BEYOND CORRELATION IN SPATIAL STATISTICS MODELING

JHAN RODRÍGUEZ AND ANDRÁS BÁRDOSSY

Abstract. This paper hast two aims. Firstly, to introduce a type of spatial model whereby joint cumulants of order greater than two (i.e. covariances) can be handled without prohibitively increasing the number of model parameters; joint cumulants were seen in Rodríguez and Bárdossy (2013) as convenient means, both conceptually and practically, for tackling interactions among more than two variables simultaneously. The type of model introduced forms a flexible subclass of the elliptical family and can be seen as an extension to the Normal model; it capitalizes on current covariance function estimation techniques, and it is seen how this model can be ensured to be spatially consistent. It is suggested that the cumulant generating function of the model can be seen as a natural extension to the covariance function. Secondly, we illustrate the need to consider spatial dependence structures beyond correlation. For this end, we present a simulation study in which relevant characteristics of a big field are to be inferred from realizations of a 30-dimensional vector ("gauging network"). Three random fields are presented having the same covariance function, and practically indistinguishable 1 and 2-dimensional marginal distributions and copulas. Yet the fields present very different dependence characteristics: In the context of daily rainfall modeling this may lead to substantial underestimation of aggregating statistics that are important for risk quantification and flood return period estimation; in mining geostatistics to substantial underestimation of the ore concentration at unwelled locations. The topics of interpolation to ungauged sites, simulation and estimation are dealt with.

1. Introduction

A generic Spatial model can be defined as follows (see Cressie (1991)): Let $\mathbf{s} \in \mathbb{R}^g$ be a generic location in the G-dimensional Euclidean space and suppose the potential datum $\mathbf{Z}(\mathbf{s})$ at spatial location $\mathbf{s}$ is a random quantity. Let $\mathbf{s}$ vary over an index set $\mathbf{D} \subset \mathbb{R}^g$, then a multivariate random field is generated

$$\{\mathbf{Z}(\mathbf{s}) : \mathbf{s} \in \mathbf{D}\} \tag{1.1}$$

For example, if $\mathbf{D} = \{\mathbf{s_1}, \mathbf{s_2}, \mathbf{s_3}, \ldots, \mathbf{s_J}\}$ is a fixed finite set, we obtain the random field

$$(\mathbf{Z}(\mathbf{s_1}), \ldots, \mathbf{Z}(\mathbf{s_J})) \tag{1.2}$$

in case set $\mathbf{D} \subset \mathbb{R}^g$ is countably infinite and fixed, the generated field can be written as

$$(\mathbf{Z}(\mathbf{s_1}), \mathbf{Z}(\mathbf{s_2}), \ldots) \tag{1.3}$$

The above two instances are usually called "lattice data models". If set $\mathbf{D} \subset \mathbb{R}^g$ is fixed and non-countable, then the model is a "Geostatistical model", the name coming from the original context in which this type of model was first developed.

If the index set $\mathbf{D} \subset \mathbb{R}^g$ is not provided in advance, and the model for the process of interest can be decomposed into two steps: 1. a location $\mathbf{s}$ is generated on $\mathbb{R}^g$; 2.

given location $\mathbf{s}$, a random quantity is generated $\mathbf{Z}(\mathbf{s})$. Then the model is a "Point pattern" model; important instances of this model are the "Poisson point process" and the "Cox Process" (see Cressie and Wikle (2011)), in which only the location generation mechanisms are random, and $\mathbf{Z}(\mathbf{s}) = 1, \forall \mathbf{s}$.

The above models do not exhaust the Spatial Statistics types of models. The reader is referred to Cressie and Wikle (2011) for more details. For the sake of simplicity, we focus in this work on simple versions of the Lattice models and $g = 2$, but the concepts and ideas presented in this paper can be also applied to the other types of Spatial Statistics models.

Association between every two components of $(\mathbf{Z}(\mathbf{s_1}), \ldots, \mathbf{Z}(\mathbf{s_J}))$ in Spatial Statistics can be modeled in terms of a covariance function, C,

$$cov(\mathbf{Z}(\mathbf{s_i}), \mathbf{Z}(\mathbf{s_j})) = C(\|\mathbf{s_i} - \mathbf{s_j}\|) \tag{1.4}$$

where $\|\mathbf{s_i} - \mathbf{s_j}\|$ is the euclidean distance between $\mathbf{s_i}$ and $\mathbf{s_j}$. This covariance function must ensure positive-definiteness of the resulting covariance matrix. By way of illustration, two popular covariance functions are:

Powered-exponential:: given by equation

$$C(d) = \sigma_0^2 . I(d = 0) + \sigma_1^2 \exp\left(-(d/\theta_1)^{\theta_2}\right) \tag{1.5}$$

where $I(*)$ stands for the indicator function.

Mátern's:: given by equation

$$C(d) = \sigma_0^2 . I(d = 0) + \sigma_1^2 \left[2^{\theta_2 - 1}\Gamma(\theta_2)\right]^{-1} [d/\theta_1]^{\theta_2} K_{\theta_2}(d/\theta_1) \tag{1.6}$$

where $\Gamma(*)$ stands for the Gamma function and $K_{\theta_2}(d/\theta_1)$ for the modified Bessel function of the second kind of order θ_2 (see, for example Abramowitz (1972)).

Parameters $\left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)$ are the covariance function parameters. Hence, only a reduced number of parameters must be estimated in order to find the covariance between every two components $\mathbf{Z}(\mathbf{s_i})$ and $\mathbf{Z}(\mathbf{s_j})$, given locations $\mathbf{s_i}$ and $\mathbf{s_j}$.

The Normal model is a common model in Spatial Statistics for components corresponding to every finite set of locations,

$$(\mathbf{Z}(\mathbf{s_1}), \ldots, \mathbf{Z}(\mathbf{s_J})) \sim N_J(\mu, \Sigma)$$

where the covariance matrix is given by $\Sigma_{ij} = C(\|\mathbf{s_i} - \mathbf{s_j}\|)$. Under the gaussian model, the whole distribution is defined in terms of a vector of means $\mu \in \mathbb{R}^J$ and parameters $\left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)$. It is often the case that the mean vector is represented as a function ξ of geographic coordinates of $\mathbf{s_j}$ or of an additional variable ("external drift") related to such location ,

$$\mu_j = \mu(\xi(\mathbf{s_j})) \tag{1.7}$$

For a new location $\mathbf{s_k} \notin \{\mathbf{s_1}, \mathbf{s_2}, \mathbf{s_3}, \ldots, \mathbf{s_J}\}$, the joint distribution of

$$(\mathbf{Z}(\mathbf{s_1}), \ldots, \mathbf{Z}(\mathbf{s_J}), \mathbf{Z}(\mathbf{s_k}))$$

can be readily found under the Normal model: one adds component $\mu_k = \mu(\xi(\mathbf{s_k}))$ to the means vector, and extends the covariance matrix by $\Sigma_{ik} = C(\|\mathbf{s_i} - \mathbf{s_k}\|)$, for each $\mathbf{s_i} \in \{\mathbf{s_1}, \mathbf{s_2}, \mathbf{s_3}, \ldots, \mathbf{s_J}\}$. The model is thus completely specified. This is one of the reasons why the Normal model is very convenient conceptually, and is often used in practice, if necessary after applying a suitable transformation to data.

The family of elliptical distributions can be seen as the wider family to which both the multivariate Normal and the multivariate Student distributions belong. The classical definition, according to Cambanis et al. (1981), is as follows:

Definition. Let $\mathbf{X}$ be a J-dimensional random vector, $\mu \in \mathbb{R}^J$ and Σ a $J \times J$, non-negative definite matrix. Let $\phi_{\mathbf{X}-\mu}(\mathbf{t}) : \mathbb{R}^J \to [0, +\infty)$ be the characteristic function of $\mathbf{X}-\mu$. If $\phi_{\mathbf{X}-\mu}(\mathbf{t}) = \Psi\left(\mathbf{t}\Sigma\mathbf{t}'\right)$ for some function $\Psi(s) : [0, +\infty) \to [0, +\infty)$, then we say that $\mathbf{X}$ has an elliptically contoured distribution with parameters μ and Σ.

In case $\mathbf{X}-\mu$ is Normally distributed with means vector $\mathbf{0}$ and covariance matrix Σ, one has of course $\Psi(s) := \exp\left(-\frac{1}{2}s\right)$.

An elliptically distributed vector $\mathbf{X} - \mu$ can always be represented as

$$\mathbf{X} - \mu = R \times \mathbf{U} \times \Sigma^{\frac{1}{2}} \tag{1.8}$$

where $\Sigma^{\frac{1}{2}}$ is a $J \times J$ matrix such that $\Sigma^{\frac{1}{2}} \times \left(\Sigma^{\frac{1}{2}}\right)^T = \Gamma$, for example its Cholesky decomposition factor; $R \geq 0$ is a non-negative random variable; and $\mathbf{U} \in \mathbb{R}^J$ is a random vector uniformly distributed on the boundary of the unit hypersphere (see Cambanis et al. (1981)). Variable R receives the name of "generating variable", and together with Σ determines the specific characteristics of $\mathbf{X}$, most remarkably, its tail behavior. The generating variable is what really marks the difference among the several elliptical distributions one might construct.

Example 1. In case $\mathbf{X} - \mu$ is Normally distributed with means vector $\mathbf{0}$, then generating variable R is distributed as a χ distribution with J degrees of freedom. That is, $R^2 \sim \chi^2_J$, a chi-squared distribution with J degrees of freedom.

Another well-known case is that of the multivariate student distribution with ν degrees of freedom, for which $R^2 \sim J \times F_{J,\nu}$, and $F_{J,\nu}$ represents the Fisher distribution with J and ν degrees of freedom.

Despite being a generalization to the Normal model, which pervades the Spatial Statistics literature, elliptical models are not part of current practice in the area. For example, among other excellent books on the subject, no mention is made about elliptical distributions at Le and Zidek (2006); Cressie and Wikle (2011); Cressie (1991); Diggle and Ribeiro (2007); Banerjee et al. (2003). This may have to do with the inconvenient the model presents for interpolation or "kriging": For the multivariate Normal and student models, it has already been seen that the distribution of the generating variable depends on the dimension of the vector $\mathbf{X} \in \mathbb{R}^J$, which means that function Ψ must also change. Since our model is defined in terms either of Ψ, as in definition 1, or in terms of generating variable R, as in representation (1.8), it is not clear into what these parameters will turn when extending the model to k "ungauged" sites, whereby $\mathbf{X} \in \mathbb{R}^{J+k}$. This issue is addressed in this paper.

Our intention in dealing with elliptical distributions is to consider interdependence among variables that cannot be quantified in terms of correlations or covariances alone, which concepts form the core of dependence modeling in current spatial statistical practice. The topic of "beyond correlation interdependence" has been addressed in itself by Rodríguez and Bárdossy (2013). We intend here to give an implementation to the ideas presented at Rodríguez and Bárdossy (2013) in the context of Spatial Statistics.

At Rodríguez and Bárdossy (2013), a distinction is drawn between interaction "parameters", and interaction "manifestations". The latter are subject-matter specific statistics connected with sub-vectors of the analyzed random vector, $\mathbf{X}$, and dependent on the type of association among the components of such sub-vectors; they have a relevant interpretation for the researcher. Interaction "parameters" can be seen as convenient building blocks of a (low dimensional) model or dependence structure that can somehow reproduce the target interaction manifestations. It is argued that the joint cumulants of $\mathbf{X}$ are legitimate extensions to correlation coefficients to more than two variables, and as the building blocks referred to above. The cumulant generating function is then accordingly referred to as a "dependence structure".

In the present paper, we show how we can build a low dimensional (regarding the number of parameters to fit) spatial model on the basis of joint cumulants, i.e. on the basis of a cumulant generating function. This model can be considered a natural extension to the Normal model. A Normal model is built on the order one and two joint cumulants only, namely a means vector μ containing the order one cumulants, and an array of covariances $cov\left(X_j, X_i\right)$ containing the order two joint cumulants. In the extension here presented, higher order joint cumulants can be considered without increasing prohibitively the number of parameters to fit.

2. The proposed model

We shall be assuming the existence of sufficiently many product moments of $\mathbf{X}$; sufficient so as to provide a practically useful approximation to the processes modeled. Then it is more convenient, for our purposes, to conceptualize elliptical distributions in terms of their moment generating function.

We say that random vector $\mathbf{X} \in \mathbb{R}^J$ is "elliptically distributed" if and only if its moment generating function can be written as

$$M_{\mathbf{X}-\mu}\left(\mathbf{t}\right) = E\left(e^{\langle \mathbf{t}, \mathbf{X}-\mu \rangle}\right) = \Upsilon\left(\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right) \tag{2.1}$$

for some function $\Upsilon : \mathbb{R} \to \mathbb{R}$, and some $\mu \in \mathbb{R}^J$. For the sake of simplicity, we assume that $\mu = \mathbf{0}$.

Consider a moment generating function of the form

$$M_{\mathbf{X}}\left(\mathbf{t}\right) = \exp\left(\delta\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right)\right) \tag{2.2}$$

for some function $\delta : \mathbb{R} \to \mathbb{R}$. Then the cumulant generating function (c.g.f.) of $\mathbf{X}$ is given by

$$K_{\mathbf{X}}\left(\mathbf{t}\right) := \log\left(M_{\mathbf{X}}\left(\mathbf{t}\right)\right) = \delta\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right) \tag{2.3}$$

This function $\delta\left(y\right)$ can be formally expanded in its Taylor Series around zero,

$$\begin{aligned} \delta\left(y\right) &= c_0 + \frac{c_1}{1!}y + \frac{c_2}{2!}y^2 + \frac{c_3}{3!}y^3 + \frac{c_4}{4!}y^4 + \dots \\ &= c_0 + \frac{c_1}{1!}\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right) + \frac{c_2}{2!}\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right)^2 + \frac{c_3}{3!}\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right)^3 + \dots \\ &= \delta\left(\left(\frac{1}{2}\mathbf{t}^{\mathbf{T}}\Sigma\mathbf{t}\right)\right) \end{aligned} \tag{2.4}$$

where $c_r = \frac{d^r}{dy^r}\delta(y)\mid_{y=0}$.

A little thought shows that the assumption $\mu = \mathbf{0}$ implies that $c_0 = 0$. Thus, by virtue of (2.3) and (2.4) combined, we have that the c.g.f can be written as

$$K_{\mathbf{X}}(\mathbf{t}) = c_1\frac{1}{2}\mathbf{t}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{t} + \frac{1}{2!}c_2\left[\frac{1}{2}\mathbf{t}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{t}\right]^2 + \frac{1}{3!}c_3\left[\frac{1}{2}\mathbf{t}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{t}\right]^3 + \dots \tag{2.5}$$

This c.g.f. was studied by Steyn (1993), in his attempt to introduce more flexibility into the elliptical distributions family. Our proposed model for spatial statistics is given by expansion (2.5), up to an (application specific) order $K \in \{1, 2, 3, 4, \dots\}$. That is, our proposed model is given by a covariance/correlation matrix, $\Gamma_{J\times J}$, together with a set of coefficients $c_1, c_2, \dots, c_K$. Coefficients corresponding to a higher order, $c_{r>K}$, are left undetermined but will be automatically fitted in the presence of data, by means of the implementation of the model given at section (3). Such an implementation circumvents the inconvenience of a model introduced in terms of a c.g.f., by dealing with the density function instead.

From the definition of our model (2.5), some remarks are immediately in place and are given below.

The introduced model as extension to the Normal model. Firstly, the c.g.f. (2.5) boils down to that of the Normal distribution by setting $c_r := 0$, for $r > 1$. Hence the proposed model can be seen as a natural extension to the Normal model which, under $\mu = \mathbf{0}$, is entirely determined by its covariance coefficients

$$\frac{\partial^2 K_{\mathbf{X}}(\mathbf{t})}{\partial t_2 \partial t_1}\mid_{\mathbf{t}=\mathbf{0}} = c_1\Sigma_{ij} = cov(X_i, X_j) \tag{2.6}$$

From (2.6), the need to assume either c_1 fixed, or Σ a correlation matrix, becomes evident: otherwise it will be impossible to identify them separately. In this research, we define Σ to be a covariance matrix, whereas $c_1 = 1$, unless otherwise stated.

Joint cumulants and product moments. Secondly, the joint cumulants of a random vector having a c.g.f as in (2.5) are readily found by differentiating $K_{\mathbf{X}}(\mathbf{t})$ with respect to the indexes of the joint cumulant, and evaluating the result at $\mathbf{t} = \mathbf{0}$. Rodríguez and Bárdossy (2013) show why it is reasonable to call joint cumulants multivariate "interaction parameters".

All joint cumulants of odd order, $\kappa^{j_1,\dots,j_k}$ (k odd), are zero for our dependence model. For k an even integer, joint cumulants are given by:

$$\begin{aligned}
\kappa^{j_1,j_2} &= \frac{c_1}{2}\left\{\Sigma_{j_1j_2} + \Sigma_{j_2j_1}\right\} \\
\kappa^{j_1,j_2,j_3,j_4} &= \frac{c_2}{2!}\frac{1}{2^2}\left\{\Sigma_{j_1j_2}\Sigma_{j_3j_4} + \Sigma_{j_1j_3}\Sigma_{j_2j_4} + \Sigma_{j_1j_4}\Sigma_{j_2j_3}\right\} \\
&\vdots \\
\kappa^{j_1,\dots,j_{2r}} &= \frac{c_{\frac{r}{2}}}{\frac{r}{2}!}\frac{1}{2^{\frac{r}{2}}}\left\{\sum_{j_1,\dots,j_r=1}^{J}\Gamma_{j_1j_2}\dots\Gamma_{j_{r-1}j_r}\right\}
\end{aligned} \tag{2.7}$$

and so on, as shown in appendix A. In this manner, interaction among sets of four, six or more variables can be conveniently summarized.

It will be convenient to define "covariance interdependence factor" $\varrho(j_1,\ldots,j_k)$ as the sum of the products of the covariance coefficients at (2.7). Specifically,

$$\begin{aligned}\varrho(j_1,j_2) &= \Sigma_{j_1j_2}\\ \varrho(j_1,\ldots,j_4) &= \Sigma_{j_1j_2}\Sigma_{j_3j_4}+\Sigma_{j_1j_3}\Sigma_{j_2j_4}+\Sigma_{j_1j_4}\Sigma_{j_2j_3}\\ \varrho(j_1,\ldots,j_6) &= \Sigma_{j_1j_2}\Sigma_{j_3j_4}\Sigma_{j_5j_6}+\Sigma_{j_1j_3}\Sigma_{j_2j_4}\Sigma_{j_5j_6}+\ldots+\Sigma_{j_1j_6}\Sigma_{j_2j_4}\Sigma_{j_5j_3}\end{aligned}$$

and so on. This is a "potential" interdependence factor, since its effect on higher order interdependence parameters (i.e. joint cumulants of order greater than 2), is only present if its corresponding coefficient $c_{k/2}$ is non-zero. So, every joint cumulant at (2.7) can be written as

$$\kappa^{j_1,\ldots,j_k} = c_{\frac{k}{2}} \times \varrho(j_1,\ldots,j_k) \tag{2.8}$$

Our interdependence parameter (i.e. joint cumulant) of order $k>2$ can then be conceptually split into two components: On the one hand, a "covariance interdependence component", $\varrho(j_1,\ldots,j_k)$, that can be estimated low-dimensionally via covariance function fitting, as usual in Geostatistics. On the other hand, an interdependence "enhancing" parameter $c_{k/2}$, whose departure from zero determines the departure from zero of the k-th order joint cumulant. As illustrated in Rodríguez and Bárdossy (2013), these joint cumulants can be connected with relevant interaction manifestations. As a consequence, one can can try fitting the problem-specific interaction manifestation, which are not explainable in terms of correlations, by fitting parameters $c_2, c_3, \ldots$

An expansion for the moment generating function (m.g.f.) for $\mathbf{X}$ will be now introduced. By setting shorthand notation

$$y := \frac{1}{2}\mathbf{t}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{t}$$

the dependence structure (2.5) can be written

$$K_{\mathbf{X}}(\mathbf{t}) = \frac{c_1}{1!}y + \frac{c_2}{2!}y^2 + \frac{c_3}{3!}y^3 + \ldots \tag{2.9}$$

On the other hand, the definition of our dependence structure, given originally by (2.2) implies that we can write, using the same shorthand notation as above,

$$\exp\left(K_{\mathbf{X}}(\mathbf{t})\right) := M_{\mathbf{X}}(\mathbf{t}) = \exp\left(\delta(y)\right) = 1 + \frac{m_1}{1!}y + \frac{m_2}{2!}y^2 + \frac{m_3}{3!}y^3 + \ldots \tag{2.10}$$

for certain coefficients $m_1, m_2, m_3, \ldots$, at least for y in a neighborhood of zero (that is, for $\mathbf{t}$ in a sufficiently small neighborhood of $\mathbf{0}$). Summarizing, we have that

$$\log\left(1 + \frac{m_1}{1!}y + \frac{m_2}{2!}y^2 + \frac{m_3}{3!}y^3 + \ldots\right) = \frac{c_1}{1!}y + \frac{c_2}{2!}y^2 + \frac{c_3}{3!}y^3 + \ldots \tag{2.11}$$

and then we can obtain, as in the case of the one-dimensional cumulants in terms of the one-dimensional moments (see, e.g. Kendall and Stuart (1969); Smith (1995)), coefficients $m_1, m_2, m_3, \ldots$ in terms of $c_1, c_2, c_3, \ldots$, by

$$
\begin{aligned}
c_1 &= m_1 \\
c_2 &= m_2 - m_1^2 \\
c_3 &= m_3 - 3m_2m_1 + 2m_1^3 \\
c_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4
\end{aligned} \tag{2.12}
$$

which after some algebraic manipulation, returns,

$$
\begin{aligned}
m_1 &= c_1 \\
m_2 &= c_2 + c_1^2 \\
m_3 &= c_3 + 3c_2c_1 + c_1^3 \\
m_4 &= c_4 + 4c_3c_1 + 3c_2^2 + 6c_2c_1^2 + c_1^4
\end{aligned} \tag{2.13}
$$

So, we have shown that the moment generating function at (2.2) can be written as

$$
M_{\mathbf{X}}(\mathbf{t}) = 1 + \frac{m_1}{1!}\left(\frac{1}{2}\mathbf{t^T\Sigma t}\right) + \frac{m_2}{2!}\left(\frac{1}{2}\mathbf{t^T\Sigma t}\right)^2 + \dots \tag{2.14}
$$

which is similar to the expansion of $K_{\mathbf{X}}(\mathbf{t})$, except for the leading term 1 and coefficients m_r, $r = 1, 2, \dots$. We express product moments analogously as joint cumulants by

$$
\mu^{j_1,\dots,j_k} := E\left(X_{j_1} \dots X_{j_k}\right) \tag{2.15}
$$

where $j_r \in \{1, \dots J\}$, $r = 1, \dots k$, allowing repetition of indexes. Then it follows, analogously to (2.7), that

$$
\mu^{j_1,j_2} = \frac{m_1}{2}\left\{\Sigma_{j_1j_2} + \Sigma_{j_2j_1}\right\} \tag{2.16}
$$

$$
\mu^{j_1,j_2,j_3,j_4} = \frac{m_2}{2!}\frac{1}{2^2}\left\{\Sigma_{j_1j_2}\Sigma_{j_3j_4} + \Sigma_{j_1j_3}\Sigma_{j_2j_4} + \Sigma_{j_1j_4}\Sigma_{j_2j_3}\right\} \tag{2.17}
$$

$$
\vdots \tag{2.18}
$$

$$
\mu^{j_1,\dots,j_{2r}} = \frac{m_{\frac{r}{2}}}{\frac{r}{2}!}\frac{1}{2^{\frac{r}{2}}}\left\{\sum_{j_1,\dots,j_r=1}^{J}\Gamma_{j_1j_2}\dots\Gamma_{j_{r-1}j_r}\right\} \tag{2.19}
$$

where m_r is as in (2.13).

We see then, for example by setting $c_1 = 1$ and $c_{r>1} = 0$, that we can have non-zero joint moments of orders greater than two, even though no dependence of order greater than two is present in the distribution of $\mathbf{X}$, according to our definition of high order dependence.

The proposed c.g.f. as an extension to the covariance function. Covariance functions, such as (1.5) or (1.6) have proved valuable tools for spatial statistics analysis. They define the order-two joint cumulant of every pair of components, e.g.,

$$
\begin{aligned}
C\left(d_{ij} \mid \left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)\right) &= \sigma_0^2.I(d=0) + \sigma_1^2 \exp\left(-\left(d/\theta_1\right)^{\theta_2}\right) \\
&= cov\left(X_i, X_j\right) = \frac{\partial^2 K_{\mathbf{X}}(\mathbf{t})}{\partial t_2 \partial t_1}\,|_{\mathbf{t=0}}
\end{aligned} \tag{2.20}
$$

where $d_{ij} \in [0, +\infty)$ denotes the distance between the sites, $\mathbf{s_i}$ and $\mathbf{s_j}$, to which X_i and X_j correspond.

Let $D = \{d_{ij}\}$ be the matrix of distances between the sites corresponding to the different components of $\mathbf{X}$. Then one has the matrix equality $\{\Sigma_{ij}\} = \left\{C\left(d_{ij} \mid \left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)\right)\right\}$. With slight abuse of notation, denote

$$C\left(D \mid \left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)\right) := \Sigma$$

The c.g.f. (2.5) can be written as a "higher order" spatial covariance function as

$$K_{\mathbf{X}}(\mathbf{t}) = c_1 \frac{1}{2}\mathbf{t}C(D)\,\mathbf{t} + \frac{1}{2!}c_2\left[\frac{1}{2}\mathbf{t}C(D)\,\mathbf{t}\right]^2 + \frac{1}{3!}c_3\left[\frac{1}{2}\mathbf{t}C(D)\,\mathbf{t}\right]^3 + \ldots$$

where the dependence on parameters $\left(\theta_1, \theta_2, \sigma_0^2, \sigma_1^2\right)$ have been obviated to avoid cumbersome notation. This higher order covariance function allows us to represent covariances in terms of the distance separating the two sites in question, and higher order (>2) joint cumulants in terms of distances among the sites involved *and* the coefficients $c_{r>2}$.

"Orthogonality" in joint cumulants. Joint cumulants of higher order do not affect lower ordered ones, as follows from inspecting (2.7). After fixing Σ, each r-ordered joint cumulant depends on a separate coefficient, $c_{\frac{r}{2}}$. Hence, one can have similar joint cumulants up to a given order K, but then different coefficients $c_{r>\frac{K}{2}}$ will lead to different joint cumulants of higher order. This results in different association types that may go unnoticed in the analysis of low dimensional marginal distributions, such as 1 and 2-dimensional ones. This topic is explored in section (5), where two random fields are equal in terms of their one and second order joint cumulants (i.e. mean and covaraince structure), and in terms of their one and two dimensional marginal distributions, but still exhibit very different clustering behaviors.

3. Model Implementation in the context of Spatial Statistics

3.1. **Spatial Consistency.** An important issue when dealing with a probability distribution for Spatial Data is that this distribution should be "consistent". If we denote by $\mathbf{X} \in \mathbb{R}^J$ our modeling vector, consistency means that any subvector $(X_{j_1}, \ldots, X_{j_K}) \in \mathbb{R}^K$, of $\mathbf{X}$, with $K \leq J$, will have the same type of distribution distribution as $\mathbf{X}$.

In order to be more specific, consider elliptically distributed vector $(X_1, \ldots, X_J) \in \mathbb{R}^J$ having a density function. This density function can be written as

$$\{f((X_1, \ldots, X_J) \mid J) \mid J \in \mathbb{N}\} \tag{3.1}$$

where dependence on dimension J has been made explicit. Kano (1994) has given a definition that can be stated as follows: The family at (3.1) possesses the consistency property if and only if

$$\int_{-\infty}^{+\infty} f((x_1, \ldots, x_{J+1}) \mid J+1)\, dx_{J+1} = f((x_1, \ldots, x_J) \mid J) \tag{3.2}$$

for any $J \in \mathbb{N}$ and almost all $(x_1, \ldots, x_J) \in \mathbb{R}^J$. We also say that such a family is consistent.

As Kano (1994) notes, not all members of the elliptical family are consistent. He gives a necessary and sufficient condition for a family such as (3.1) to be consistent.

The family is consistent if and only if, for each $J \in \mathbb{N}$, random vector $\mathbf{X} \in J$ can be stochastically written as

$$\mathbf{X} = \sqrt{V} \times \mathbf{Z} \tag{3.3}$$

where $\mathbf{Z} \sim N_J(\mathbf{0}, \Sigma)$ stands for a normally distributed vector with the same covariance matrix as $\mathbf{X}$, and $V > 0$ is a univariate random variable independent of $\mathbf{Z}$ and *unrelated* to J.

As Kano (1994) reminds us, the construction at (3.3) produces distributions with tails at least as heavy as the Normal distribution, whereby Normal tail dependence (i.e. "zero" tail dependence) can only be achieved for the case where V is a positive constant.

3.2. **Relation between R^2 and coefficients $c_1, c_2, c_3, \ldots$.** In appendix B, it is shown that if $\mathbf{X} \in \mathbb{R}^J$ has a c.g.f. as in (2.5), and consequently a m.g.f. as in (2.2), then the following relation between the k-th order moments of its squared generating variable, R^2, and coefficients $m_1, m_2, m_3, \ldots$ exists:

$$E\left(\left(R^2\right)^k\right) = \frac{m_k}{c_1^k} \frac{2^k \Gamma\left(k + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)} \tag{3.4}$$

The expression is conveniently expressed in terms of $m_1, m_2, m_3, \ldots$, but it can be written in terms of the c_r coefficients by virtue of (2.13),

$$E\left(\left(R^2\right)^1\right) = \frac{c_1}{c_1} \frac{2^1 \Gamma\left(1 + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)} \tag{3.5}$$

$$E\left(\left(R^2\right)^2\right) = \frac{\left(c_2 + c_1^2\right)}{c_1^2} \frac{2^2 \Gamma\left(2 + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)} \tag{3.6}$$

$$E\left(\left(R^2\right)^3\right) = \frac{\left(c_3 + 3c_2c_1 + c_1^3\right)}{c_1^3} \frac{2^3 \Gamma\left(3 + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)} \tag{3.7}$$

$$\vdots \quad \vdots \quad \vdots \tag{3.8}$$

and still further simplified by substituting 1 for c_1.

If we consider (1.8) and example 1, then construction (3.3) indicates that the generating variable of $\mathbf{X}$ can be represented as follows :

$$R \underset{d}{=} \sqrt{V \times \chi_J^2} \tag{3.9}$$

and hence

$$R^2 \underset{d}{=} V \times \chi_J^2 \tag{3.10}$$

where V and χ_J^2 are independent (see item *iii* at theorem 1 of Kano (1994)). Due to this independence,

$$E\left(\left(R^2\right)^k\right) = E(V) \times E\left(\chi_J^2\right) \tag{3.11}$$

Note that the moments of χ_J^2 are given by

$$E\left(\left(\chi_J^2\right)^k\right) = \frac{2^k \Gamma\left(k + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)}$$

Equation (3.4) then means that the moments of V are given by $m_1 = 1, m_2, m_3, \ldots$, whereas its cumulants are given by $c_1 = 1, c_2, c_3, \ldots$. We have then identified a sufficient condition under which both the expression at (2.5) is a legitimate cumulant generating function *and* it produces a consistent model, useful for spatial statistics: the coefficients $c_1 = 1, c_2, c_3, \ldots$ must be the cumulants of some random variable, $V > 0$, whereas $m_1 = 1, m_2, m_3, \ldots$ must be its moments.

Remark. Note that a scaling variable $V > 0$ having a very small variance, c_2, will produce a random field very similar to a Gaussian random field in its one and two dimensional marginal distributions (which is all that current Geostatistical techniques fit and check for goodness of fit). This is because the (common) kurtosis of each marginal distribution, given by

$$\frac{\kappa_4(X_j)}{Var(X_j)^2} = \frac{\kappa^{j,j,j,j}}{Var(X_j)^2} = \frac{3c_2}{8Var(X_j)}$$

will be very close to zero, as in the Normal model. But if coefficients $c_{r>2}$ are relatively big, then (2.7) indicates that the joint cumulants of higher order, involving the interaction of 4, 6 and more components of $\mathbf{X}$, will be considerably altered. As the dimension of the field increases, important characteristics of the field constructed via (3.3) will be totally different from those of the Gaussian field (see example below), though these differences cannot be noticed from the one and two dimensional marginals. Additionally, conditional distributions (i.e. at "ungauged sites") will also be different, as the number of conditioning values increases. This has important consequences for interpolation.

3.3. **Parameter Estimation.** Apart from the estimation of covariance matrix Σ, estimation of the model defined by (2.5) amounts to estimating coefficients $c_2, c_3, \ldots$, or equivalently, coefficients $m_2, m_3, \ldots$.

If we assign a flexible model to (squared) scaling variable V, such as a mixture of gamma distributions,

$$f_V(V) = \sum_{s=1}^{S} \pi_s \frac{\beta_s^{-\alpha_s}}{\Gamma(\alpha_s)} V^{\alpha_s - 1} e^{-\frac{V}{\beta_s}} \tag{3.12}$$

then parameter estimation for our model can be effected as follows:

(1) First, we estimate covariance matrix Σ, for which we may use standard covariance function models, such as (1.5) or (1.6). We can do this in a first, independent step, because of the "orthogonality" property of the joint cumulants of $\mathbf{X}$ referred to in section (2) remarks.
(2) Second, we fit the density of V conditioned on $E(V) = c_1 = m_1 = 1$, thus fitting the parameters present at density function (3.12). One must impose some restrictions on these latter parameters, in order to avoid lack of identifiability; we impose at the example below that weights $\pi_1, \ldots, \pi_{S-1}$ must be in decreasing order, whereas $\pi_S := 1 - \sum_{s=1}^{S-1} \pi_s$.

The parameters estimation at step 2 is effected by estimating computing estimators $\hat{m}_2, \hat{m}_3, \ldots$ and then finding $\hat{\pi}_1, \ldots, \hat{\pi}_{S-1}, \hat{\beta}_1, \ldots, \hat{\alpha}_S$, such that

$$\hat{m}_k \approx \sum_{s=1}^{S} \hat{\pi}_s \frac{\hat{\beta}_s^k \Gamma(\hat{\alpha}_s + k)}{\Gamma(\hat{\alpha}_s)} \tag{3.13}$$

for $k = 2, 3, \ldots$, where the "hat" symbol can be read as "estimator of" the parameter it covers. This is an instance of the method of moments.

Note also that estimation at step 2 above does not alter in any manner the already estimated covariance matrix, containing the joint cumulants of order two. Step 2 is concerned with estimating coefficients, $c_2, c_3, c_4, \ldots$, affecting joint cumulants of higher orders, only.

The estimation technique will be now explained in more detail.

Assume one has a sample $\mathbf{x}_1, \ldots, \mathbf{x}_I$ of $\mathbf{X} \in \mathbb{R}^J$. This sample might represent, for example, I observations of the (spatially associated) residual process obtained by applying a daily time series model to each of J precipitation gauging stations spread over sites with coordinates $\mathbf{s_1}, \ldots, \mathbf{s_J}$, $\mathbf{s_j} \in \mathbb{R}^2$. The fact that precipitation demands a truncated model will be ignored for now, since this issue will be briefly considered in section (6). Begin by standardizing data, so that each component has mean zero.

Perform the following estimating steps:

(1) Apply any transformation to data that might be necessary (cf. Sanso and Guenni (1999)), in order to make data approximately Gaussian in its one-dimensional marginals.

(2) Fit a multivariate Normal model to $\mathbf{X}$, on the basis of $\mathbf{x_1}, \ldots, \mathbf{x_I}$. A standard covariance model, such as (1.5), can be used to estimate covariance structure of $\mathbf{X}$. The covariance between every two components of $\mathbf{X}$ referred to locations $\mathbf{s_{j_1}}$ and $\mathbf{s_{j_2}}$, are then estimated as a function of the distance between the locations by $\hat{\Sigma}_{j_1 j_2} = C\left(\|\mathbf{s_{j_1}} - \mathbf{s_{j_2}}\| \mid \hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_0^2, \hat{\sigma}_1^2\right)$.

(3) Compute $r_i^2 = \mathbf{x_i}\hat{\Sigma}^{-1}\mathbf{x_i}^T$, for $i = 1, \ldots, I$. These are approximate samples of R^2, the squared generating variable of $\mathbf{X}$, as can be seen by an argument similar to that presented in appendix B.

(4) Compute $\hat{\vartheta}_k = \frac{1}{I}\sum_{i=1}^{I}\left(r_i^2\right)^k$, the estimates of the moments of squared generating variable R^2, up to a prudent order, say $K = 5$.

(5) By virtue of (3.4) and remembering that $c_1 = m_1 = 1$, one has estimates for m_k, for $k = 2, \ldots, K$, given by

$$\hat{m}_k = \frac{\Gamma\left(\frac{J}{2}\right)}{2^k \Gamma\left(k + \frac{J}{2}\right)}\hat{\vartheta}_k \tag{3.14}$$

(6) Apply the method of moments to estimate the parameters of the density of scaling variable V. That is, solve the following minimization problem:

$$\min_{\vec{\alpha},\vec{\beta},\vec{\pi}_{-S\cdot}} \sum_{k=2}^{K}\left(\hat{m}_k - \sum_{s=1}^{S}\pi_s \frac{\beta_s^k \Gamma\left(\alpha_s + k\right)}{\Gamma\left(\alpha_s\right)}\right)^2 \tag{3.15}$$

subject to

$$\begin{aligned} \sum_{s=1}^{S}\pi_s \frac{\beta_s \Gamma\left(\alpha_s + 1\right)}{\Gamma\left(\alpha_s\right)} &= m_1 = 1 \\ \pi_s &\geq \pi_{s+1} \geq 0 \\ \sum_{s=1}^{S-1}\pi_s &\leq 1 \end{aligned}$$

where $\vec{\pi}_{-S} = (\pi_1, \ldots, \pi_{S-1})$, $\pi_S = 1 - \sum_{s=1}^{S} \pi_s$, and the inequalities at the second constraint are valid for $1 \leq s \leq S-2$.

For step 6 above, the Lagrange multipliers approach can be employed.

As output to the procedure outlined by steps 1 through 6, one has an estimation of the covariance model, and of the distribution of the squared scaling variable, V. With these, simulation and interpolation can be performed, as explained subsequently.

Remark. The representation of the density of scaling variable V as a mixture of gamma distributions, indicates that the model here presented can approximate a wide spectrum of tail dependence association, which includes that of the Normal and the Student-t distribution.

4. Simulation, interpolation and inference of field characteristics

4.1. **Random Simulation.** The decomposition (3.3) can be conveniently used both for simulation and for interpolation.

In order to simulate a realization of vector $\mathbf{X} \in \mathbb{R}^J$:

(1) Sample a realization $\mathbf{z}_i$ from a multivariate Normal distribution with means vector $\mathbf{0} \in \mathbb{R}^J$ and covariance matrix $\hat{\Sigma}$.
(2) Sample a realization v_i of V. To this end sample an index, $s \in \mathbb{N}$, from a multinomial distribution with class probabilities $(\hat{\pi}_1, \ldots, \hat{\pi}_S)$ and then sample $v_i \sim Gamma\left(\hat{\alpha}_s, \hat{\beta}_s\right)$.
(3) The realization of $\mathbf{X}$ is given by $\mathbf{x}_i := \sqrt{v_i} \times \mathbf{z}_i$. Add a means vector, $\mu \in \mathbb{R}^J$, to $\mathbf{x}_i$, if necessary.

Note that a field of dimension $J^* \neq J$ can be simulated in the same manner, since the distribution of V does not depend on J. Hence, one can simulate a big random field by obtaining (approximately) a realization of a Gaussian random field using some fast generation mechanism, such as turning bands (see, for example, Ripley (1981)), and then multiplying it by a realization of $\sqrt{V}$. This is done for section (5), and the consequences on some manifestations of interaction, as compared to the original Gaussian field, are there illustrated.

4.2. **Interpolation via the saddlepoint approximation.** The distribution of the environmental variable of interest at a new location can be described with little inconvenience. This is because we are building on the idea of the covariance function, and so we can extend the covariance matrix Σ to include the covariance between the variable of interest at any gauged site and any new location. Denote by j_1 any generic component of $\mathbf{X}$. The correlation matrix components corresponding to site $\mathbf{s}_{j^*}$ are given by

$$\Sigma_{j^* j_1} = C\left(\|\mathbf{s}_{j_1} - \mathbf{s}_{j^*}\| \mid \hat{\theta}_1, \hat{\theta}_2, \hat{\sigma}_0^2, \hat{\sigma}_1^2\right) \tag{4.1}$$

For the subsequent discussion, we shall denote the extended covariance matrix by $\Sigma^* \in \mathbb{R}^{J+1 \times J+1}$.

Suppose that the distribution of the variable is desired for a new location with coordinates $\mathbf{s}_{j^*} \in \mathbb{R}^2$, given that one has observed a realization $\mathbf{x}$ of $\mathbf{X}$ at sites $\mathbf{s}_1, \ldots, \mathbf{s}_J$. We present now a method for obtaining the approximate distribution of $X_{j^*} = X\left(\mathbf{s}_{j^*}\right)$ given $\mathbf{x}$.

Since the distribution of $\mathbf{X}$ is consistent, the cumulant generating function of $\mathbf{Y} := (X_{j^*}, \mathbf{X})$ is of the same form as that of $\mathbf{X}$ (cf. Kano (1994)). This new c.g.f. can then be written as,

$$K_{\mathbf{Y}}(\mathbf{w}) = 1\frac{1}{2}\mathbf{w}^{\mathbf{T}}\boldsymbol{\Sigma}^{*}\mathbf{w} + \frac{1}{2!}\hat{c}_2\left[\frac{1}{2}\mathbf{w}^{\mathbf{T}}\boldsymbol{\Sigma}^{*}\mathbf{w}\right]^2 + \frac{1}{3!}\hat{c}_3\left[\frac{1}{2}\mathbf{w}^{\mathbf{T}}\boldsymbol{\Sigma}^{*}\mathbf{w}\right]^3 + \dots \tag{4.2}$$

where $\mathbf{w} \in \mathbb{R}^{J+1}$. As shown by Skovgaard (1987) (see also: Kolassa (2006); Barndorff-Nielsen and Cox (1990)), we have that:

$$\Pr(X_{j^*} \leq a \mid \mathbf{X} = \mathbf{x}) \approx \Phi(r) + \phi(r)\left(\frac{1}{r} - q\right) \tag{4.3}$$

where

$$r = sign(\hat{\mathbf{w}}_1)\sqrt{2\left\{\hat{\mathbf{w}}^T(a, \mathbf{x}) - \hat{\mathbf{w}}_{-1}^T\mathbf{x} - K_{\mathbf{Y}}(\hat{\mathbf{w}}) + K_{\mathbf{X}}(\hat{\mathbf{w}}_{-1})\right\}} \tag{4.4}$$

$$q = \frac{1}{\hat{\mathbf{w}}_1}\det\left(K_{\mathbf{X}}^{''}(\hat{\mathbf{w}}_{-1})\right)\det\left(K_{\mathbf{Y}}^{''}(\hat{\mathbf{w}})\right)^{-\frac{1}{2}} \tag{4.5}$$

and $\hat{\mathbf{w}} \in \mathbb{R}^{J+1}$, $\hat{\mathbf{w}}_{-1} \in \mathbb{R}^{J}$ are the solutions to equations

$$\begin{aligned} \nabla K_{\mathbf{Y}}(\hat{\mathbf{w}}) &= (a, \mathbf{x}) \\ \nabla K_{\mathbf{X}}(\hat{\mathbf{w}}_{-1}) &= \mathbf{x} \end{aligned}$$

Additionally, $\hat{\mathbf{w}}_1$ is the first component of $\hat{\mathbf{w}}$, and $K_{\mathbf{X}}^{''}(\hat{\mathbf{w}})$ stands for the matrix of second derivatives on the c.g.f. evaluated at $\hat{\mathbf{w}}$.

We can apply this approximation to $\Pr(X_{j^*} \leq a \mid \mathbf{X} = \mathbf{x})$ directly, using the extended c.g.f. given by (4.2). This is done in section (5.8) below.

In case one wishes the distribution of the environmental variable at several new locations $\mathbf{s}_{j_1^*}, \dots, \mathbf{s}_{j_K^*}$ simultaneously,

$$\Pr\left(X_{j_1^*} \leq a_1, \dots, X_{j_K^*} \leq a_K \mid \mathbf{X} = \mathbf{x}\right)$$

one can apply the the extension to this approach presented by Kolassa and Li (2010). A conceptually easier approach would be to run the Gibbs sampler repeatedly, using (4.3) to sample from each (approximate) full conditional distribution (see Kolassa and Tanner (1994)). After sufficiently many iterations, the samples obtained can be considered approximate realizations of the conditional distribution desired.

5. A simulation-based illustration

In this section, we present a simulation study of the type of interdependence that can be generated using a model having c.f.g. as (2.5). The study is built so as to mimic the model building process in Spatial Statistics: from data obtained at a limited number of locations ("gauging stations"), we intend to infer a model for the interesting variable over the whole region to which these locations belong.

It will be noted, that the additional characteristics the field possesses can be unnoticeable from the one and two dimensional marginal distributions. In this example, they are indistinguishable from those of a Gaussian field. However, specific characteristics of the underlying field, which are relevant for applications such as rainfall modeling and mining geostatistics, will be considerably different.

5.1. **Scaling variable used and simulated fields employed.** We generated 3650 realizations of a $J^* = 300 \times 300$ Gaussian field, using the circular embedding method as implemented in package *RandomFields* of the statistical software R. The specification of the field is: $\mu = 0$, $\theta_1 = 20$, $\sigma_0^2 = 0$ and $\sigma_1^2 = 1$, where μ, θ_1, σ_0^2 and σ_1^2 denote the field mean, the range parameter, the nugget effect and the field variance, respectively. Additionally, the covariance function model was the exponential model, given by setting $\theta_2 = 0$ and using equation (1.5).

In order to apply (3.3) we simulated 3650 realizations of a mixture of 5 Gamma distributions, as in (3.12), with the following parameters, rounded up to the fourth decimal place:

Mixture Weights: $\vec{\pi} = (0.7137, 0.1697, 0.1094, < 0.0000, < 0.0000)$
Shape Parameters: $(\alpha_1, \ldots, \alpha_5) = (32.5168, 25.0004, 27.4404, 0.3582, 11.3288)$
Scale Parameters: $(\beta_1, \ldots, \beta_5) = (0.0302, 0.0393, 0.0357, 0.6012, 0.2975)$

This amounts to V having moments $(m_1, \ldots, m_5) = (0.9986, 1.0766, 1.3856, 2.6163, 8.0863)$, and cumulants $(c_1, \ldots, c_5) = (0.9986, 0.0795, 0.1519, 0.5210, 1.9712)$. A plot of the density of V, together with a boxplot based on 10000 realizations, is presented at figure (5.1). Note that the field then constructed according to (3.3) may be a multiplication of a Gaussian field by a value of $\sqrt{V} = \sqrt{7} \approx 2.65$. This is the kind of event that makes the so constructed field different dependence properties from a Gaussian field. One has a non-neglectable probability of $V > 4$, which implies that many fields will be multiplied by values $\sqrt{V} > 2$.

The behavior of scaling variable V influences the tail behavior of the resulting vector $\mathbf{X}$. A typical representation of the multivariate Student distribution with correlation matrix Σ and ν degrees of freedom is (see Kotz and Nadarajah (2004)):

$$\mathbf{X} = \sqrt{S}\mathbf{Z}$$

where $\mathbf{Z}$ is a normally distributed vector with vector of means $\mathbf{0}$ and correlation matrix Σ, and

$$S \sim \frac{\nu}{\chi^2_\nu}$$

Hence we can compare the distribution of squared scaling variable V, presented at figure (5.1) with the distribution of a multivariate Student distribution, for various degrees of freedom. The distributions of the squared scaling variables are presented at figure (5.2), for a multivariate Student distribution with $\nu \in \{10, 15, 20, 35\}$ degrees of freedom.

It is noteworthy that scaling variable V seems to have the lightest tail, if you focus on the left hand panel of figure (5.2). However, the uppermost part of the distribution of V is similar to that of $\frac{\nu}{\chi^2_\nu}$ with $\nu = 15$. That is, the tail dependence of our model is actually similar to that of the multivariate Student distribution with $\nu = 15$ degrees of freedom. This fact goes unnoticed in the 1 and 2-dimensinal marginal distributions, as we shall see.

The non-Gaussian fields were generated by taking each realization of the Gaussian field, and multiplying it by a realization of $\sqrt{V}$.

5.2. **Partial observation of the fields: A network of 30 stations.** Let us denote by $\mathbf{Z}^* \in \mathbb{R}^{J^*}$ and $\mathbf{X}^* \in \mathbb{R}^{J^*}$ the random fields generated as a gaussian field, and by multiplication of the latter by $\sqrt{V}$, respectively. In this case, $J^* = 300 \times 300 = 90000$. We selected 30 components of the field, corresponding in a Spatial

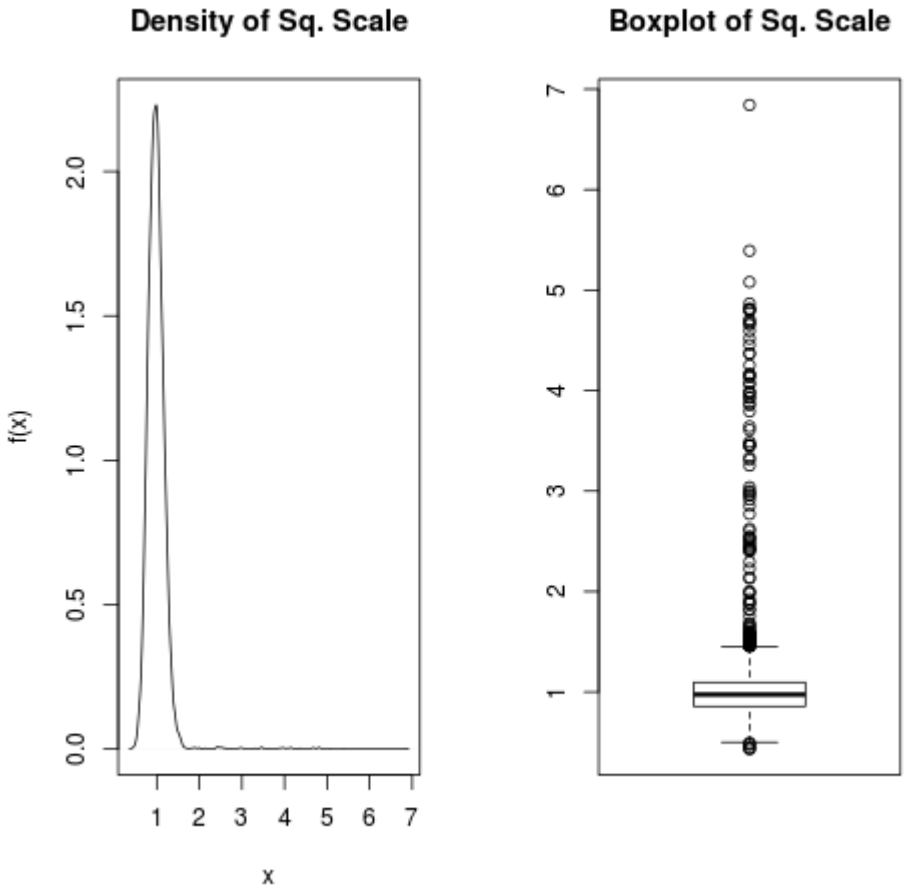

FIGURE 5.1. Density and simulation-based boxplot (n=10000) of the squared scaling variable, $V > 0$, used for the example in this section.

context to 30 locations on the plane, and stored the data of these components. The setting is illustrated in figure (5.3).

The n=3650, 30-dimensional observations thus available from field $\mathbf{Z}^*$ are in the following considered realizations from sub-vector $\mathbf{Z} \in \mathbb{R}^{30}$ of $\mathbf{Z}^*$, whereas those from field $\mathbf{X}^*$ are considered as realizations of sub-vector $\mathbf{X} \in \mathbb{R}^{30}$.

Data from these vectors, $\mathbf{Z}$ and $\mathbf{X}$, represent the data available at a limited number of gauging stations. As usual in Spatial Statistics, we intend to identify characteristics of the whole fields, $\mathbf{Z}^*$ and $\mathbf{X}^*$.

A third vector dealt with in this section is $\mathbf{W} \in \mathbb{R}^{30}$, of which each component is given by

$$W_j = F_{Z_j}^{-1}\left(F_{X_j}\left(X_j\right)\right) \tag{5.1}$$

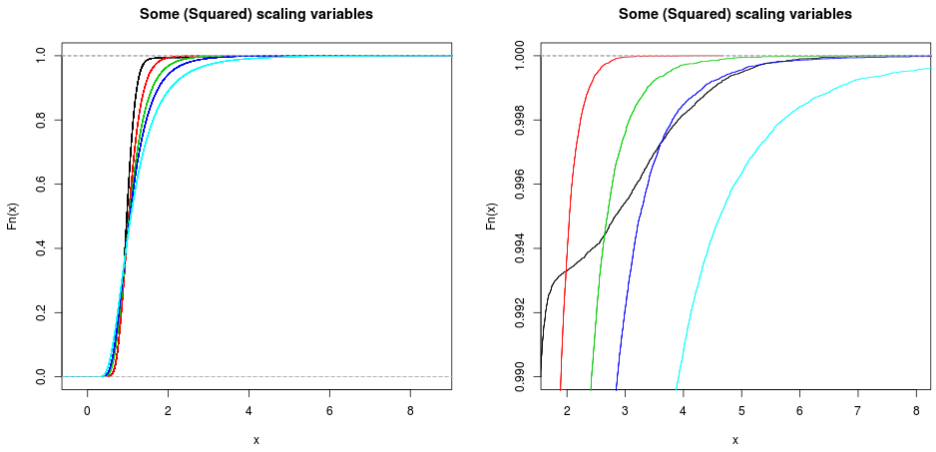

FIGURE 5.2. Comparison of the distribution of squared scaling variable V (black) with distributions of the scaling variable of the multivariate Student distribution for degrees of freedom: 10 (light blue), 15 (dark blue), 20 (green), 35 (red). The uppermost part of the distribution of V produces a tail behaviour similar to that of a multivariate Student distribution with 15 degrees of freedom.

that is, $\mathbf{W}$ is the vector resulting from applying the quantile-quantile transformation to each component of $\mathbf{X}$, mapping these into the quantiles of the components of $\mathbf{Z}$. Hence, each marginal distribution of $\mathbf{W}$ is standard normal, like those of $\mathbf{Z}$, but the joint distribution of its ranks (the copula), is like that of $\mathbf{X}$.

5.3. **Analysis of one dimensional marginal distributions.** Comparison of the 1-dimensional marginal distributions of $\mathbf{Z}$ and $\mathbf{X}$ is performed in this sub-section. At figure (5.4) we present four quantile-quantile plots. Each of these plots corresponds to data from (Z_j, X_j), where j has been randomly selected from $\{1, \ldots, 30\}$. The Anderson-Darling test for equality in distributions was applied to data involved in each plot, and the resulting p-value (n=3650) has been written on each plot title. Both visually and from the testing viewpoint, the marginal distributions considered at each plot seem to be the same.

Additionally, the Anderson-Darling test was applied to data from every pair (Z_j, X_j), for $j = 1, \ldots, 30$, n=3650. The minimum p-value obtained from all 30 tests was 0.642. Hence $\mathbf{X}$ and $\mathbf{Z}$ can be considered to have the same 1-dimensional marginals, namely, standard normal marginal distributions.

5.4. **Analysis of two dimensional marginal distributions.** Data corresponding to two components of both $\mathbf{X}$ and $\mathbf{Z}$, namely 3 and 28, are shown at figure (5.5) for illustration. The multivariate version of Shapiro-Wilks test for normality introduced by Villasenor Alva and Estrada (2009), as implemented in the R package *mvShapiro.Test*, was applied to a randomly selected sample (n=500) of (X_3, X_{28}). This test resulted in a p-value of 0.191, whereby (X_3, X_{28}) can be considered a

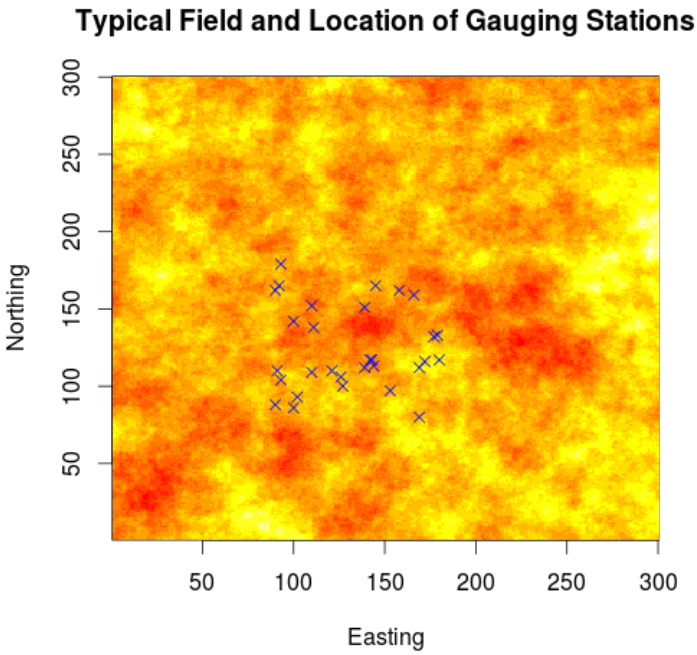

FIGURE 5.3. Typical (non-Gaussian) field of the n=3650 generated, and the 30 locations at which data was recorded to form **X** and **Z**.

α-Level	(Z_{j_1}, Z_{j_2})	(X_{j_1}, X_{j_2})	(W_{j_1}, W_{j_2})
0.01	433 (99.54%)	402 (92.41%)	433 (99.54%)
0.05	419 (96.32%)	360 (82.76%)	421 (96.78%)
0.10	398 (91.49%)	323 (74.25%)	405 (93.10%)

TABLE 5.1. Summary of multivariate Shapiro-Wilks test applied on all bivariate marginal distributions of **Z**, **X** and **W**. A random sub-sample (n=500) from the available data was used for each testing. Out of the total $\binom{30}{2} = 435$ bivariate combinations, the total number (and percentage) of combinations by which the Normality hypothesis cannot be rejected at the respective α-level are shown.

Gaussian 2-dimensional vector[1]. The same test procedure was performed on all $\binom{30}{2} = 435$ pairs of marginals, for **Z**, **X** and **W**. The results are summarized at table (5.1). It can be seen that non-Gaussian vectors, **X** and **W**, exhibit gaussian bivariate marginals most of the time. Results are qualitatively similar to those of **Z**, in particular for **W**.

[1]This procedure was repeated several times, and some of the p-values obtained were rightly under 0.05.

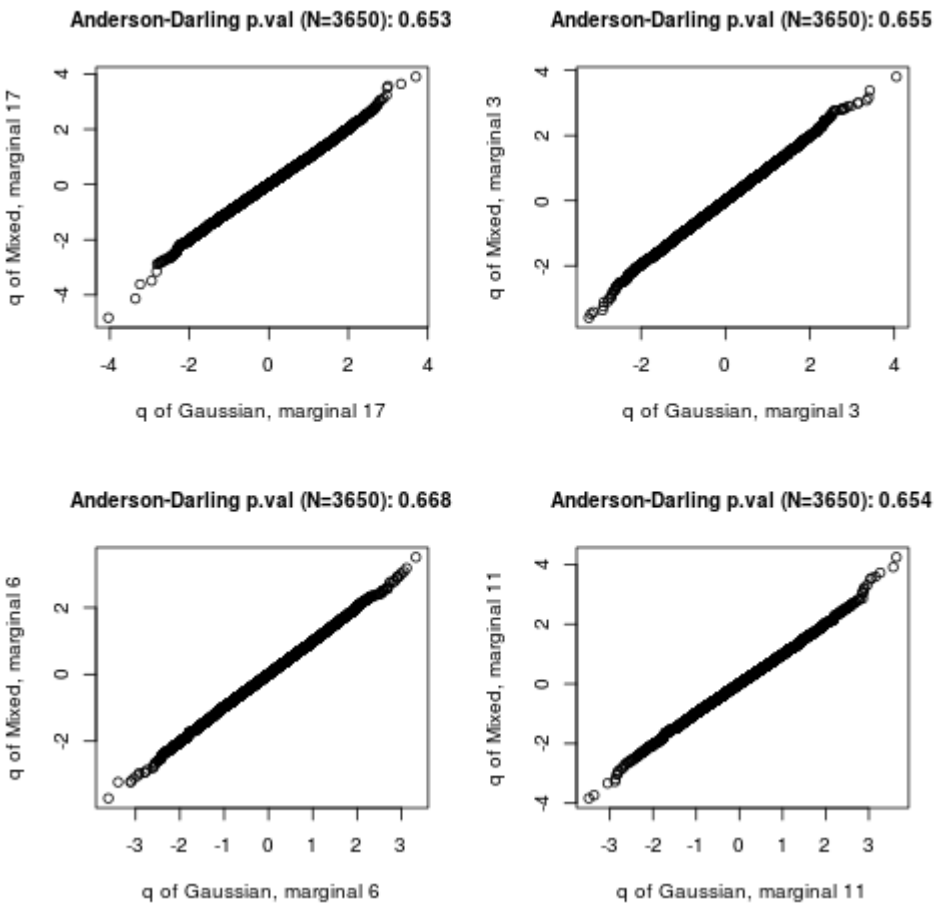

FIGURE 5.4. Quantile-quantile plots of four randomly selected components of $\mathbf{Z}$ and $\mathbf{X}$. The p-values of the Anderson-Darling test for equality in distribution (n=3650) are given. The marginal distributions illustrated can be reasonably accepted to be equal.

A more detailed analysis of the 2-dimensional components of $\mathbf{Z}$ and $\mathbf{X}$, comprises the study of their respective empirical copulas. Data plotted at figure (5.6) is given, exemplifying for data of vector $\mathbf{X}$, by

$$u_{i,j} = \hat{F}_j\left(x_{i,j}\right)$$

where

$$\hat{F}_j\left(a\right) := \frac{\#\left\{x_{i,j} : x_{i,j} \leq a\right\}}{n+1}$$

stands for the empirical cumulative distribution function of component X_j, for $j = 1, \ldots, 30$, and $i = 1, \ldots, n$. Visually, both data sets seem to have the same empirical copula. The test proposed by Kojadinovic and Yan (2011) and implemented for package *copula* of R, was applied to a randomly selected sub-sample of size n=500 of data from (X_3, X_{28}), with the number of multipliers replications set to N=1000.

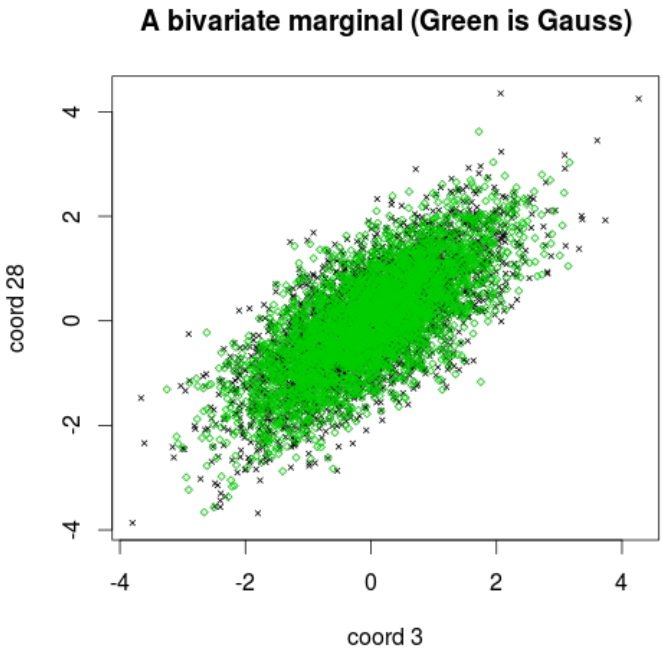

FIGURE 5.5. Dispersion plot of two typical components of $\mathbf{Z}$ and $\mathbf{X}$. The p-value of the multivariate Shapiro-Wilks test applied to 500 randomly selected samples of (X_3, X_{28}) test is 0.191.

α-Level	(Z_{j_1}, Z_{j_2})	(X_{j_1}, X_{j_2})
0.01	433 (99.54%)	433 (99.54%)
0.05	413 (94.94%)	416 (95.63%)
0.10	392 (90.11%)	397 (91.26%)

TABLE 5.2. Summary of goodness of fit test for the Gaussian copula applied on all bivariate marginal distributions of $\mathbf{Z}$ and $\mathbf{X}$. A random sub-sample (n=500) from the available data was used for each testing. Out of the total $\binom{30}{2} = 435$ bivariate combinations, the total number (and percentage) of combinations by which the Normality hypothesis cannot be rejected at the respective α-level are shown.

The resulting p-value is 0.955, whereby Gaussianity in the underlying copula seems an acceptable hypothesis. Note that this test is already very efficient under sample sizes of n=300 (see Kojadinovic and Yan (2011)).

The same testing procedure was applied to all possible pair-wise combinations of components of $\mathbf{Z}$ and $\mathbf{X}$, as had been done with the multivariate Shapiro-Wilks test. Results are summarized at table (5.2). Again, the bivariate sub-vectors of $\mathbf{X}$ are most of the time considered to have the Gaussian copula, in a qualitatively similar proportion as the 2-dimensional sub-vectors of $\mathbf{Z}$.

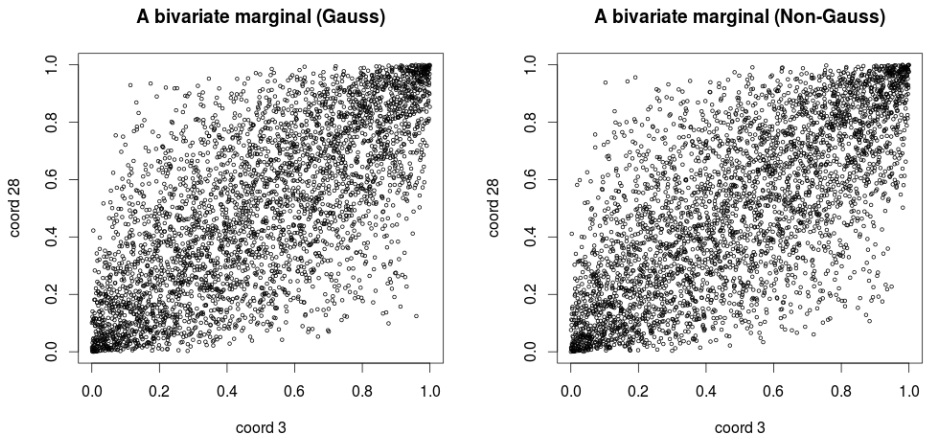

FIGURE 5.6. Empirical copula plots for: (left) data from (Z_3, Z_{28}), and (right) data from (X_3, X_{28}). The p-value of goodness of fit test for Gaussianity on a randomly selected subsample (n=500) is 0.955.

D.o.f quantile (%)	(Z_{j_1}, Z_{j_2})	(X_{j_1}, X_{j_2})
5%	35.75	17.97
50%	500.00	34.67
95%	500.00	500.00

TABLE 5.3. Quantiles of the degrees of freedom fitted to each of the 435 pairs combinations (Z_{j_1}, Z_{j_2}) and (X_{j_1}, X_{j_2}). Using all data, the fitted degrees of freedom are 500 and 45.75 for $\mathbf{Z}$ and $\mathbf{X}$, respectively.

We also fitted T-copulas to the data of all 435 pairs of components, using the data available (n=3650). The idea is to find out how many degrees of freedom would be an optimal assignment for each pair of components, both of $\mathbf{Z}$ and of $\mathbf{X}$. The fitting method employed is described at section 4.2 of Demarta and McNeil (2005), and named "method of moments using Kendall's Tau".

The 5%, 50% and 95% quantiles of the fitted degrees of freedom are shown at table (5.3). By fitting all data one gets to a T-copula with 500 and 34.62 degrees of freedom for $\mathbf{Z}$ and $\mathbf{X}$, respectively[2]. As seen in section (5.1), however, the tail dependence of $\mathbf{X}$ is comparable to that of a multivariate T distribution with 15 degrees of freedom, a fact totally invisible for the T-copula fitting procedure, even with a sample size of n=3650. Such a tail behavior, which has gone mostly unnoticed in the one and two dimensional marginals (what Geostatistics check!), may have a great impact on the wider field, of which the data from $\mathbf{X}$ constitute but a partial observation. See section (5.7).

[2]500 degrees of freedom were the highest possible attainable with the employed fitting algorithm.

Vector	Mean	Range par.	Nugget	Var
Z	0.007	19.966	0.000	1.000
X	0.006	19.992	0.000	0.992
W	0.006	19.876	0.000	0.994

TABLE 5.4. Gaussian field specification, as estimated by maximum likelihood (n=3650), and using the exponential covariance function model.

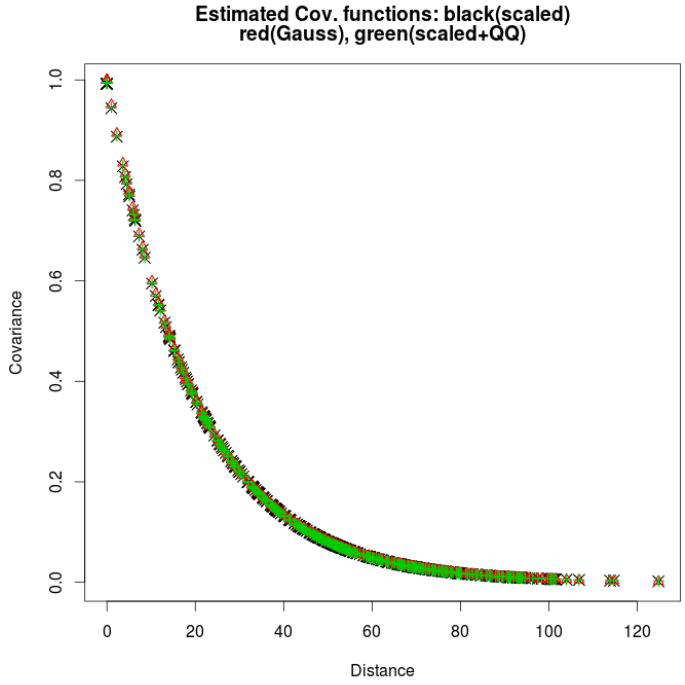

FIGURE 5.7. Plots of fitted exponential covariance functions for **Z** (red), **X** (black) and **W** (green). Plots are practically identical.

5.5. **The fitted covariance function.** On the basis of the analysis of the one and two dimensional marginal distributions, we deem adequate to fit a multivariate Normal distribution to **Z**, **X** and **W**.

Since data comes from the Spatial context illustrated at figure (5.3), we fit covariance matrices, $cov\,(\mathbf{Z})$, $cov\,(\mathbf{X})$ and $cov\,(\mathbf{W})$, using an exponential covariance function. The estimation method was maximum likelihood using the Normal distribution as model. Estimated parameters are shown in table (5.4), whereas plots of the resulting covariance functions appear at figure (5.7).

Note that both the parameter estimates and the covariance function plots are practically identical. As was true during the analysis of the one and two dimensional marginal distributions, there is little evidence that the distributions of **Z**, **X** and **W** are not the same.

5.6. **Analysis of Aggregating statistics.** We have seen that both the 1-dimensional and the 2-dimensional marginal distributions of $\mathbf{X}$ and $\mathbf{W}$ seem to indicate that these vectors can be safely modeled by a multivariate Normal model, like the one suitable for $\mathbf{Z}$. We know, however, that the distributions of $\mathbf{Z}$ and $\mathbf{X}$ are not the same.

In this sub-section we compute some statistics that indicate that the probability distributions of $\mathbf{X}$ and $\mathbf{W}$ may actually be different from that of $\mathbf{Z}$. In Rodríguez and Bárdossy (2013) these statistics are called *interactions manifestations*.

The first aggregating statistic we consider, is the number of components trespassing a given threshold a. Data observed from $\mathbf{X}$ lead to realizations of random variable $L_{\mathbf{X}}$, defined as

$$L^{\mathbf{X}} = \sum_{j=1}^{J} 1\left\{X_j > a\right\}$$

where

$$1\left\{X_j > a\right\} = \begin{cases} 1, & X_j > a \\ 0, & X_j \leq a \end{cases} \tag{5.2}$$

Similar constructions lead to $L^{\mathbf{Z}}$ and $L^{\mathbf{W}}$ from $\mathbf{Z}$ and $\mathbf{W}$, respectively. Denote by $l_1^{\mathbf{Z}}, \ldots, l_n^{\mathbf{Z}}$; $l_1^{\mathbf{X}}, \ldots, l_n^{\mathbf{X}}$ and $l_1^{\mathbf{W}}, \ldots, l_n^{\mathbf{W}}$ the samples of $L^{\mathbf{Z}}$, $L^{\mathbf{X}}$ and $L^{\mathbf{W}}$. These are plotted in figure (5.8). Note that the difference among the plots begins to be quite apparent for thresholds 2.326 through 3.09. As opposed to what was seen when analyzing the one and two dimensional marginals separately, there seems to be a difference among the distributions of $L^{\mathbf{Z}}$, $L^{\mathbf{X}}$ and $L^{\mathbf{W}}$, and hence of $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{W}$.

The second statistic we mention, is the "congregation measure" used by Bárdossy and Pegram (2009) and Bárdossy and Pegram (2012), for the sake of model validation. This is a measure not affected by monotonic transformations on the components of the vector analyzed.

The congregation measure referred to is constructed as follows. Set a threshold percentile, $b \in (0,1)$. Select a set of indexes $(j_{i_1}, \ldots j_{i_K})$, with $1 \leq j_{i_1} < \ldots < j_{i_K} \leq J$. For the analysis of the components of $\mathbf{X}$, define binary random variables

$$\varsigma_b\left(j_{i_k}\right) = \begin{cases} 1, & F_{j_{i_k}}\left(X_{j_{i_k}}\right) > b \\ 0, & F_{j_{i_k}}\left(X_{j_{i_k}}\right) \leq b \end{cases} \tag{5.3}$$

This results in a discrete random vector $\varsigma_b = \left(\varsigma_b\left(j_{i_1}\right), \ldots, \varsigma_b\left(j_{i_K}\right)\right)$. The congregation measure referred to is defined to be the entropy of a sub-vector of ς,

$$\begin{aligned} congr_b\left(X_{j_{i_1}}, \ldots, X_{j_{i_K}}\right) = \\ - \sum_{j_{i_1}, \ldots, j_{i_K}} \Pr\left(\varsigma_b\left(j_{i_1}\right), \ldots, \varsigma_b\left(j_{i_K}\right)\right) \log\left(\Pr\left(\varsigma_b\left(j_{i_1}\right), \ldots, \varsigma_b\left(j_{i_K}\right)\right)\right) \end{aligned} \tag{5.4}$$

That is, the measure is defined as the entropy of the joint distribution of the binary variables just defined. A higher value of this measure indicates less association. A similar definition applies to $congr_b\left(Z_{j_{i_1}}, \ldots, Z_{j_{i_K}}\right)$. Note that this measure is not affected by the marginal distributions of the components employed, hence

$$congr_b\left(W_{j_{i_1}}, \ldots, W_{j_{i_K}}\right) = congr_b\left(X_{j_{i_1}}, \ldots, X_{j_{i_K}}\right)$$

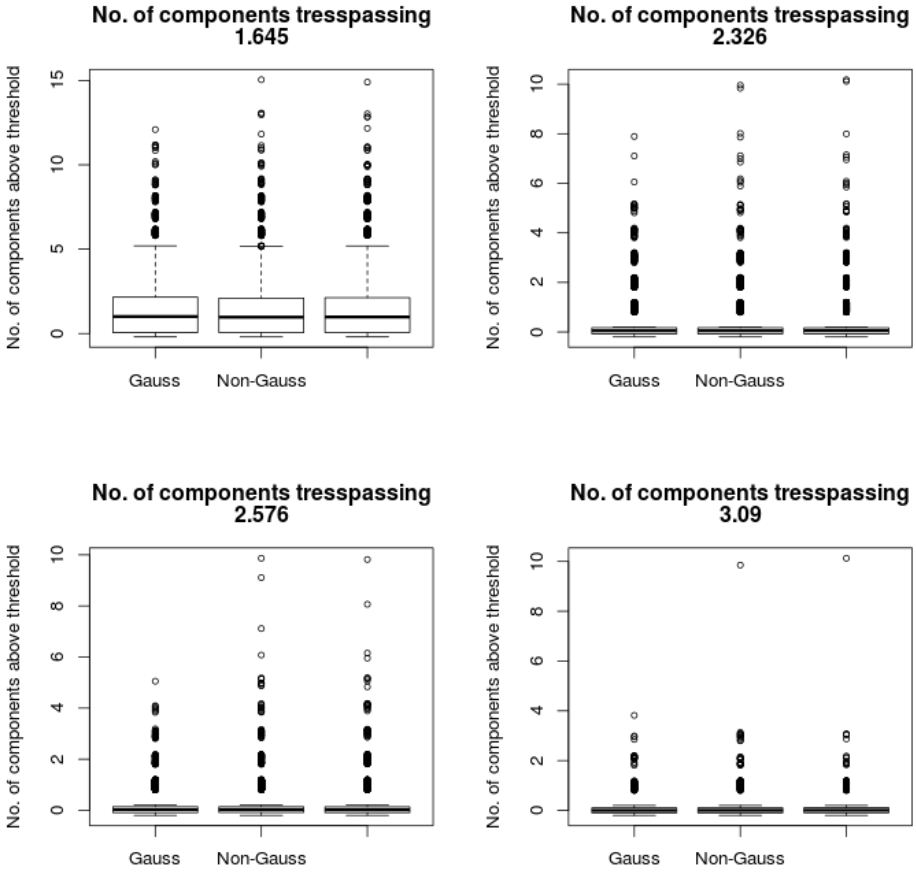

FIGURE 5.8. Boxplots of number of components trespassing the four thresholds indicated (1.645, 2.326, 2.576 and 3.09), for samples from $\mathbf{Z}$ (left), $\mathbf{X}$ (middle) and $\mathbf{Z}$ (right). Data has been jittered for visualization purposes. The difference among the plots becomes most apparent as the threshold is pulled up.

We applied this measure to three components of vectors $\mathbf{Z}$ and $\mathbf{X}$, namely components 3, 28 and 19, of which the former two were visualized at figure (5.5). We used percentiles $b \in \{0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 0.99, 0.995, 0.999\}$, and computed

$$r_b = \frac{congr_b\left(Z_{j_{i_1}}, \ldots, Z_{j_{i_K}}\right)}{congr_b\left(X_{j_{i_1}}, \ldots, X_{j_{i_K}}\right)}$$

The resulting ratio values are shown in figure (5.9). The estimated association of the components (Z_3, Z_{28}, Z_{19}), as quantified by this measure, does not seem to decrease considerably as compared to that of (X_3, X_{28}, X_{19}). A "parametric bootstrap" (see Efron and Tibshirani (1993)) confidence interval has been added

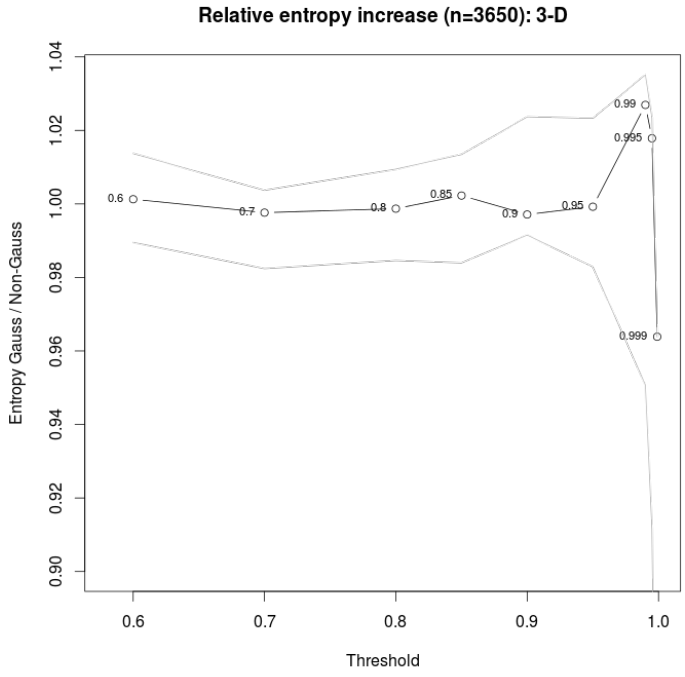

FIGURE 5.9. Ratio, r_b, of congregation measures for the percentiles $b \in \{0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 0.99, 0.995, 0.999\}$. Association among the three components (Z_3, Z_{28}, Z_{19}) decreases considerably as compared to that of (X_3, X_{28}, X_{19}) from the 99% percentile on. A bootstrap based confidence interval has been added for significance assessment.

for the ratio of the entropies, computed by simulating a Normal sample of size $n = 3650$ with zero means and correlation matrix the sample correlation matrix of (Z_3, Z_{28}, Z_{19}). The procedure is repeated 10000 times to create the confidence interval.

However, as shown in figure (5.10), if the sample size is increased to n=10000, the association among subvectors of $\mathbf{X}$ can be seen to increase considerably as compared to that of subvectors of $\mathbf{Z}$. In particular this is the case as one approaches the uppermost tail of the 2, 3, 4 and 5-dimensional distributions. Subvectors employed for the evaluation are indicated in figure (5.10). This was to be expected in view of the uppermost tail of scaling variable V, see the right panel of figure (5.2). Hence, on the basis of the analysis of only three through five components, it is possible to notice a difference in the dependence structure of the fields (compare Bárdossy and Pegram (2009)), provided the sample size is sufficiently large.

A third kind of aggregating statistic comprises the quantiles of the sum of components above given thresholds. To this end, we took the n=3650 observations of each vector, $\mathbf{Z}$, $\mathbf{X}$ and $\mathbf{W}$, and obtained from them $s_1^{\mathbf{Z}}, \ldots, s_n^{\mathbf{Z}}, \ldots, s_n^{\mathbf{W}}$. Where, for

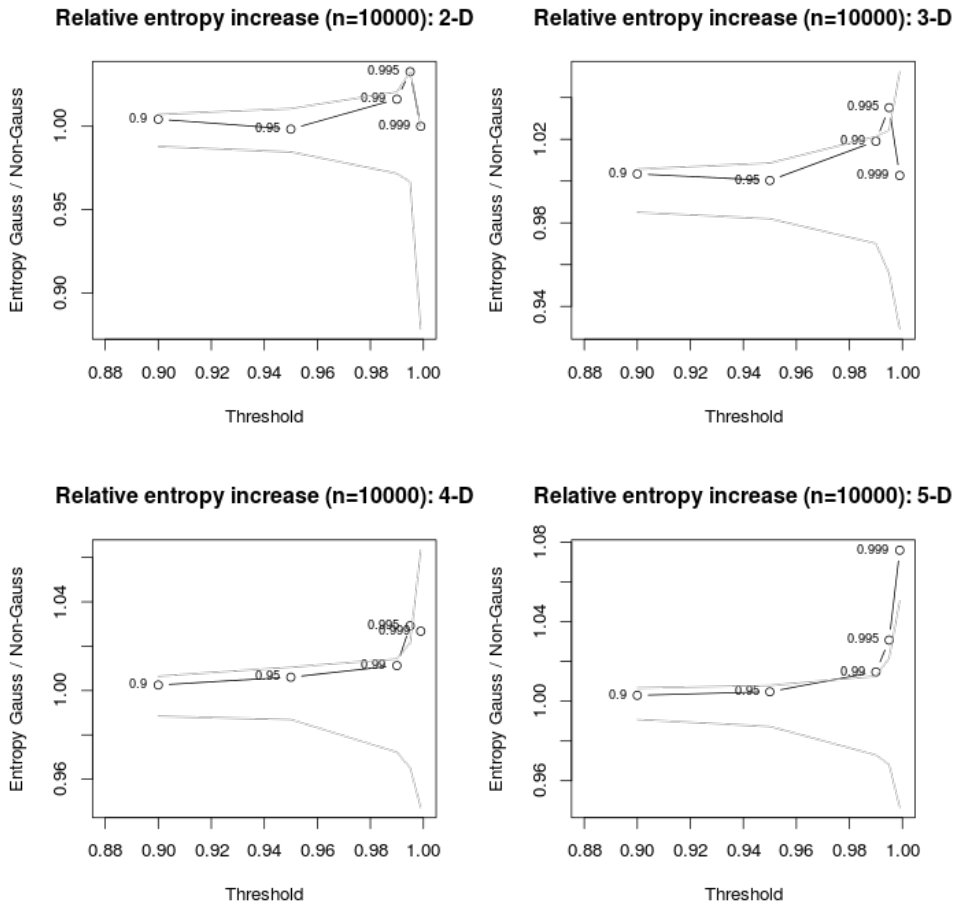

FIGURE 5.10. Ratio, r_b, of congregation measures for the percentiles $b \in \{0.6, 0.7, 0.8, 0.85, 0.9, 0.95, 0.99, 0.995, 0.999\}$. Association among two, three, four and five components of $\mathbf{Z}$ decreases considerably as compared to that of subvectors of $\mathbf{X}$ from the 99% percentile on. A bootstrap based confidence intervals have been added for significance assessment. Subvectors used have indexes $(3, 28)$, $(3, 28, 19)$, $(3, 28, 19, 16)$ and $(3, 28, 19, 16, 25)$.

a given threshold a, one has for example,

$$s_i^{\mathbf{X}} = \sum_{j=1}^{30} 1\{x_{ij} > a\} \times x_{ij} \tag{5.5}$$

with $i = 1, \ldots, 3650$.

The empirical cumulative distribution functions built from $s_1^{\mathbf{X}}, \ldots, s_{3650}^{\mathbf{X}}$ and $s_1^{\mathbf{W}}, \ldots, s_{3650}^{\mathbf{W}}$ are presented in figure (5.11), for thresholds $a = \{1.04, 1.28, 2.5, 3\}$. Simulation based 90% confidence intervals (appearing in red) for the empirical cumulative distribution function of $s_1^{\mathbf{Z}}, \ldots, s_{3650}^{\mathbf{Z}}$ were also added. These confidence

intervals were created for each threshold, a, by repeting 1000 times the following procedure: simulate n=3650 realizations of a gaussian random vector with mean and covariance as estimated for $\mathbf{X}$ in section (5.5), and then apply construction (5.5). In this manner we obtain 1000 empirical cumulative distribution functions; at each percentile $u \in [0,1]$, we compute the values of all 1000 e.c.d.f. and take from them the 5% and 95% quantile values.

It is clear from figure (5.11), that tails of the distribution functions obtained for $s_1^{\mathbf{X}}, \ldots, s_{3650}^{\mathbf{X}}$ for thresholds $a \in \{2.5, 3\}$ are heavier than expected from a gaussian vector having the same means, covariance matrix, and approximately the same marginal distributions as $\mathbf{X}$. Hence, the analysis of these statistics is valuable for diagnosis of higher order interaction.

5.7. **Applications-relevant discrepancies in the underlying fields.** The object of this section and of section (5.8) is to show what kind of inference can go (unnoticed) wrong, if one does not pay attention to the discrepancies pointed out by the aggregating statistics shown above. We focus on characteristics of the whole underlying fields, relevant for hydrological applications, in this section. In section (5.8) we deal with conditional distributions, more relevant for mining geostatistics.

We present two aggregating statistics of the complete fields, $\mathbf{Z}^*$, $\mathbf{X}^*$ and $\mathbf{W}^*$, portrayed in figure (5.3). We indicate the kind of global statistics that would go totally unnoticed, if we were to check only the one and two dimensional marginal distributions of the data available, and fit a Gaussian field to the whole region.

Sums of positive values of the whole field. The first interaction manifestation we investigate is the distribution of

$$S^+_{\mathbf{X}^*} = \sum_{j=1}^{J=90000} \max(X_j, 0) \tag{5.6}$$

that is, of the sum of positive values of the whole field, $\mathbf{X}^*$. Similarly, we define $S^+_{\mathbf{Z}^*}$ and $S^+_{\mathbf{W}^*}$ on the basis of fields $\mathbf{Z}^*$ and $\mathbf{W}^*$, respectively.

The distribution is investigated in terms of the sample quantiles of $S^+_{\mathbf{X}^*}$, $S^+_{\mathbf{Z}^*}$ and $S^+_{\mathbf{W}^*}$. These are important statistics for rainfall modeling over a basin, for example, since a value proportional to this sum must find its way through the outlet of the basin, possibly causing a flood.

Boxplots illustrating the distribution of the sum of positive values are given in figure (5.12), whereas a table with some important sample quantiles of $S^+_{\mathbf{Z}^*}$, $S^+_{\mathbf{X}^*}$ and $S^+_{\mathbf{W}^*}$ are given in table (5.5). Notice that the sample quantiles of $S^+_{\mathbf{Z}^*}$ begin to deviate from those of of $S^+_{\mathbf{X}^*}$ and $S^+_{\mathbf{W}^*}$ from the 99% quantile on. The relative percentage increase of the quantiles of $S^+_{\mathbf{X}^*}$ and $S^+_{\mathbf{W}^*}$ with respect to those of $S^+_{\mathbf{Z}^*}$ are given within parentheses in table (5.5).

A sample of size $n = 3650$ would amount to a 10-year period, if data were to represent some daily measured variable. If the field $\mathbf{X}^*$ were to represent daily rainfall over a basin, the maximum total rainfall would be 47.58% higher than one would expect by fitting a gaussian model with adequate one and two dimensional marginal distributions and covariance function. By letting the simulation run up to $n = 10000$ (roughly ten years data), the increase in the maximum sum ascends to 61.11% for $\mathbf{X}^*$ and to 50.36% for $\mathbf{W}^*$, as compared to the maximum sum produced by field $\mathbf{Z}^*$. These possibilities are completely missed by an analysis based on one and two dimensional marginal distributions, and the field's covariance function.

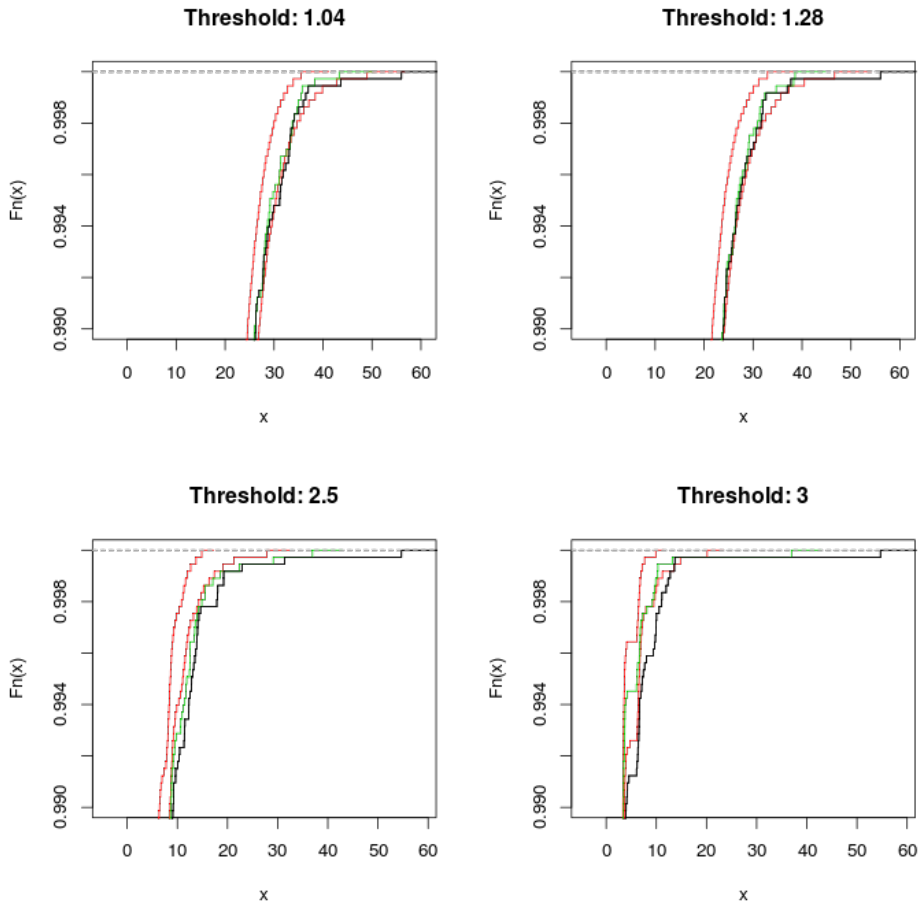

FIGURE 5.11. Empirical distributions of sums above various thresholds (1.04, 1.28, 2.5 and 3) for the n=3650 observations of random vectors $\mathbf{X}$ (black) and $\mathbf{W}$ (green). Simulation based 90% confidence intervals for the empirical distributions arising from a gaussian vector with the same means and covariance matrix as $\mathbf{X}$ appear in red. The tail of the sums above thresholds 2.5 and 3 is significantly heavier than the gaussian model would prescribe.

Number of components of the whole field trespassing a given threshold. A second interaction manifestation we shall investigate for the whole fields, is the distribution of the number of components trespassing a given threshold. Analogously to (5.2), we define for $\mathbf{X}^* \in \mathbb{R}^{J^*}$,

$$L^{\mathbf{X}^*} = \sum_{j=1}^{J^*} 1\left\{X_j^* > a\right\}$$

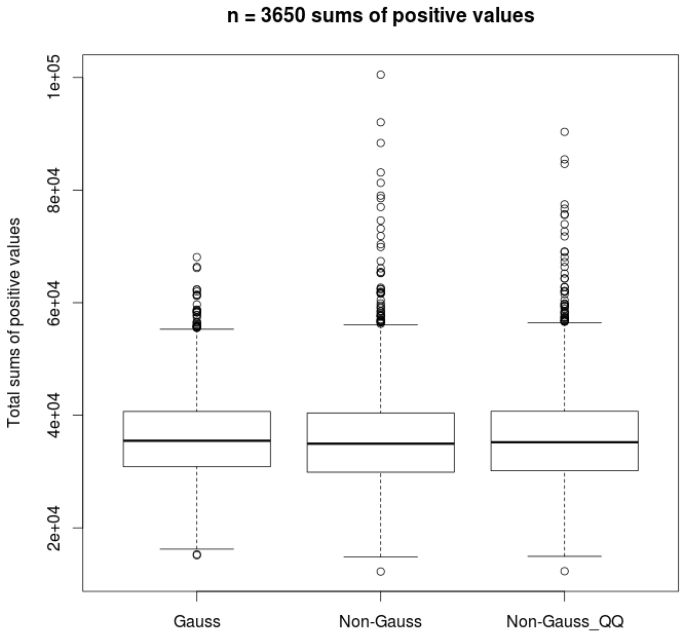

FIGURE 5.12. Boxplots of sums of positive values for n realizations of $\mathbf{Z}^*$, $\mathbf{X}^*$ and $\mathbf{W}^*$, with $n = 3650$.

Quantile	Gauss	Non-Gauss	Non-Gauss_QQ
80%	41901.96	41992.56 (+0.22%)	42288.41 (+0.92%)
90%	45342.97	46334.13 (+2.19%)	46576.23 (+2.72%)
95%	48609.21	49678.96 (+2.20%)	49910.30 (+2.68%)
99%	54906.90	57634.93 (+4.97%)	57583.49 (+4.87%)
99.5%	57660.03	62627.25 (+8.61%)	62794.43 (+8.90%)
99.9%	62331.09	81939.63 (+31.46%)	76972.03 (+23.49%)
100%	68099.17	100503.12 (+47.58%)	90353.53 (+32.68%)

TABLE 5.5. From left to right: Sample quantiles ($n = 3650$) for $S^+_{\mathbf{Z}^*}$, $S^+_{\mathbf{X}^*}$ and $S^+_{\mathbf{W}^*}$. Percentages within parentheses indicate percentage increase with respect to data from the Gaussian field.

where $J^* = 300 \times 300$, and

$$1\left\{X_j^* > a\right\} = \begin{cases} 1, & X_j^* > a \\ 0, & X_j^* \leq a \end{cases} \tag{5.7}$$

In the context of spatial statistics, $L^{\mathbf{X}^*}$ can be interpreted as the total area over which the environmental variable of interest realizes "extreme" values. Similar constructions define $L^{\mathbf{Z}^*}$ and $L^{\mathbf{W}^*}$. We have a total of n=3650 samples from each

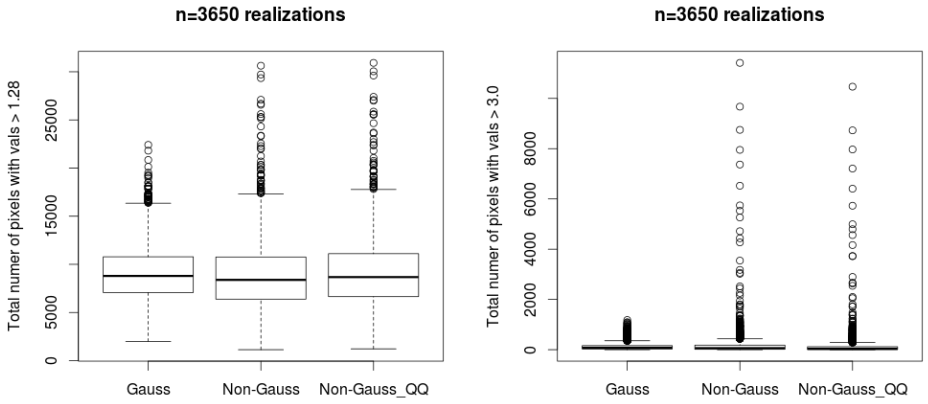

FIGURE 5.13. Boxplots of the number of components above thresholds 1.28 (left) and 3.0 (right). The two non-gaussian fields are very different from the gaussian one with respect to this interaction manifestation. The difference is exacerbated as the threshold is pulled upwards. More extensive areas with very high values are to be expected for the non-gaussian fields.

of these three random variables, which are plotted at figure (5.13) for thresholds 1.28 (left) and 2.5 (right).

Notice the great difference between the samples of $L^{\mathbf{Z}^*}$ and $L^{\mathbf{W}^*}$ (labeled "Gauss" and "Non-Gauss_QQ", respectively) when we use 2.5 as threshold. This occurs even though marginally fields $\mathbf{Z}^*$ and $\mathbf{W}^*$ have exactly the same distribution, and the covariance function of both fields is the same. In more practical terms, the difference in this variable amounts to $\mathbf{Z}^*$ and $\mathbf{W}^*$ having very different types of clusters, as illustrated in figure (5.14). Field $\mathbf{W}^*$ can exhibit big clusters of values above 4 (99.99683% quantile of its marginal distribution), even though marginally and in terms of its covariance function it has the same specification as $\mathbf{Z}^*$.

5.8. **Conditional distributions and interpolation.** The conditional distribution of the random quantity at a new location, given a partial observation of the field will now be analyzed, by using the approximation given at equation (4.3). This is an important type of distribution in mining geostatistics, where inference on the random quantity investigated is necessary at unbored locations. We shall illustrate the type of discrepancy between the conditional distribution arising from a gaussian model as compared with model given by equation (2.5), by focusing on a realization of a sub-vector of $\mathbf{X}$.

Vector

$$\mathbf{z}_I = (-1.489, -0.626, -0.050, 0.068, 0.491, 0.832, -0.666) \tag{5.8}$$

constitutes the first realization of $\mathbf{Z}_I = (Z_3, Z_{28}, Z_{19}, Z_{16}, Z_{25}, Z_9, Z_{21})$ used in the illustrative example of this section. Due to the mechanism depicted by equation (3.3), one can have immediately a realization, $\mathbf{x}_I$, of $\mathbf{X}_I = (X_3, \ldots, X_{21})$ by multiplying $\mathbf{Z}_I$ by probable values of scaling variable $\sqrt{V}$. For our analysis we employ

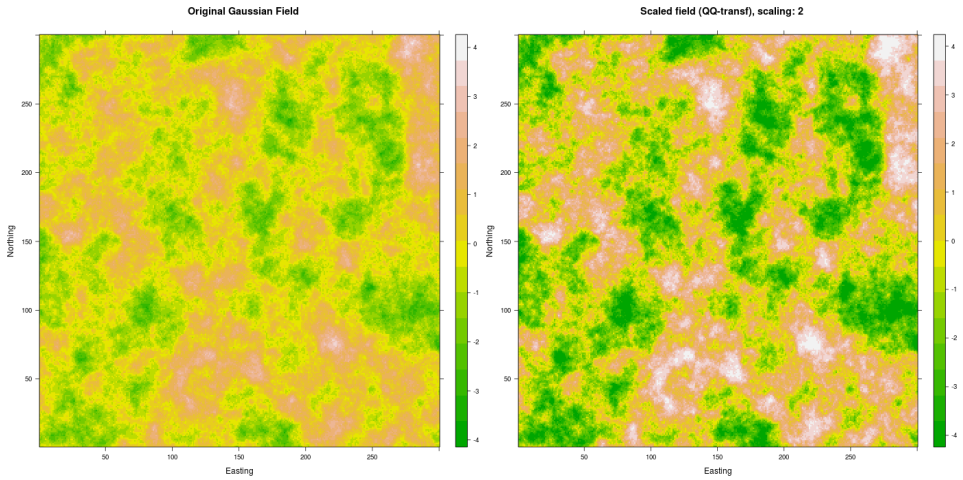

FIGURE 5.14. Different clustering characteristics between one realization of the gaussian (left) and non-gaussian (right) fields $\mathbf{Z}^*$ and $\mathbf{W}^*$. Scaling originally corresponding to non-gaussian field is $\sqrt{V} = 2$. Field $\mathbf{W}^*$ can exhibit big clusters of values around 4, even though marginally and in terms of the covariance function it has the same specification as $\mathbf{Z}^*$.

the following values as realizations of $\sqrt{V}$: 0.64, 1 and 2; hence obtaining three different realizations of $\mathbf{X}_I$.

We analyze the distribution of Z_3 and of X_3, conditioned on an increasing number of values. Such conditioning values are given by multiplying (5.8) times 0.64, 1 and 2. We plot percentiles: 80%, 90%, 95%, 99%, 99.5%, 99.9%, 99.99% and 99.999%.

In figure (5.15) we show the conditional distributions using $\sqrt{v} = 0.64$ for $\mathbf{X}_I$. A moderate increase in the discrepancy between the conditional distributions is seen as the number of conditioning values increases, while the tail of the non-gaussian distribution becomes lighter and lighter as compared with that of the conditional gaussian.

In figure (5.16) we show the conditional distributions using $\sqrt{v} = 1$ for $\mathbf{X}_I$. Note that the non-gaussian conditional distribution keeps its similarity to the gaussian conditional, though it has higher quantiles for all conditioning schemes.

In figure (5.15) we show the conditional distributions using $\sqrt{v} = 2$ for $\mathbf{X}_I$. The conditional distribution given one conditioning value is very similar for both models, but this situation quickly changes, as more conditioning values are considered. The quantiles of the upper part of the conditional distribution for X_3 become sensibly bigger already for 2 conditioning values. Note that the values one may expect for the conditioned ("ungauged") variable is considerably greater for X_3 than for Z_3. This is a relevant issue for applications.

5.9. **Estimated parameters.** We employ now the simple technique given in section (3.3) to estimate the additional parameters, $c_2, \ldots, c_5$, corresponding to the cumulants of a non-degenerate (constant) scaling variable V. The estimated cumulants and moments of squared scaling variable V are presented in table (5.6).

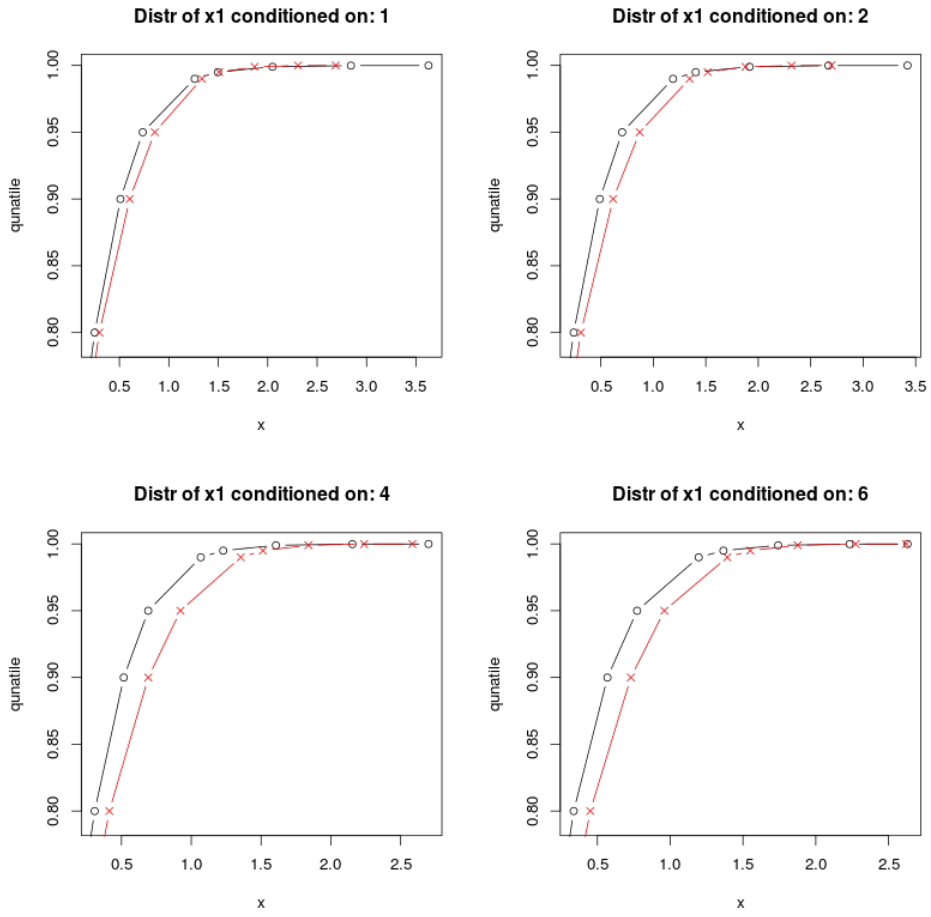

FIGURE 5.15. From left to right and downwards: Comparison of upper parts of conditional distributions for Z_3 (red) and X_3 (black), for $n = 1, 2, 4, 6$ conditioning values. Realization of $\mathbf{X}_I$ is given by $0.64 \times \mathbf{z}_I$ (small scaling variable).

Using the method of moments, and the n=3650 data values, we fitted a mixture of 5 gamma distributions to each the series of moments shown in table (5.6). The resulting distributions, together with the distribution of the original squared scaling variable $V > 0$ are shown and compared in figure (5.18). The distribution of the scaling variable of a student multivariate distribution with 15 degrees of freedom, shown in blue, has been added for comparison.

We notice that the scaling variable is approximately recovered by this technique. However this technique cannot be used, for example, in the context of rainfall modeling, where data is constrained to be positive. Even though a latent variable approach (cf. Sanso and Guenni (1999)) can be employed for fitting the best gaussian model to data (step 1 of estimation), the step effecting the estimation of additional parameters $c_2, \ldots, c_5$ determining important interaction manifestations

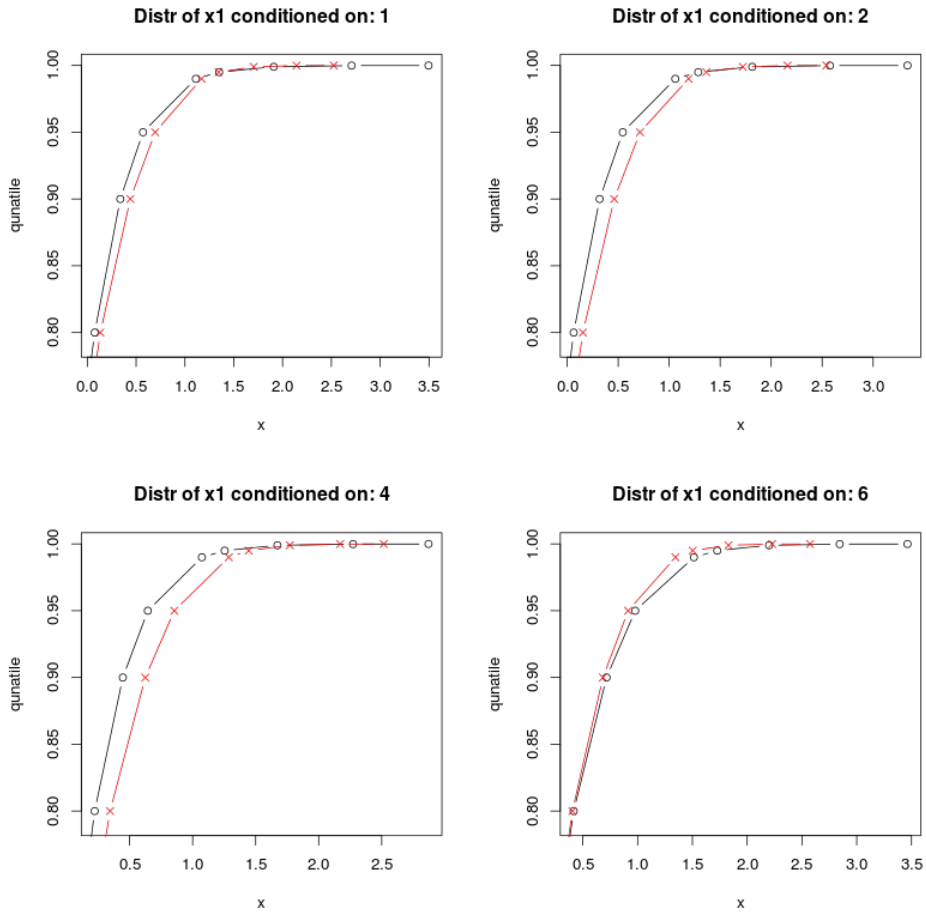

FIGURE 5.16. From left to right and downwards: Comparison of upper parts of conditional distributions for Z_3 (red) and X_3 (black), for $n = 1, 2, 4, 6$ conditioning values. Realization of $\mathbf{X}_I$ is given by $1 \times \mathbf{z}_I$ (middle-valued scaling variable).

of the field, cannot be executed via the method of moments: the latent imputed data correspond to a Normal distribution and hence does not produce valid realizations of squared generating variable, R^2.

A paper describing an alternative second step for the estimation method, which circumvents this difficulty, is already in preparation.

6. Discussion and work in progress

We have seen that modeling of spatial data that attempts to fit properly one and two dimensional marginal distributions only, may lead to erroneous specification of the underlying mechanism which generates data. As an illustration we have seen that data from random vectors $\mathbf{Z}$, $\mathbf{X}$ and $\mathbf{W}$, which represent a network of 30

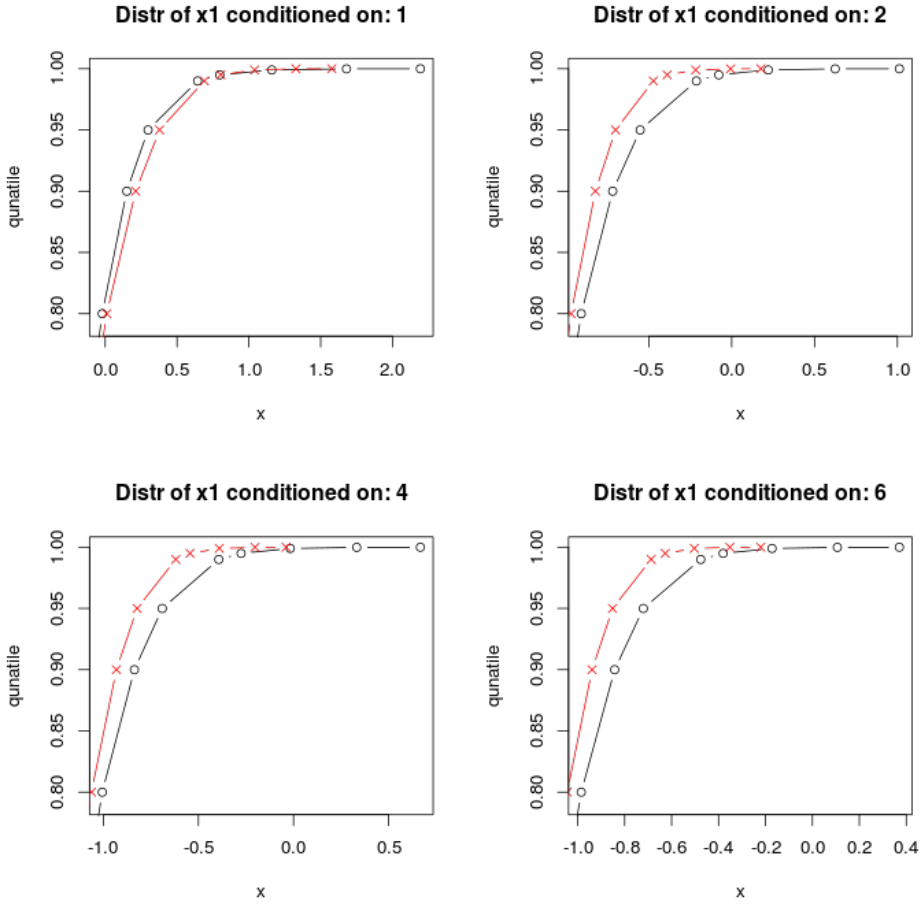

FIGURE 5.17. From left to right and downwards: Comparison of upper parts of conditional distributions for Z_3 (red) and X_3 (black), for $n = 1, 2, 4, 6$ conditioning values. Realization of $\mathbf{X}_I$ is given by $2 \times \mathbf{z}_I$ (high-valued scaling variable).

gauging stations each, had to be more carefully analyzed in order to notice possible discrepancies in their dependence structure. Conveniently selected aggregating statistics can help to notice dependence structure differences that are not visible in the one and two dimensional marginals.

In particular, the congregation measure (5.4) and the distribution of the number of components trespassing a given threshold (5.2) pointed to discrepancies between gaussian field $\mathbf{Z}^*$ and non-gaussian fields $\mathbf{X}^*$ and $\mathbf{W}^*$, on the basis of 3650 realizations of respective sub-vectors $\mathbf{Z}$, $\mathbf{X}$ and $\mathbf{W}$.

The model introduced by (2.5) allows for the consideration of joint cumulants of order greater than two (covariances) in a manner that is convenient for spatial modeling: building on available geostatistical techniques, requiring a minimum of extra parameters, and respecting the principle of consistency.

Estimated	Gauss	Non-Gauss	Non-Gauss_QQ
m.1	1	1	1
m.2	0.9998	1.0792	1.0547
m.3	0.9988	1.4098	1.2378
m.4	0.9958	2.8072	1.8129
m.5	0.9897	9.0952	3.6831
c.1	1	1	1
c.2	-2e-04	0.0792	0.0547
c.3	-7e-04	0.1722	0.0738
c.4	-3e-04	0.6244	0.1806
c.5	2e-04	2.2289	0.4100

TABLE 5.6. Coefficients estimated by the method of moments, rounded up to four decimal places.

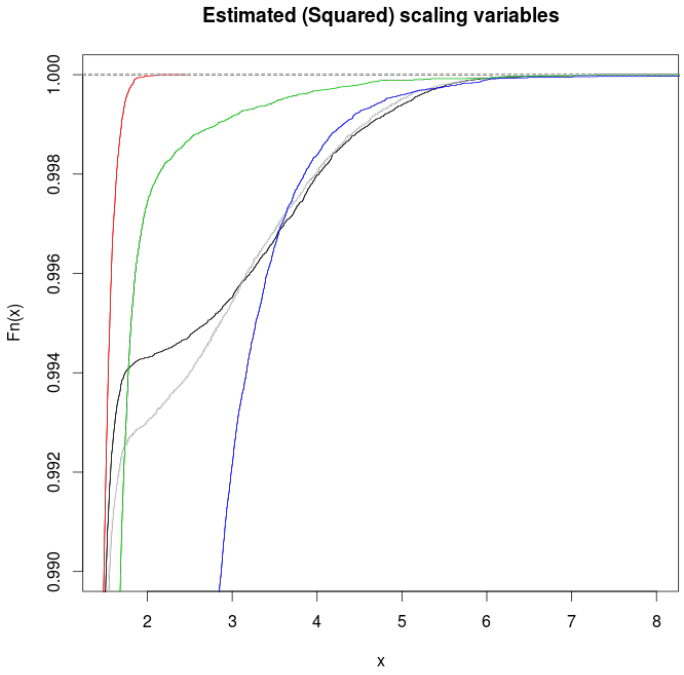

FIGURE 5.18. Estimated squared scaling variables, uppermost part: for $\mathbf{X}$ (black), $\mathbf{Z}$ (red), $\mathbf{W}$ (green), and a Student with 15 degrees of freedom (blue). The squared scaling variable originally employed for the simulation case study is shown in gray. The method of moments estimation was successful in capturing the uppermost behavior of the scaling variable.

The basic estimation technique outlined in section (3.3) was successful in recovering scaling variable V of (3.3), which is responsible for beyond-covariance interactions among components of field $\mathbf{X}^*$. An alternative method is under testing, which can allow estimation even if we need to deal with truncated data, as in the case of daily rainfall modeling.

The specific implementation of model (2.5), whereby the squared scaling variable, V, is modeled by a mixture of gamma distributions, will be explored in future research.

Acknowledgments. This research forms part of the Ph.D thesis of the first author, which was funded by a scholarship of the German Academic Exchange Service (DAAD).

Appendix A. Derivation of joint cumulants

Our object of study is the cumulant generating function of a random variable $\mathbf{X} \in \mathbb{R}^J$. We shall be interested in joint cumulants such as

$$cum\left(X_{j_1}, \ldots, X_{j_r}\right) \tag{A.1}$$

where some, or all, of the indexes can be repeated. Hence it is convenient to refer to a random vector $\mathbf{X}^* \in \mathbb{R}^{J^*}$ having the components of $\mathbf{X}$, even repeated, and then find the joint cumulants that appear with degree at most one, of this "new" random vector. Thus we can, without loss of generality, focus on finding the joint cumulants with degree not greater than one, given by

$$\frac{\partial^r}{\partial t_{j_r} \ldots \partial t_{j_1}} K_{\mathbf{X}^*}\left(\mathbf{t}\right) |_{\mathbf{t}=\mathbf{0}} := cum\left(X_{j_1}, \ldots, X_{j_r}\right) \tag{A.2}$$

where no t_j, for $j \in \{j_1, \ldots, j_r\}$, is repeated.

For example, when computing the variance of a component, X_j, of $\mathbf{X}$, one would rather compute the covariance of vector $\mathbf{X}^* = (X_j, X_j)$, namely $\kappa_{11}\left(\mathbf{X}^*\right)$.

The archetypal dependence structure advocated for in this work is given by

$$K_{\mathbf{X}^*}\left(\mathbf{t}\right) = c_1 \frac{1}{2} \mathbf{t}^{\mathbf{T}} \mathbf{\Gamma} \mathbf{t} + \frac{1}{2!} c_2 \left[\frac{1}{2} \mathbf{t}^{\mathbf{T}} \mathbf{\Gamma} \mathbf{t}\right]^2 + \frac{1}{3!} c_3 \left[\frac{1}{2} \mathbf{t}^{\mathbf{T}} \mathbf{\Gamma} \mathbf{t}\right]^3 + \ldots \tag{A.3}$$

for some coefficients $c_1, c_2, c_3, \ldots$ and covariance matrix $\Gamma_{J^* \times J^*}$, and $\mathbf{t} \in \mathbb{R}^{J^*}$.

By expansion, the above expression can be written as

$$\begin{aligned} K_{\mathbf{X}^*}\left(\mathbf{t}\right) = \frac{c_1}{1!} \frac{1}{2} \sum_{j_1, j_2 = 1}^{J} t_{j_1} t_{j_2} \Gamma_{j_1 j_2} + \frac{c_2}{2!} \frac{1}{2^2} \sum_{j_1, \ldots, j_4 = 1}^{J} t_{j_1} \ldots t_{j_4} \Gamma_{j_1 j_2} \Gamma_{j_3 j_4} + \\ \frac{c_3}{3!} \frac{1}{2^3} \sum_{j_1, \ldots, j_6 = 1}^{J} t_{j_1} \ldots t_{j_6} \Gamma_{j_1 j_2} \Gamma_{j_3 j_4} \Gamma_{j_5 j_6} + \ldots \end{aligned} \tag{A.4}$$

For each coefficient $c_{\frac{r}{2}}$, for r even, there appears a sum of the form

$$\frac{c_{\frac{r}{2}}}{\frac{r}{2}!} \frac{1}{2^{\frac{r}{2}}} \sum_{j_1 = 1}^{J} \cdots \sum_{j_{2r} = 1}^{J} t_{j_1} \ldots t_{j_r} \Gamma_{j_1 j_2} \ldots \Gamma_{j_{r-1} j_r} \tag{A.5}$$

This is the only block-summand of (A.4) that does not vanish upon differentiation with respect to each variable and equation to zero, as in (A.2). Other blocks will vanish either upon differentiation with respect to a variable that does not appear

in them, or upon equation to zero, since such blocks become a sum of zeroes. So, it suffices to focus on this block, to differentiate it and equate it with zero.

Let each member of the (A.5) be labeled

$$s_{j_1\ldots,j_r} = t_{j_1}\ldots t_{j_r}\Gamma_{j_1j_2}\ldots\Gamma_{j_{r-1}j_r}$$

then, we have stated that,

$$\frac{\partial^r}{\partial t_{j_r}\ldots\partial t_{j_1}}K_{\mathbf{X}^*}(\mathbf{t}) \qquad |_{\mathbf{t}=\mathbf{0}}= \qquad \frac{c_{\frac{r}{2}}}{\frac{r}{2}!}\frac{1}{2^{\frac{r}{2}}}\sum_{j_1=1}^{J}\cdots\sum_{j_{2r}=1}^{J}\frac{\partial^r}{\partial t_{j_r}\ldots\partial t_{j_1}}s_{j_1,\ldots,j_r} \tag{A.6}$$

Partial differentiation of $s_{j_1,\ldots,j_r}$ is readily found to be

$$\frac{\partial^r}{\partial t_{j_r}\ldots\partial t_{j_1}}s_{j_1,\ldots,j_r} = \Gamma_{j_1j_2}\ldots\Gamma_{j_{r-1}j_r} \tag{A.7}$$

Sub-indexes appearing in the factors, $\Gamma_{j_1j_2}$, $\Gamma_{j_3j_4},\ldots$ constitute a partition of size $\frac{r}{2}$ of the set $A=\{j_1,j_2,\ldots,j_r\}$. That is, the union of the $\frac{r}{2}$ non-overlapping sets

$$\{j_1,j_2\},\{j_3,j_4\},\ldots,\{j_{r-1},j_r\}$$

formed with elements of set $A=\{j_1,j_2,\ldots,j_r\}$, is equal to that set:

$$\{j_1,j_2\}\cup\{j_3,j_4\}\cup\ldots\cup\{j_{r-1},j_r\}=A$$

Since the sum at (A.6) runs over all indexes in A, the sum returning the joint cumulant in question comprises all partitions of size two of A. How many different partitions of size two can be obtained for A, by forming sets out of different combinations of indexes? In general, a set with n elements, n even, can be seen to have

$$1\times 3\times\ldots\times(n-1)$$

such partitions.

We have shown that joint cumulants of the archetypal dependence structure are given by

$$cum\left(X_{j_1},\ldots,X_{j_r}\right) = \frac{c_{\frac{r}{2}}}{\frac{r}{2}!}\frac{1}{2^{\frac{r}{2}}}\sum_{j_1,\ldots,j_r=1}^{J}\Gamma_{j_1j_2}\ldots\Gamma_{j_{r-1}j_r} \tag{A.8}$$

Appendix B. Relation between moments of squared scaling variable and generating variable

Assume that we have random vector $\mathbf{Z}\in\mathbb{R}^J$ with c.g.f (2.5), with $\mu=\mathbf{0}$ and covariance matrix equal to the identity matrix, $\Sigma=I_{J\times J}$. For this special case, in agreement with representation (1.8), we have

$$\|\mathbf{Z}\|_2=\sqrt{\|\mathbf{Z}\|_2\|\mathbf{Z}\|_2}=\sqrt{\|R\times\mathbf{U}^{J-1}\|_2\|R\times\mathbf{U}^{J-1}\|_2}=R\times 1$$

since $\left\|\mathbf{U}^{J-1}\right\|_2=1$. Then,

$$R^2=\sum_{j=1}^{J}Z_j^2 \tag{B.1}$$

which in turn means that,

(B.2)

$$E\left(\left(R^2\right)^k\right) = E\left(\left(\sum_{j_1=1}^{J} Z_{j_1}^2\right) \times \ldots \times \left(\sum_{j_k=1}^{J} Z_{j_k}^2\right)\right) = \sum_{j_1=1}^{J} \ldots \sum_{j_k=1}^{J} E\left(Z_{j_1}^2 \ldots Z_{j_k}^2\right)$$

Since $\mathbf{Z}$ has c.g.f. given by

$$K_{\mathbf{Z}}(\mathbf{t}) = \frac{c_1}{1!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right) + \frac{c_2}{2!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^2 + \frac{c_3}{3!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^3 + \ldots$$

it follows, as seen in section 2, that

$$M_{\mathbf{Z}}(\mathbf{t}) = 1 + \frac{m_1}{1!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right) + \frac{m_2}{2!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^2 + \frac{m_3}{3!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^3 + \ldots$$

with coefficients given by

$$\begin{aligned} m_1 &= c_1 \\ m_2 &= c_2 + c_1^2 \\ m_3 &= c_3 + 3c_2c_1 + c_1^3 \\ m_4 &= c_4 + 4c_3c_1 + 3c_2^2 + 6c_2c_1^2 + c_1^4 \end{aligned} \tag{B.3}$$

and so on. A particular case of this function is the Gaussian moment generating function, for which all $c_{r>1}$ are set to zero. In particular, for $\xi \sim N_J(\mathbf{0}, I_{J\times J})$,

$$M_{\xi}(\mathbf{t}) = 1 + \frac{c_1}{1!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right) + \frac{c_1^2}{2!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^2 + \frac{c_1^3}{3!}\left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^3 + \ldots \tag{B.4}$$

with $c_1 = 1$. Hence joint moments of $\mathbf{Z}$ and ξ are similar, except for what pertains to coefficients $c_2, c_3, \ldots$. In fact, calling

$$h_r(\mathbf{t}) = \left(\frac{1}{2}\mathbf{t}'\mathbf{t}\right)^r$$

one has

$$\begin{cases} \frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} M_{\xi}(\mathbf{t}) = \frac{c_1}{1!}\frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} h_1(\mathbf{t}) + \frac{c_1^2}{2!}\frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} h_2(\mathbf{t}) + \ldots \\ \frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} M_{\mathbf{Z}}(\mathbf{t}) = \frac{m_1}{1!}\frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} h_1(\mathbf{t}) + \frac{m_2}{2!}\frac{\partial^{r_1+\ldots+r_k}}{\partial t_{j_1}\ldots\partial t_{j_k}} h_2(\mathbf{t}) + \ldots \end{cases}$$

Hence, for odd orders joint moments of both random vectors are zero, and for even orders

$$\begin{aligned} E(\xi_i\xi_j) &= \frac{c_1}{m_1} E(Z_iZ_j) \\ E(\xi_i\xi_j\xi_k\xi_l) &= \frac{c_1^2}{m_2} E(Z_iZ_jZ_kZ_l) \\ &\vdots \\ E\left(\xi_{j_1}^{r_1}\ldots\xi_{j_k}^{r_k}\right) &= \frac{c_1^{\frac{1}{2}\sum_{j=1}^k r_j}}{m_{\frac{1}{2}\sum_{j=1}^k r_j}} E\left(Z_{j_1}^{r_1}\ldots Z_{j_k}^{r_k}\right) \end{aligned}$$

whenever $order = \sum_{i=1}^{k} r_i$ is an even integer. It is then clear that the following relation holds, for joint moments of even order:

$$\begin{aligned} \frac{m_1}{c_1} E\left(\xi_i \xi_j\right) &= E\left(Z_i Z_j\right) \\ \frac{m_2}{c_1^2} E\left(\xi_i \xi_j \xi_k \xi_l\right) &= E\left(Z_i Z_j Z_k Z_l\right) \\ &\vdots \\ \frac{m_{\frac{1}{2}\sum_{j=1}^{k} r_j}}{c_1^{\frac{1}{2}\sum_{j=1}^{k} r_j}} E\left(\xi_{j_1}^{r_1} \ldots \xi_{j_k}^{r_k}\right) &= E\left(Z_{j_1}^{r_1} \ldots Z_{j_k}^{r_k}\right) \end{aligned} \tag{B.5}$$

Product moments appearing on the left hand side of equation (B.5) can be readily found, since they are the moments of a multivariate Gaussian distribution with covariance matrix equal to identity matrix $I_{J\times J}$.

Coefficients $m_1, m_2, m_3, \ldots$ are given in terms of $c_1, c_2, c_3, \ldots$ (and vice versa). Hence we have, by virtue of (B.2), identified requirements on all moments of (squared) generating variable R^2, so that the resulting multivariate distribution $\mathbf{X}$ has cumulant generating function (2.5).

Summarizing these results: First, since the multivariate Gaussian distribution referred to at equation B.5 has covariance matrix equal to identity, one can write for any set of components $(j_1, \ldots, j_k)$,

$$\frac{m_k}{c_1^k} E\left(\xi_{j_1}^2 \ldots \xi_{j_k}^2\right) = E\left(Z_{j_1}^2 \ldots Z_{j_k}^2\right) \tag{B.6}$$

where ξ is a J-dimensional normally distributed vector with mean vector $\mathbf{0}$ and covariance matrix $I_{J\times J}$, the identity matrix on $\mathbb{R}^{J\times J}$. Equation (B.2) holds in particular for vector ξ, in which case $R^2 \sim \chi_J^2$, and

$$\sum_{j_1=1}^{J} \cdots \sum_{j_k=1}^{J} E\left(\xi_{j_1}^2 \ldots \xi_{j_k}^2\right) = E\left(\left(\chi_J^2\right)^k\right) = \frac{2^k \Gamma\left(k + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)}$$

Second and more importantly, by virtue of (B.6), one can re-write (B.2) as

$$E\left(\left(R^2\right)^k\right) = \sum_{j_1=1}^{J} \cdots \sum_{j_k=1}^{J} \frac{m_k}{c_1^k} E\left(\xi_{j_1}^2 \ldots \xi_{j_k}^2\right) = \frac{m_k}{c_1^k} \frac{2^k \Gamma\left(k + \frac{J}{2}\right)}{\Gamma\left(\frac{J}{2}\right)} \tag{B.7}$$

which expresses the moments of R^2 in terms of parameters m_k (hence indirectly of c_k) and the dimension of the random vector $\mathbf{X}$.

References

Abramowitz, M. (1972). *Handbook of Mathematical Functions: with formulas, graphs, and mathematical tables.* Dover Publications, New York.

Banerjee, S., Gelfand, A. E., and Carlin, B. P. (2003). *Hierarchical modeling and analysis for spatial data.* CRC Press.

Bárdossy, A. and Pegram, G. (2009). Copula based multisite model for daily precipitation simulation. *Hydrology and Earth System Sciences Discussions*, 6(3):4485–4534.

Bárdossy, A. and Pegram, G. (2012). Multiscale spatial recorrelation of RCM precipitation to produce unbiased climate change scenarios over large areas and small. *Water Resources Research*, 48(9).

Barndorff-Nielsen, O. E. and Cox, D. R. (1990). *Asymptotic techniques for use in statistics*. Chapman and Hall, London [u.a.].

Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11(3):368 – 385.

Cressie, N. and Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Wiley.

Cressie, N. A. C. (1991). *Statistics for spatial data*. Wiley series in probability and mathematical statistics. Wiley, New York.

Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas. *International Statistical Review*, 73(1):111–129.

Diggle, P. J. and Ribeiro, P. J. (2007). *Model-based Geostatistics*. Springer.

Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*, volume 57. CRC press.

Kano, Y. (1994). Consistency property of elliptic probability density functions. *Journal of Multivariate Analysis*, 51(1):139 – 147.

Kendall, M. G. and Stuart, A. (1969). *The advanced theory of statistics Vol. 1, Distribution theory*. Griffin, London.

Kojadinovic, I. and Yan, J. (2011). A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. *Statistics and Computing*, 21(1):17–30.

Kolassa, J. (2006). *Series approximation methods in statistics*, volume 88. Springer.

Kolassa, J. and Li, J. (2010). Multivariate saddlepoint approximations in tail probability and conditional inference. *Bernoulli*, 16(4):1191–1207.

Kolassa, J. E. and Tanner, M. A. (1994). Approximate conditional inference in exponential families via the gibbs sampler. *Journal of the American Statistical association*, 89(426):697–702.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t-distributions and their applications*. Cambridge University Press.

Le, N. D. and Zidek, J. V. (2006). *Statistical analysis of environmental space-time processes*. Springer series in statistics. Springer, New York.

Ripley, B. D. (1981). *Spatial statistics*. Wiley series in probability and mathematical statistics. Wiley, New York.

Rodríguez, J. and Bárdossy, A. (2013). Modeling of multivariate interactions through their manifestations and low dimensional model building via the cumulant generating function. Technical report, Institute for Modelling Hydraulic and Environmental Systems, Universität Stuttgart.

Sanso, B. and Guenni, L. (1999). Venezuelan rainfall data analysed by using a bayesian space-time model. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 48(3):345–362.

Skovgaard, I. (1987). Saddlepoint expansions for conditional distributions. *Journal of Applied Probability*, pages 875–887.

Smith, P. J. (1995). A recursive formulation of the old problem of obtaining moments from cumulants and vice versa. *The American Statistician*, 49(2):217–218.

Steyn, H. (1993). On the problem of more than one kurtosis parameter in multivariate analysis. *Journal of Multivariate Analysis*, 44(1):1 – 22.

Villasenor Alva, J. A. and Estrada, E. G. (2009). A generalization of shapiro–wilk's test for multivariate normality. *Communications in Statistics Theory and Methods*, 38(11):1870–1883.