. *Continuous Multivariate Distributions*. Wiley, New York, NY.

Long D. and Krzysztofowicz R. 1995. "A family of bivariate densities constructed from marginals." *Jour of the Amer. Statist. Assoc*. 90(430): 739-746.

Mikosch T. 2006a. "Copulas: tales and facts." *Extremes* 9: 3–20.

Mikosch T. 2006b. "Copulas: tales and facts–rejoinder." *Extremes* 9: 55–62.

Nelsen, R. 1999. *An Introduction to Copulas*. Springer, New York, NY.

Olea R. A., 1999. *Geostatistics for Engineers and Earth Scientists.* Kluwer Acad. Publ., Boston, MA.

Parkin R., Savelieva E., and Serre M. L. 2005. "Soft geostatistical analysis of radioactive soil contamination." In GeoEnv V-Geostatistics for environmental applications. Renard Ph. (ed.). Kluwer Acad., Dordrecht, the Netherland.

Pearson K. 1901. "Mathematical contributions to the theory of evolution, VII: On the correlation of characters not quantitatively measurable:" *Philosophical Trans. Royal Soc. of London, Series A* 195: 1-47.

Savelieva E., Demyanov V., Kanevski M., Serre M. L., and Christakos G. 2005. "BME-based uncertainty assessment of the Chernobyl fallout." *Geoderma* 128: 312–324.

Scholzel, C., & Friederichs, P. (2008). Multivariate non-normally distributed random variables in climate research – introduction to the copula approach. *Nonlinear Processes in Geophysics* 15: 761–772.

Serinaldi, F. 2008. "Analysis of inter-gauge dependence by Kendall's τk, upper tail dependence coefficient, and 2-copulas with application to rainfall fields." *Stochastic Environmental Research and Risk Assessment* 22(6): 671–688.

Serinaldi, F. 2009. "Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective." *Stochastic Environmental Research and Risk Assessment* 23(5): 677-693.

Sklar A. 1959. "Fonctions de repartition à n dimensions et leurs marges." *Publications de l'Institut de Statistique de L'Universite de Paris* 8: 229–231, Paris, France.

Song S. and Singh V.P. 2010a. "Meta-elliptical copulas for drought frequency analysis of periodic hydrologic data." *Stochastic Environmental Research and Risk Assessment* 24(3): 425-444.

Song S. and Singh V.P. 2010b. "Frequency analysis of droughts using the Plackett copula and parameter estimation by genetic algorithm." *Stochastic Environmental Research and Risk Assessment* 24(5): 783-805.

Sungur E.A. 1990. "Dependence information in parameterized copulas." *Commun. Statist.-Simula*. 19(4): 1339-1360.

Shvidler M. I. 1965. "Sorption in a plane-radial filtration flow." *Journal of Applied Mechanics and Technical Physics* 6(3): 77–79.

Vyas V. M., Tong S. N., Uchrin C., Georgopoulos P. G., and Carter G. P. 2004. "Geostatistical estimation of horizontal hydraulic conductivity for the Kirkwood-Cohansey aquifer." *Jour of the American Water Resources Associates* 40(1): 187–195.

Yu H.-L., Christakos G., Modis K., and Papantonopoulos G. 2007a. "A composite solution method for physical equations and its application in the Nea Kessani geothermal field (Greece)." *Jour of Geophysical Research-Solid Earth* 112: B06104. doi:10.1029/2006JB004900.

Yu H.-L., Kolovos A., Christakos G., Chen J.-C., Warmerdam S., and Dev B. 2007b. "Interactive spatiotemporal modelling of health systems: the SEKS-GUI framework." *Stochastic Environmental Research and Risk Assessment* 21(5): 555–572.

Recibido: febrero 2011
Revisado: agosto 2011
Aceptado: septiembre 2011
Publicado: octubre 2011