Distribution of the maximum in air pollution fields

Sofia Aberg[1] and Peter Guttorp[2]

[1] Lund University, Lund, Sweden
[2] University of Washington, Seattle, USA

Air quality standards are set to protect public health. The values of the standards are often based on health effect studies, without any statistical considerations. In order to judge if a standard is met measurements of ambient air quality are taken at monitoring stations, and these measured values are used to decide whether or not the standard has been violated. In this paper we examine the statistical quality of some air quality standards by taking both measurement error and variability of the ambient field away from the monitoring sites into account. In particular we study the distribution of the maximum of the ambient field conditional on a measured monitoring value at the value prescribed by the standard. The distribution of the maximum is computed using Rice's method and relies on a generalization of upcrossings of a level in one dimension to two dimensions.

KEY WORDS: Air quality standards, Distribution of the maximum, Random fields, Rice's formula.

1 Introduction

Air quality standards are set by governments to protect public health, see e.g. US Clean Air Act (1990) and COM (2000). Typically, the standard is a value which the governing agency feels public exposure should not be exceeded too often.

Examples:

- The European Union standard for sulphur dioxide, implemented on January 1, 2005, has two standard values: a 1-hour-average maximum of 350 g/m^3, and a 24-hour-average maximum of 125 g/m^3. The health effects studies on which this standard is based were summarized by Bertollini et al. (1997) in a World Health Organization publication.

- The United States Environmental Protection Agency (US EPA) standard for ozone, implemented on June 15, 2004, is based on an maximum 8-hour-average of 0.08 ppm.The previous standard, which is being phased out, was a maximum 1-hour-average of 0.12 ppm.

In order to judge whether the standard has been violated, measurements of ambient air quality are taken at some monitoring stations at regular times. Most implementations of air quality standards use the measured values at these monitoring sites to decide whether or not the standard has been violated. This approach can have some difficulties:

- Most people do not live at or even near the monitoring site.
- Due to measurement error, observed values are not actual values.
- Ambient air is only part of the actual exposure of any individual.

Due to these difficulties, people may be exposed to dangerous concentrations of pollutants, although the standard is attained. If this is the case, the implementation of the standards needs to be changed in order to protect people.

There has been some work on the statistical quality of air quality standards. Barnett and O'Hagan (1997), in a Royal Commission Report, discussed a variety of environmental standards, and developed a theoretical framework for setting standards which included a statistical quality assessment. Cox et al. (1999) and Thompson et al. (2002) discussed the US EPA ozone standard, as did Guttorp (2006). Among the main findings of these authors were that the effect of measurement error, temporal and spatial correlation made the ozone standard, both the old 1-hour standard and the new 8-hour standard, inadequate for protecting individuals from the intended exposure limit, even if the monitoring stations could be considered representative for the actual ambient pollution field.

None of these studies took into account the variability of the ambient field away from the monitoring sites. This is an important issue to consider since it is closely related to the first difficulty stated above, namely that people do not live exactly at the monitored sites. For example, suppose that a certain measurement is attained at one of the monitors. Given that measurement, what concentrations levels could then be expected in a region around the monitoring site where people actually live? Due to the variability of the ambient field, these concentrations may be much higher than the observed ones. The purpose of this paper is to answer such questions and to illuminate the effects of taking the variability of the ambient field away from the monitoring sites into account. In particular, we will study this by looking at the maximum of the ambient field, conditional on a measured monitoring value being at the value prescribed by the standard, i.e. the site is just inside the region of compliance with the standard.

In order to compute the distribution of the maximum, conditioned on some measurements, we use Rice's method. The methodology we propose can be used for a fairly large class of air pollution models, although it relies on assumptions of Gaussianity. The theoretical foundations of our methodology can be found in Azaïs and Delmas (2002) and in a recent paper by Mercadier(2005). Other approaches to approximate the distribution of the maximum are taken by Adler (1981), Sun (1993), Takemura and Kuriki (2002) and Piterbarg (1996).

The paper is organized as follows. First the modeling framework is presented in quite general terms that easily can be extended to different applied situations. Then we derive upper and lower bounds for the distribution of the maximum of a random field. To do so we first look at the corresponding problem in one dimension and then we generalize it, using the same kind of arguments, to two dimensions. Finally we consider two important examples, the first one dealing with the 1-hour standard for ozone and the second one with $PM_{2.5}$, i.e. particulate matter with median aerodynamic diameter of at most 2.5 μm, in southern California.

2 Model of the observed ambient pollution field

The general situation that we will consider is that we have measurements from an ambient air pollution field. Given these we want to compute the distribution of the maximum of the ambient field in a compact region of the geographical plane. For example, it might be of interest to study a region around the monitoring station where people actually live or an area for which, in some other sense, the monitored value is representative.

Let $\xi(\mathbf{x})$ be a Gaussian random field representing either the ambient air pollution field or a transformed version of it. For example, $\xi(\mathbf{x})$ could be the square root or the logarithm of the ambient pollution field. The reason to allow for such transformations is that they are commonly used in air pollution models, since they make the Gaussian assumption more realistic. Further, denote the mean and covariance functions of $\xi(\mathbf{x})$ by $m_\xi(\mathbf{x})$ and $r_\xi(\mathbf{s}, \mathbf{x})$, respectively. Now assume that the field is being observed at the monitoring points $(\chi_1, \cdots, \chi_n) = \mathcal{X}$ with an independent additive Gaussian measurement error having mean zero and variance $\sigma^2_\varepsilon(\mathbf{x})$. If the vector of observations is denoted by $z(\mathcal{X})$, the vector of field values at the monitoring points by $\xi(\mathcal{X})$ and the vector of measurement errors by $\varepsilon(\mathcal{X})$, then

$$z(\mathcal{X}) = \xi(\mathcal{X}) + \varepsilon(\mathcal{X}),$$

i.e., the measurements are just the sum of the underlying pollution field and a measurement error. Writing $\sigma^2_\varepsilon(\mathcal{X})$ for the vector of measurement error variances, the covariance matrix of the observations becomes

$$\Sigma_\mathcal{X} = (r_\xi(\chi_i, \chi_j)) + \mathrm{diag}(\sigma^2_\varepsilon(\mathcal{X})),$$

where by $\mathrm{diag}(\sigma^2_\varepsilon(\mathcal{X}))$ is meant the diagonal matrix having the vector $\sigma^2_\varepsilon(\mathcal{X})$ as its diagonal elements.

The dependence between the observations and the field at an arbitrary point $\mathbf{x}$ is, due to the Gaussian assumption, totally defined by the vector of covariances

$$r_\mathcal{X}(\mathbf{x}) = \mathrm{Cov}(z(\mathcal{X}), \xi(\mathbf{x})) = (r_\xi(\chi_i, \mathbf{x})),$$

and thus this vector contains valuable information about how the distribution of $\xi(\mathbf{x})$ is affected by observing the field at points $\mathcal{X}$. In fact, the random field $W(\mathbf{x})$ defined as the

(transformed) ambient field conditional on the observations, viz.

$$W(\mathbf{x}) = \xi(\mathbf{x}) \big| \xi(\mathcal{X}) + \varepsilon(\mathcal{X}) = z(\mathcal{X}) \tag{1}$$

is a Gaussian random field with mean given by

$$m(\mathbf{x}) = m_\xi(\mathbf{x}) + r_\mathcal{X}(\mathbf{x})^T \Sigma_\mathcal{X}^{-1}(z(\mathcal{X}) - m_\xi(\mathcal{X})) \tag{2}$$

and covariance function

$$r(\mathbf{s}, \mathbf{x}) = r_\xi(\mathbf{s}, \mathbf{x}) - r_\mathcal{X}(\mathbf{s})^T \Sigma_\mathcal{X}^{-1} r_\mathcal{X}(\mathbf{x}). \tag{3}$$

Note that if the covariance $r_\mathcal{X}(\mathbf{x})$ is different from zero the effect of the conditioning is that the mean $m(\mathbf{x})$ is adjusted in proportion to the measured values and that the variance $r(\mathbf{s}, \mathbf{s})$ gets smaller since more information is brought to our knowledge. From (3) it also follows that $W(\mathbf{x})$ is a non-homogeneous random field regardless whether the underlying process $\xi(\mathbf{x})$ is homogeneous or not.

Another interpretation of the conditional expectation (2) is that it is the best prediction of $\xi(\mathbf{x})$ given the observations, in the sense that it minimizes the mean squared error. In the geostatistical litterature this spatial prediction method is refered to as *kriging*, see e.g. Cressie (1993).

3 Distribution of the maximum in one dimension

Our main purpose is to examine the distribution of the maximum of a Gaussian random field in a compact region of the plane. In our case the field of interest is the air pollution field conditional on some measured values given by (1), but of course the theory is much more general than that. Before discussing this complicated two-dimensional problem it is appropriate to start with a simpler problem, namely its one-dimensional counterpart. Besides being a simpler problem, it will also be needed later when dealing with boundary behaviour in two dimensions.

Let $X(t)$ be a Gaussian process with $t \in [0, T]$, having differentiable sample paths, and denote its maximum on this interval by $M_T(X) = \max_{t \in [0,T]} X(t)$. If the maximum exceeds the level u then there are two possibilities, either the process starts above u or it starts below u and has at least one upcrossing of u, viz.

$$P(M_T(X) > u) = P(X(0) > u) + P(N_T(u) \geq 1, X(0) < u), \tag{4}$$

writing $N_T(u)$ for the number of upcrossings of u. Note that since $X(0)$ has a continuous distribution $P(N_T(u) \geq 1, X(0) \leq u) = P(N_T(u) \geq 1, X(0) < u)$ for each fixed u. Clearly the last term in (4) can be bounded from above by the expected number of upcrossings of u giving rise to the upper bound

$$P(M_T(X) > u) \leq P(X(0) > u) + \mathsf{E}[N_T(u)]. \tag{5}$$

This bound is good if the probability of having more than one upcrossing is small, and it is usually used for high levels u. To evaluate (5) the expected number of upcrossings must be computed. This problem was first considered in the pioneer work of Kac (1943) and Rice (1944, 1945) and the original result is that, for a zero mean stationary Gaussian process, the expected number of upcrossings in the interval $[0, 1]$ is given by

$$\mathsf{E}[N_1(u)] = \frac{1}{2\pi}\sqrt{\frac{\mathrm{Var}(X'(0))}{\mathrm{Var}(X(0))}}\exp\left(-\frac{u^2}{2\mathrm{Var}(X(0))}\right).$$

This classical formula, called Rice's formula, can be written in a much more general form, see e.g. Leadbetter et al. (1983), and it can also be used in a non-stationary setting. In fact it turns out that

$$\mathsf{E}[N_T(u)] = \int_0^T \int_0^\infty z f_{X'(t),X(t)}(z,u)\,dz\,dt = \int_0^T \mathsf{E}[X'(t)^+ \,|\, X(t) = u] f_{X(t)}(u)\,dt \qquad (6)$$

where $x^+ = \max(0, x)$. Now formula (6) provides a tool to compute the upper bound (5).

For moderately high levels the bound (5) is too crude and another approach is needed. One alternative is to study the time τ_u, say, when the process first upcrosses u. Since the event that the process starts below u and has at least one upcrossing of u is equivalent to the event that the time of the first upcrossing of u, τ_u, is in the interval $[0, T]$ we may rewrite (4) as

$$P(M_T(X) \geq u) = P(X(0) > u) + P(\tau_u \in [0, T]).$$

At first glance this does not seem to leed to any simplification of the problem. However it turns out that the density function of τ_u can be computed and is given by Durbin's formula, see Rychlik (1987), viz.

$$f_{\tau_u}(t) = \mathsf{E}\left[X'(t)^+ \mathbf{1}_{\{X(s)<u, \forall s<t\}} \,|\, X(t) = u\right] f_{X(t)}(u),$$

where $\mathbf{1}_{\{condition\}}$ is the indicator function equal to one if *condition* holds and zero otherwise. Thus we have the following equality for the distribution of the maximum on the interval $[0, T]$

$$P(M_T(X) \geq u) = P(X(0) > u) + \int_0^T \mathsf{E}[X'(t)^+ \mathbf{1}_{\{X(s)<u, \forall s<t\}} \,|\, X(t) = u] f_{X(t)}(u)\,dt. \qquad (7)$$

Numerical evaluation of this integral turns out to be possible and will be discussed later.

We have now presented two alternative ways of computing the distribution of the maximum for a Gaussian process having differentiable sample paths, the upper bound (5) and the expression (7). These can be used for high and moderate to high levels u respectively. Extensions to non-Gaussian processes are possible but then more restrictive assumptions are required. Next we will extend the arguments and the formulae to a two-dimensional framework.

4 Distribution of the maximum in two dimensions

4.1 Derivation of an upper bound

In the one dimensional case we argued that if the maximum exceeds level u then either the process starts above u or it starts below u and has at least one upcrossing of the level u. In the two dimensional case we can use similar arguments but we need to generalize the concept of upcrossings of a level.

Let $W(\mathbf{x})$ be a Gaussian random field, having continuously differentiable sample paths, where $\mathbf{x} = (x_1, x_2) \in \mathbf{S}$ is a space coordinate and $\mathbf{S}$ is a compact subset of $\mathbb{R}^2$ with boundary $\partial\mathbf{S}$. Sufficient conditions for a random field to have continuously differentiable sample paths are given in Adler (1981), Kent (1989) and Banerjee and Gelfand (2003), where continuity properties are related to the smoothness of the covariance function. Further, partial derivatives of $W(\mathbf{x})$ with respect to x_1 and x_2 will be denoted by $W_{10}(\mathbf{x})$ and $W_{01}(\mathbf{x})$, respectively. Analogously W_{20} and W_{02} are second derivatives with respect to x_1 and x_2, respectively, and W_{11} is the mixed second derivative.

If the maximum of the field $W(\mathbf{x})$ exceeds a threshold u, then u can either be exceeded on the boundary, or the field can stay below u on the boundary and have a local maximum higher than u in the interior of $\mathbf{S}$. Writing $M_{\mathbf{S}}(W) = \max_{\mathbf{x}\in\mathbf{S}} W(\mathbf{x})$ and $M_{\partial\mathbf{S}}(W) = \max_{\mathbf{x}\in\partial\mathbf{S}} W(\mathbf{x})$ we thus have

$$\begin{aligned} P\left(M_{\mathbf{S}}(W) > u\right) &= P\left(M_{\partial\mathbf{S}}(W) > u\right) + P\left(M_{\mathbf{S}}(W) > u, M_{\partial\mathbf{S}}(W) \leq u\right) \\ &= P\left(M_{\partial\mathbf{S}}(W) > u\right) + P\left(M_{\mathbf{S}}(W) > u, M_{\partial\mathbf{S}}(W) < u\right), \end{aligned} \tag{8}$$

where the last equality holds due to a lemma by Bulinskaya (1961) stating that under mild conditions a random process on the line is, with probability one, never tangential to a given level u in any finite interval. Now suppose that $M_{\partial\mathbf{S}}(W) < u$, that is the field is below level u on the boundary, and that the maximum is above u, i.e. $M_{\mathbf{S}}(W) > u$. Then, due to the smoothness assumptions, there must be at least one level curve $W(\mathbf{x}) = u$ in the interior of $\mathbf{S}$, and on each level curve there must be at least one point satisfying $W_{01}(\mathbf{x}) = 0$, $W_{02}(\mathbf{x}) \leq 0$ and $W_{10}(\mathbf{x}) \geq 0$. Such points, where the process upcrosses u in the x_1-direction, can be seen as a generalization of the upcrossings in one dimension to two dimensions. However, it should be emphasized that the choice to favour a certain direction is totally arbitrary and any other direction could have been chosen. Let $N_{\mathbf{S}}(A|\mathbf{u})$ denote the number of points in $\mathbf{S}$ such that $W(\mathbf{x}) = u$, $W_{01} = 0$ and the statement $A =$ "$W_{02}(\mathbf{x}) \leq 0, W_{10}(\mathbf{x}) \geq 0, W(\mathbf{s}) < u, \forall \mathbf{s} \in \partial\mathbf{S}$" is fulfilled. Then, following the arguments above, (8) can be rewritten as

$$\begin{aligned} P\left(M_{\mathbf{S}}(W) > u\right) &= P\left(M_{\partial\mathbf{S}}(W) > u\right) + P\left(N_{\mathbf{S}}(A|\mathbf{u}) \geq 1\right) \\ &\leq P\left(M_{\partial\mathbf{S}}(W) > u\right) + \mathsf{E}[N_{\mathbf{S}}(A|\mathbf{u})], \end{aligned} \tag{9}$$

which is a direct analogue to (4) and (5). In (9) there are two terms. The first one is a contribution from the boundary and will in principle be calculated using (7). The second term, which is a contribution from the interior of the region, will be computed by using a generalized version of Rice's formula presented next.

4.2 A generalized Rice's formula

Rice's formula, given by (6), can be extended to include so called marked crossings, see Leadbetter et al. (1983). This means that not only the occurence of a crossing is registered, but also the value of some other random variable connected with the crossing. For example, one may be interested in the expected number of upcrossings such that the derivative is larger than some specific value or, as is our case, the expected number of points such that $(W(\mathbf{x}), W_{01}(\mathbf{x})) = (u, 0)$ and a statement A concerning derivatives and process values is fulfilled. The following generalized Rice's formula, valid for Gaussian random fields, can be found in Mercadier (2005).

Theorem 1 *Let $\mathbf{Z}(\mathbf{x}) \in \mathbb{R}^2$ be an a.s. continuously differentiable Gaussian random field defined on a compact subset $\mathbf{S}$ of $\mathbb{R}^2$, whose boundary has Lebesgue measure zero. Denote by $\mathbf{Z}'(\mathbf{x})$ the matrix of partial derivatives of $\mathbf{Z}$ and by $f_{\mathbf{Z}(\mathbf{x})}(\mathbf{z})$ the density of $\mathbf{Z}(\mathbf{x})$. Further, let $\mathbf{Y}(\mathbf{x})$ be a random vector field, a.s. continuos and jointly Gaussian with $\mathbf{Z}(\mathbf{x})$. Let $g(\mathbf{Y}, \mathbf{x})$ be a positive, bounded and continuous function which can be a function of the whole sample path of $\mathbf{Y}$, not only of the value at a single point. If there are a.s. no points in the interior of $\mathbf{S}$ such that $\mathbf{Z}(\mathbf{x}) = \mathbf{u}$ and $\det(\mathbf{Z}'(\mathbf{x})) = 0$, then , for every $\mathbf{u}$,*

$$\mathsf{E}\left[\sum_{\{\mathbf{x}\in\mathbf{S};\mathbf{Z}(\mathbf{x})=\mathbf{u}\}} g(\mathbf{Y}, \mathbf{x})\right] = \int_{\mathbf{S}} \mathsf{E}\left[|\det(\mathbf{Z}'(\mathbf{x}))| g(\mathbf{Y}, \mathbf{x}) \middle| \mathbf{Z}(\mathbf{x}) = \mathbf{u}\right] f_{\mathbf{Z}(\mathbf{x})}(\mathbf{u})\, d\mathbf{x}, \qquad (10)$$

where both members are finite.

Remark 1 *To prove generalized versions of Rice's formula that holds for almost every $\mathbf{u}$ is often a fairly easy task and such results are valid under mild conditions on the process. However, to prove it for every $\mathbf{u}$ is usually much harder and more restricitve conditions are needed. To get from the almost everywhere result to the everywhere result it is sufficient to show that both sides of (10) are continuous functions of $\mathbf{u}$. In order to prove that, one uses the local inversion theorem explaining why the condition that there are a.s. no points in the interior of $\mathbf{S}$ such that $\mathbf{Z}(\mathbf{x}) = \mathbf{u}$ and $\det(\mathbf{Z}'(\mathbf{x})) = 0$ is needed. Sufficient conditions for $\mathbf{Z}(\mathbf{x}) = \mathbf{u}$ and $\det(\mathbf{Z}'(\mathbf{x})) = 0$ to hold are given in Azaïs and Wschebor (2005).*

Identify the processes $\mathbf{Z}$ and $\mathbf{Y}$ from the theorem with $\mathbf{Z}(\mathbf{x}) = (W(\mathbf{x}), W_{01}(\mathbf{x}))$ and $\mathbf{Y}(\mathbf{x}) = (W_{02}(\mathbf{x}), W_{10}(\mathbf{x}), W(\mathbf{x}))$, respectively. Further, let $\mathbf{u} = (u, 0)$ and let $A =$ "$W_{02}(\mathbf{x}) \leq 0, W_{10}(\mathbf{x}) \geq 0, W(\mathbf{s}) < u, \forall \mathbf{s} \in \partial\mathbf{S}$" be a condition on $\mathbf{Y}(\mathbf{x})$ and define an indicator function by $g(\mathbf{Y}, \mathbf{x}) = \mathbf{1}_{\{A\}}$. With this notation we have

$$\mathsf{E}\left[N_{\mathbf{S}}(A|\mathbf{u})\right] = \mathsf{E}\left[\sum_{\{\mathbf{x}\in\mathbf{S};\mathbf{Z}(\mathbf{x})=\mathbf{u}\}} g(\mathbf{Y}, \mathbf{x})\right].$$

However, the theorem cannot be used straight away since the indicator function is clearly not a continuous function. In order to show that the theorem remains true with $g(\mathbf{Y}, \mathbf{x}) = \mathbf{1}_{\{A\}}$

one can pick a sequence $\{g_n\}$ of continuous functions that turns to $\mathbf{1}_{\{A\}}$ as $n \to \infty$, and use monotone and dominated convergence arguments to show that (10) holds with $g(\mathbf{Y}, \mathbf{x}) = \mathbf{1}_{\{A\}}$. Leaving out the details, we thus have

$$\mathsf{E}\left[N_{\mathbf{S}}(A|\mathbf{u})\right] = \int_{\mathbf{S}} \mathsf{E}\left[|\det(\mathbf{Z}'(\mathbf{x}))\mathbf{1}_{\{A\}}| \,\middle|\, \mathbf{Z}(\mathbf{x}) = \mathbf{u}\right] f_{\mathbf{Z}(\mathbf{x})}(\mathbf{u})\, d\mathbf{x}. \tag{11}$$

Now, note that since

$$\mathbf{Z}'(\mathbf{x}) = \begin{pmatrix} W_{10}(\mathbf{x}) & W_{01}(\mathbf{x}) \\ W_{11}(\mathbf{x}) & W_{02}(\mathbf{x}) \end{pmatrix}$$

and it is conditioned on that $W_{01} = 0$ it holds that $|\det(\mathbf{Z}'(\mathbf{x}))| = |W_{10}(\mathbf{x})W_{02}(\mathbf{x})|$ so that (11) gives

$$\begin{aligned} \mathsf{E}[N_{\mathbf{S}}(A|\mathbf{u})] &= \int_{\mathbf{S}} \mathsf{E}\left[|W_{10}(\mathbf{x})W_{02}(\mathbf{x})|\mathbf{1}_{\{A\}} \,\middle|\, W_{01}(\mathbf{x}) = 0, W(\mathbf{x}) = u\right] f_{W_{01}(\mathbf{x}),W(\mathbf{x})}(0,u)\, d\mathbf{x} \\ &= \int_{\mathbf{S}} \mathsf{E}\left[W_{10}(\mathbf{x})^{+}W_{02}(\mathbf{x})^{-}\mathbf{1}_{\{W(\mathbf{s})<u,\forall \mathbf{s}\in\partial\mathbf{S}\}} \,\middle|\, W_{01}(\mathbf{x}) = 0, W(\mathbf{x}) = u\right] f_{W_{01}(\mathbf{x}),W(\mathbf{x})}(0,u)\, d\mathbf{x}. \end{aligned} \tag{12}$$

Thus the generalized Rice's formula provides a tool to compute the second term in the upper bound (9).

4.3 A lower bound

So far we have discussed how to compute an upper bound for the distribution of the maximum. However, to get an idea about how close we are to the true distribution it would be useful to also have a lower bound. In this way the difference between the bounds can be used as an indicator of the accuracy of the approximation.

To get a lower bound, choose N points that are inside $\mathbf{S}$ and call the set of those points $\mathbf{S}_N$. Then

$$P\left(M_{\mathbf{S}}(W) > u\right) \geq 1 - P(W(\mathbf{x}) \leq u, \forall \mathbf{x} \in \mathbf{S}_N), \tag{13}$$

constitutes a lower bound. The higher value of N the closer the lower bound will be to the true distribution. How the upper and lower bounds can be evaluated is the topic of the next section.

5 Computing the upper and lower bounds

In the previous sections the upper bound (9) for the distribution of the maximum for a random field was derived and theoretical results that show how to compute this bound were given in terms of (7) and (12). However, to be able to compute these bounds numerically they have to be put in a discrete form that is suitable for numerical evaluation. We start by having a closer look at the first term in (9), which is a contribution from the boundary, and then we turn our attention to the contribution from the interior of the region, i.e. the second term. After that we discuss how to numerically evaluate both the upper and lower bounds presented earlier.

5.1 Contribution from the boundary

Suppose that the boundary $\partial\mathbf{S}$ is a closed curve γ in $\mathbb{R}^2$, having a parametrisation with parameter θ taking values in $[0,\Theta]$. Define the process $X(\theta)$ as the process one gets by restricting the field $W(\mathbf{x})$ to the boundary of $\mathbf{S}$, i.e. $X(\theta) = W(\gamma(\theta))$. The process $X(\theta)$ is a process of one parameter and hence the one-dimensional result (7) can be used to compute the probability that its maximum exceeds a level u on the interval $[0,\Theta]$. Since $M_\Theta(X) = M_{\partial\mathbf{S}}(W)$ it follows that

$$P(M_{\partial\mathbf{S}}(W) \geq u) = P(X(0) > u) + \int_0^\Theta \mathsf{E}\left[X'(\theta)^+ \mathbf{1}_{\{X(\theta')<u, \forall \theta'<\theta\}} \middle| X(\theta) = u\right] f_{X(\theta)}(u)\, d\theta, \quad (14)$$

where $X'(\theta)$ is the directional derivative of the process $W(\gamma(\theta))$ in the direction of the tangent to the curve at θ. In practice, however, this formula can not be used directly because of the continuous path in the indicator function. To get a discretized version the interval $[0,\Theta]$ is replaced by a subdivision $\{\theta_1,\ldots,\theta_k\}$ and the indicator is approximated by $\mathbf{1}_{\{X(\theta_j)<u, \forall j:\theta_j<\theta\}}$. Since $\{X(\theta') < u, \forall \theta' < \theta\}$ is a subset of $\{X(\theta_j) < u, \forall j : \theta_j < \theta\}$, the discretization of the indicator in (14) gives rise to an upper bound, viz.

$$P(M_{\partial\mathbf{S}}(W) \geq u) \leq P(X(0) > u) + \int_0^\Theta \mathsf{E}\left[X'(\theta)^+ \mathbf{1}_{\{X(\theta_j)<u, \forall j:\theta_j<\theta\}} \middle| X(\theta) = u\right] f_{X(\theta)}(u)\, d\theta. \quad (15)$$

This is the formulation that we will use to compute an upper bound of the first term in (9).

5.2 Contribution from the interior of the region

Previously it was derived that the second contribution to the upper bound (9), i.e. the term $\mathsf{E}[N_{\mathbf{S}}(A|\mathbf{u})]$, is given by (12). Again, due to the continuous path in the indicator, this formula is not directly applicable to numerical computations. However, by introducing a subdivision $\{\mathbf{s}_1,\ldots,\mathbf{s}_n\}$ of the boundary, where $\mathbf{s}_i$ and $\gamma(\theta_i)$ are not necessarily equal, and approximating the indicator by $\mathbf{1}_{\{W(\mathbf{s}_j)<u, j=1,\ldots,n\}}$ an upper bound that is suitable for numerical evaluation is obtained, namely

$$\mathsf{E}[N_{\mathbf{S}}(A|\mathbf{u})] \leq$$
$$\int_{\mathbf{S}} \mathsf{E}\left[W_{10}(\mathbf{x})^+ W_{02}(\mathbf{x})^- \mathbf{1}_{\{W(\mathbf{s}_j)<u, j=1,\ldots,n\}} \middle| W_{01}(\mathbf{x}) = 0, W(\mathbf{x}) = u\right] f_{W_{01}(\mathbf{x}),W(\mathbf{x})}(0,u)\, d\mathbf{x}. \quad (16)$$

Altogether, combining the contributions from the boundary and from the interior of the region, an upper bound for the distribution of the maximum of a random field in a compact subset of the plane is given by

$$P\left(M_{\mathbf{S}}(W) > u\right) \leq P(X(0) > u) + \int_0^\Theta \mathsf{E}\left[X'(\theta)^+ \mathbf{1}_{\{X(\theta_j)<u, \forall j:\theta_j<\theta\}} \middle| X(\theta) = u\right] f_{X(\theta)}(u)\, d\theta +$$
$$\int_{\mathbf{S}} \mathsf{E}\left[W_{10}(\mathbf{x})^+ W_{02}(\mathbf{x})^- \mathbf{1}_{\{W(\mathbf{s}_j)<u, j=1,\ldots,n\}} \middle| W_{01}(\mathbf{x}) = 0, W(\mathbf{x}) = u\right] f_{W_{01}(\mathbf{x}),W(\mathbf{x})}(0,u)\, d\mathbf{x}. \quad (17)$$

5.3 Numerical evaluation of the upper and lower bounds

To evaluate the upper bound for the distribution of the maximum, given by (17), multivariate normal expectations must be computed. This can be done by using the routine `rind`, custom made for this type of computations, from the MATLAB toolbox WAFO[1], see Brodtkorb et al. (2000). The input to the `rind` function are the means and covariances of all the variables involved in the expectations in (17). Expressions for these are given in the appendix. In particular, if the underlying field $\xi(\mathbf{x})$ is assumed to be isotropic, the more explicit forms for this special case can be used.

The `rind`-function can also be used to compute multivariate normal probabilities. Thus it can be used to compute the lower bound (13). In our applications we chose to evaluate it with $N = 50$ and $N = 100$ points chosen at random within the region $\mathbf{S}$. These points are only chosen once and then used for all levels u for which the lower bound is evaluated for. The reason to choose at most 100 points to integrate over is that this is the upper limit the authors of the `rind` routine recommend. From our experience, however, the number of points could probably be increased and still give a reliable result for moderately high levels u.

An alternative to exact computation of the lower bound using `rind`, would be to use Monte Carlo simulation. To do so one should simulate the process in N points within the region $\mathbf{S}$ a large number of times, and then compute an empirical distribution of the maximum based on these simulated samples. The advantages with the simulation method is that it does not have the same kind of limitation in the number of points N that could be chosen. The drawback is that it would be much heavier computationally. However, for our purposes, we were satisfied with the results that we got by using low values of N and did not pursue the simulation.

6 Examples

In this section we will apply the previously derived bounds for the distribution of the maximum to air pollution fields. The purpose of these examples are twofold. Firstly they are intended to illustrate the methodology and secondly they illuminate the importance of considering the variability of the ambient field away from the monitored sites when designing the air quality standard. Thus, these examples should not be seen as a full solution to the standard setting problem, but rather as an illustration of the fact that there is reason for concern about some existing standards.

Two different examples will be presented. The first one deals with ozone in Harris County, Texas. In this example the simplest possible model for the ambient field is used and hence it should mainly be seen as an illustrating example and no major conclusions can therefore be drawn. However, we use this example to introduce a new way of thinking of the standard

[1] Available free of charge at http://www.maths.lth.se/matstat/wafo

setting problem. The second example, on $PM_{2.5}$ in the Los Angeles area, uses a much more realistic model showing the flexibility of the modeling framework. In this case we also investigate the sensitivity of the methodology to parameter uncertainties.

6.1 Example 1: Ozone

The US EPA 1-hour ozone standard, which is currently being replaced by a more stringent 8-hour standard, is that a region is in violation if the expected number of days per year with exceedances of 0.12 ppm is more than one. Consider a measurement that is exactly 0.12 ppm. Such a measurement would not be considered in violation of the standard. We are interested in the probability that someone in the region for which the measurement is considered representative of ambient air concentration is actually subjected to ambient concentrations in excess of $u > 0.12$ ppm. In particular, is it possible that anyone in that region has been subjected to concentrations as high as 0.18 ppm, considered a serious health risk by the US EPA? We need to define what region is representative for the measurement station, and define this, somewhat arbitrarily, as the region of the ambient field with correlation at least 0.7 with the measurement site.

6.1.1 Model of mean and covariance

Carroll et al. (1997) proposed a model for ozone in Harris County, Texas. Their model was quite involved and used temperature as covariate. Since this is an illustrating example we choose not to work with their full model but rather use a couple of their findings to make our example slightly more realistic. One of the features that we borrow from their work is that a square root transformation makes a Gaussian assuption more realistic. The other thing that we use from their work is the estimate of the measurement standard deviation, which was estimated to 0.020 $\text{ppm}^{1/2}$ at the square root scale for the year 1993. This value is also in good agreement of the measurement standard deviation found in Cox et al. (1999) for southern California data, which was $0.020 - 0.027$ $\text{ppm}^{1/2}$ at the square root scale.

In this example we use the simplest possible model, that is a model with constant mean and stationary isotropic covariance. To find out reasonable values for mean and variance we choose the annual mean and variance for a particular monitor in Harris County for 1993, which were $m_\xi(\mathbf{x}) = 0.235$ $\text{ppm}^{1/2}$ and $\sigma^2 = 0.0648^2$ ppm, according to the US EPA AIRS data base. What is left to specify is the covariance function. Here it is assumed that the covariance is of the Gaussian type, i.e. it has the parametric form $r_\xi(\mathbf{s}, \mathbf{x}) = \sigma^2 \exp(-||\mathbf{s} - \mathbf{x}||^2/\phi^2)$. Due to the smoothness properties of this covariance at the origin (it is in fact analytic), the smoothness properties of the random field and its derivatives required in Theorem 1 follow. In the expression for the covariance there is a parameter ϕ defining the range of the covariance. We have empirically found that if we choose the region over which the maximum is computed as the points being correlated at least 0.7 with the monitoring point, on the square root scale, then the result will not depend on the choice of ϕ. This

is thus an effect of how the region is chosen in this particular example and not a general result for this methodology. However, in the computations we used $\phi = 0.4$ length units, corresponding to a circular region having radius $r = \phi\sqrt{-\ln 0.7} \approx 0.24$ length units.

6.1.2 Computational issues

The region over which the maximum is taken is in this case a circular disc in the plane and its boundary can be parametrized using the angle $\theta \in [0, 2\pi]$ from the positive x-axis as parameter. In this example the parameter space is subdivided into 100 equally spaced points, $\{\theta_1, \ldots, \theta_{100}\}$, and for simplicity the same subdivision is used to compute the second integral in (17), i.e. $\{\mathbf{s}_1, \ldots, \mathbf{s}_n\} = \{\gamma(\theta_1), \ldots, \gamma(\theta_{100})\}$.

The computations are carried out on an equally spaced grid, constructed by placing 100 grid points in each direction over the square $[-0.5, 0.5]^2$, where the monitoring site is assumed to be at the origin. The integration is done by using the simplest possible numerical quadrature, i.e. by multiplying the value in each grid cell with its area and summing over all grid cells that falls in the integration region.

6.1.3 Distribution of the maximum

Having defined our model we are now ready to compute the distribution of the maximum in the region with correlation at least 0.7 with the measurement site, given a measurement that is exactly 0.12 ppm. However, to get a better understanding of what is actually computed, we start by showing what the integrands in (17), computed by `rind` for a fixed value of u, typically look like. In order to avoid confusion, values on the transformed scale are in the following denoted by u and values on the original scale by v. In Figure 1 (*Left*) the first integrand in (17), viz.

$$\mathsf{E}[X'(\theta)^+ \mathbf{1}_{\{X(\theta_j)<u, \forall j:\theta_j<\theta\}} \,|\, X(\theta) = u] f_{X(\theta)}(u)$$

is plotted against the parameter θ for $u = \sqrt{0.12}$ ppm$^{1/2}$, i.e. the standard level on the transformed scale. The function is clearly decreasing in a smooth way as the value of the parameter increases and by integrating this function the boundary contribution to the upper bound is obtained. In the same figure (*Right*) the second integrand of (17), namely

$$\mathsf{E}[W_{10}(\mathbf{x})^+ W_{02}(\mathbf{x})^- \mathbf{1}_{\{W(\gamma(\theta_j))<u, j=1,\ldots,100\}} \,|\, W_{01}(\mathbf{x}) = 0, W(\mathbf{x}) = u] f_{W_{01}(\mathbf{x}), W(\mathbf{x})}(0, u)$$

is shown. From this picture it is obvious that this function is not symmetric around the origin, although the problem that is solved is totally symmetric. This artifact is due to that a certain direction, the x_1-direction in this case, was chosen when the upcrossings from one dimension were generalized to two dimensions. Thus, if another direction would have been chosen another integrand had been obtained. In this case, however, since this problem is symmetric their integrals would have been the same and thereby also the upper bound.

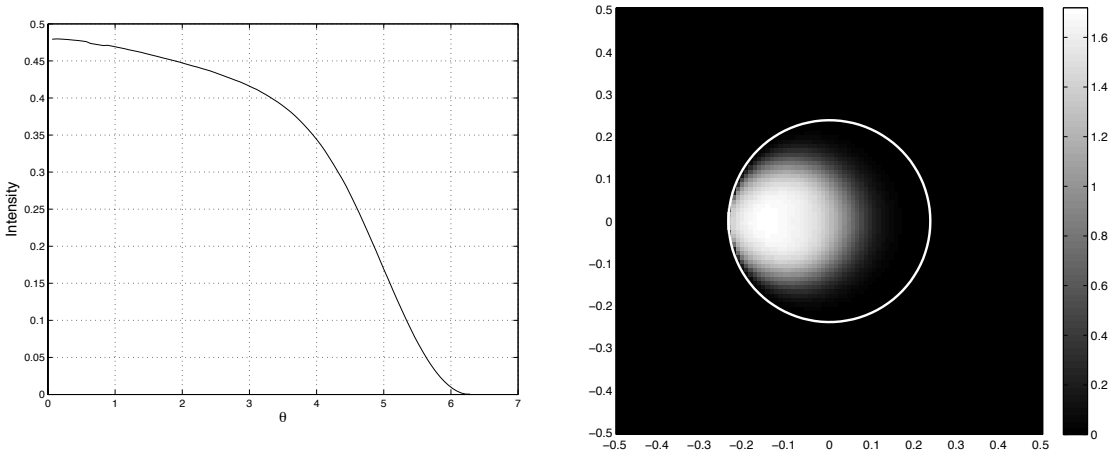

Figure 1: *Left:* The first integrand in (17) as a function of the angle θ. *Right:* Second integrand in (17). In both cases the integrands are computed for $u = \sqrt{0.12}$ ppm$^{1/2}$.

Unfortunately, for a non-symmetric problem, this may not be the case and the choice of direction can influence the upper bound.

Next we turn our attention to the distribution of the maximum. The upper and a lower bounds for the distribution of the maximum are shown on a linear and a log scale respectively in Figure 2. For the values of v of interest to us, i.e. for v up to 0.2 ppm, the result is very accurate. Also note that the lower bound approaches the upper bound as the number of points N increases as could be expected.

In order to study the implications of this distribution to the statistical quality of the ozone standard the important thing to study is the decay of the distribution. A fast decay would mean that the effect of the spatial variability of pollutant concentration is not so important to consider, whereas a slow decay would indicate that this really is something that should be taken into account. In this case the decay is pretty slow. In fact, there is almost as much as a 5% risk of exceeding 0.18 ppm, which is the limit for a serious violation and obviously considered as a very dangerous concentration level to be exposed to. Thus, in this case, the consequences of not considering the variability of the concentration field away from the monitor when designing the standard could be very serious.

So, how should then the standards be designed? One way of approaching that problem would be to first specify how large the probability of the maximum exceeding the level 0.18 ppm is allowed to be, and then setting the standard by finding out what observed value at the monitoring site this corresponds to. In Figure 3 the result from such a computation is shown. From this curve one sees that, if the probability of having a serious exceedance is to be decreased to for example 1% then the 1-hour standard level needs to be changed to around $0.095 - 0.10$ ppm.

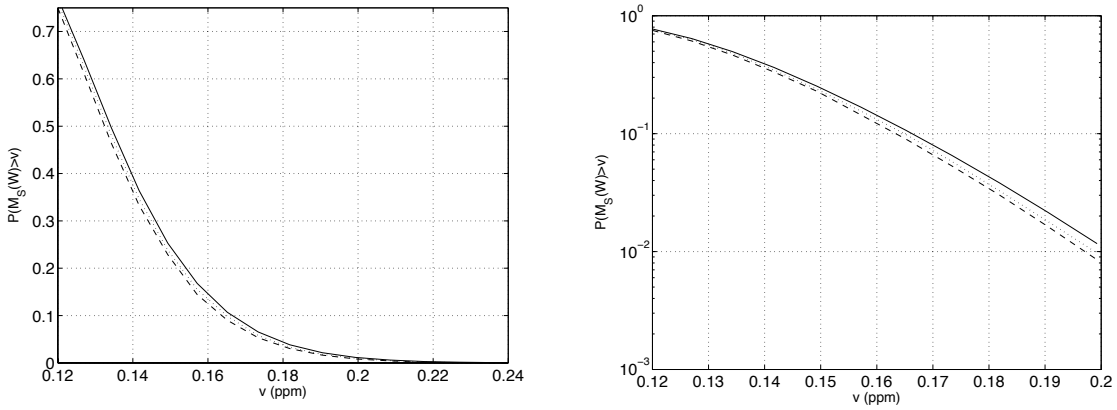

Figure 2: Upper (solid) and lower bounds, $N = 50$ (dashed) and $N = 100$ (dotted) for the probability of the maximum in the circular region to exceed level v on a linear scale to the *left* and on a logarithmic scale to the *right*.

6.2 Example 2: $PM_{2.5}$ in southern California

The US EPA has a 24-hour standard for $PM_{2.5}$. To check whether this standard is met one should, for each station, compute the 98th percentile of daily measurements (24-hour averages) over a year and then compute the average of such 98th percentiles for three consecutive years. If this average is equal to or below 65 $\mu g/m^3$ the standard is attained, otherwise not. The health effects caused by exposure to $PM_{2.5}$ are not totally clear, but the latest US studies indicate that serious health effects can result from PM_{10} exposure to as little as 25 $\mu g/m^3$, see the Staff document US EPA (2005) used in the latest revision of the PM standard.

In this example a more sophisticated model, described below, will be used. This model was fitted using four years of daily observations from 21 monitoring stations in the Los Angeles region, see the left part of Figure 5. One of the stations, station number 3, is located in central Los Angeles and in this example we will consider the distribution of the maximum in a circle centered at this station with a radius of about 20 miles, covering most of Los Angeles city. The important question that we will answer is: How high concentrations of $PM_{2.5}$ can anyone in Los Angeles be exposed to given that the standard level 65 $\mu g/m^3$ is observed at station number 3 on a particular day?

6.2.1 Model of mean and covariance

To model the ambient $PM_{2.5}$ concentration field a space-time model described in Guttorp et al. (2006) will be used. Here, for completeness, we review the main features of the model but for any details regarding the model or how it is fitted we refer to their work. Since the model is in space and time the pollution field will be denoted by $\xi(\mathbf{x}, t)$, where t is time. However, when the model is employed to compute the distribution of the maximum a fixed

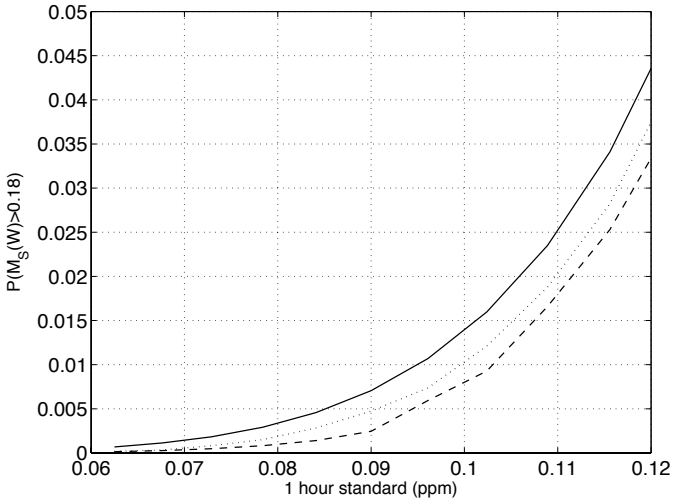

Figure 3: The probability that the maximum in the region exceeds 0.18 ppm for different 1 hour standards. Upper bound (solid) and lower bounds, $N = 50$ (dashed) and $N = 100$ (dotted).

time point t_0, say, is considered and then it is understood that $\xi(\mathbf{x}) = \xi(\mathbf{x}, t_0)$. To obtain the parameters used in this example the model was fitted using four years of daily measurements from 21 monitoring stations in southern California, see the left part of Figure 5.

Guttorp et al. (2006) found that after a logarithmic transformation the data was well approximated by a Gaussian distribution. Therefore, in their model, they choose the field $\xi(\mathbf{x}, t)$ as the logarithm of the ambient $PM_{2.5}$ concentration field and it is assumed to be Gaussian. The mean function is modeled as a linear combination of smoothed temporal empirical orthogonal functions (EOFs), i.e. smoothed eigenvectors of the empirical covariance matrix, with spatially varying coefficients viz.

$$m_\xi(\mathbf{x}, t) = \beta_0(\mathbf{x}) + \sum_{j=1}^{J} \beta_j(\mathbf{x}) f_j(t),$$

where $f_j(t)$ are smoothed EOFs. The gain in using EOFs is that a small number of orthogonal functions can describe most of the trend in the data and in this example only two EOFs are used. In Figure 4 these two trend components are shown together with the mean function at the third monitoring station. From these pictures it is clear that the EOFs manage to capture the seasonal trend in the data and also that the seasonal pattern is allowed to change from year to year.

From the model fit the spatial coefficients $\beta_0(\mathbf{x})$, $\beta_1(\mathbf{x})$ and $\beta_2(\mathbf{x})$ are only specified at the monitored sites, i.e. the spatial points where the measurements are taken. To predict the trend at unmonitored sites the spatial coefficients are spatially interpolated using kriging, see Cressie (1993). However, for our purposes, the important thing is that the mean can be computed at any spatial point at any time point.

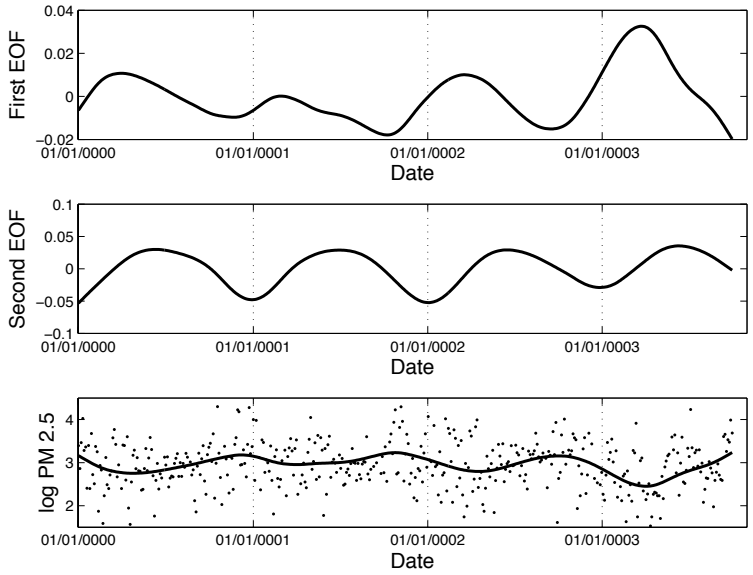

Figure 4: *Top:* The first annual trend component. *Middle:* The second annual trend component. *Bottom:* Measurements and fitted trend at monitoring station 3.

The underlying process $\xi(\mathbf{x}, t)$ is assumed to have a Sampson-Guttorp nonstationary spatial covariance, see Damian et al. (2001). This means that in the geographic coordinates the covariance is nonstationary but if the geographic plane is deformed using a function f, say, then, in the deformed geography, the covariance is approximately isotropic. If we assume that the covariance in the deformed plane is of the Gaussian type we can write

$$\mathrm{Cov}(\xi(\mathbf{s}, t), \xi(\mathbf{x}, t)) = \sigma^2 \exp\left(-||f(\mathbf{s}) - f(\mathbf{x})||^2/\phi^2\right). \tag{18}$$

The function f is modeled using a pair of thin-plate splines, see Bookstein (1989), which is a parametric function $\mathbb{R}^2 \mapsto \mathbb{R}^2$. The specific deformation f used in this example is depicted in Figure 5. The interpretation of this figure is that areas with high correlation are compressed while areas with low correlation are stretched out, so that the correlation in the deformed geography becomes isotropic. The other parameters, obtained by fitting the model to the data, used in this example are $\sigma^2 = 0.2910\ (\log(\mu g/m^3))^2$, $\phi = 2.0328$ and the variance of the measurement error is $\sigma_\varepsilon^2 = 0.0128\ (\log(\mu g/m^3))^2$. (The geographical coordinates of the monitor sites are given in so-called Lambert coordinates. These are obtained by locally projecting the longitude/latitude coordinates on the sphere onto a plane. Thus the unit for ϕ is the unit for distance in the Lambert-projected plane.)

6.2.2 Computational issues

The computations are, in principle, carried out in the same manner as in the ozone example. However, the model for mean and covariance that is used for PM concentrations makes the

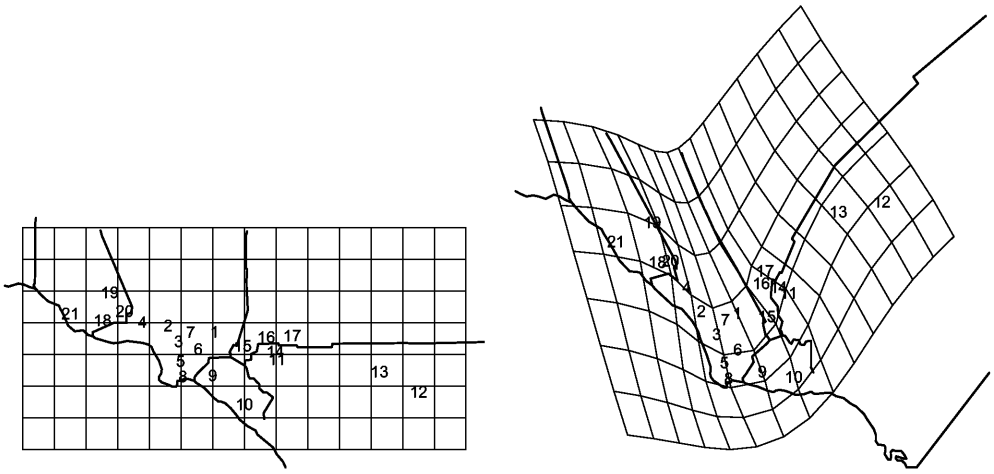

Figure 5: Spatial deformation representing nonstationary spatial covariance structure.

computations a bit more involved. In order to evaluate the distribution of the maximum, given by (17), spatial derivatives of the mean of the ambient pollution field are needed, see e.g. formulas (21)-(25) in the appendix. Since, by the way the model for the mean is constructed, the mean can be evaluated in every point in time and space so that these derivatives can be evaluated numerically by using e.g. symmetric differences. Another problem that must be dealt with is the non-stationary covariance model. To derive expressions for the derivatives of the covariance function (18), necessary to evaluate the distribution of the maximum (17), is quite complicated. However, in the deformed geometry, see Figure 5, the underlying field is isotropic and therefore the problem is greatly simplified if all computations are carried out in the deformed geometry and then transformed back.

6.2.3 Distribution of the maximum in Los Angeles

In this section the distribution of the maximum in the Los Angeles region, i.e. a circular region with radius 20 miles centered around station number 3, will be computed for a particular day. The day we choose to study is November 8, 2001. The reason for this choice is that the mean level has a peak at this date, see Figure 4, so that we may expect to get high measured values for this particular day. In fact, the measurement that was obtained exceeded the standard level 65 $\mu g/m^3$. In Figure 6 the mean around monitor station 3, according to the model, is shown together with the region over which the maximum will be computed. Note that this figure is plotted in deformed coordinates so that the circular region is transformed to a somewhat more elliptical shape.

Suppose that we got a measeurement exactly at the standard level 65 $\mu g/m^3$, corresponding to 4.17 $\log(\mu g/m^3)$ on the log-scale. Given that measurement, what is then the

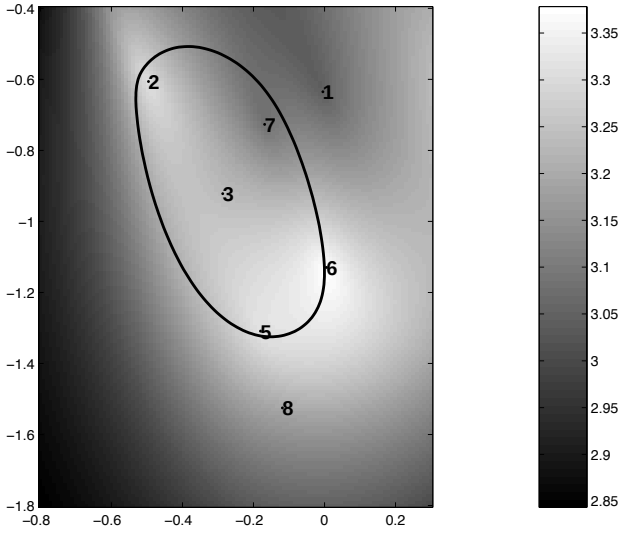

Figure 6: The mean surface according to the model at November 8, 2001, together with the Los Angeles region in transformed geographical coordinates. The concentrations are plotted on the log-scale and the x- and y-axis are in Lambert coordinates.

distribution of the maximum in Los Angeles? In Figure 7 upper and lower bounds for distribution of the maximum are shown. From this figure one can see that the decay of the distribution is quite slow. For example, the probability of having concentrations that are 30% higher than the standard is as much as 0.2. This again shows the importance of taking the variability of the field away from the monitoring sites into account. By only using the measured values without considering the spatial variability away from the monitoring site, one is not getting a good idea of what concentrations people actually are exposed to. Thus, we believe that there is major reason for concern about the $PM_{2.5}$ standard.

6.2.4 How much should the pollution concentration be decreased?

Los Angeles is a city with high concentration of $PM_{2.5}$, being a health threat to its inhabitants. Consequently it would be interesting to know how much the overall mean concentration in the city should be decreased in order for the concentrations not to exceed dangerous levels, at least with high probability. One of the advantages with our methodology is its flexibility which makes it possible to answer such questions. In Figure 8 the probability of exceeding the standard in Los Angeles given a measurement exactly at the standard level is shown for different values of reduction percentages of the overall mean surface. For example, one can see that in order to decrease the probability of exceeding the standard to 5%, the overall mean needs to be reduced to about a factor of $0.88 - 0.89$ of its current value.

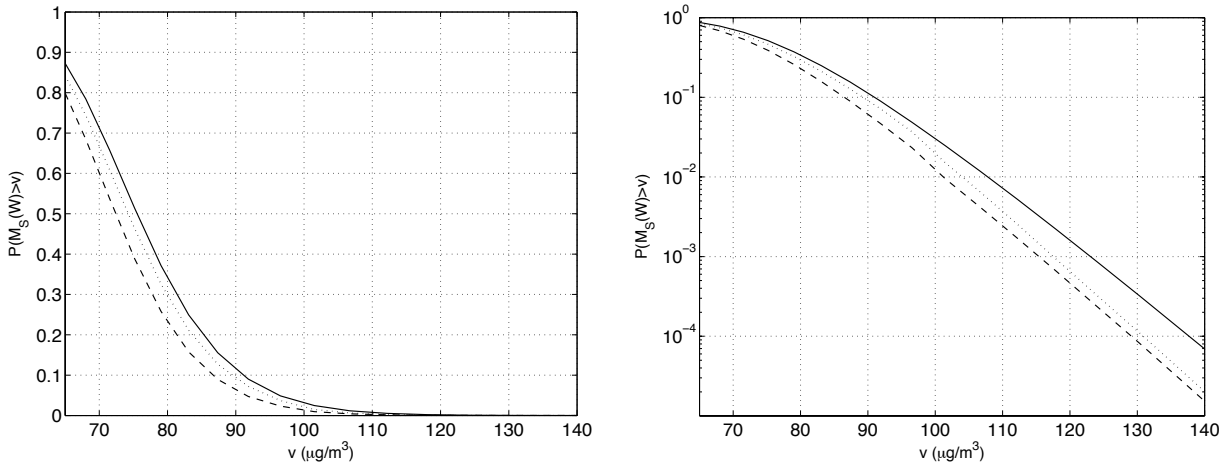

Figure 7: Upper (solid) and lower bounds, $N = 50$ (dashed) and $N = 100$ (dotted), for the probability of the maximum in the Los Angeles region to exceed level v on a linear scale to the *left* and on a logarithmic scale to the *right*.

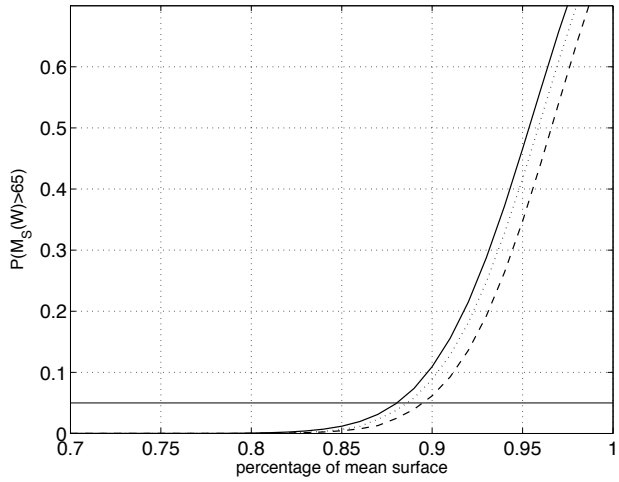

Figure 8: Upper (solid) and lower bounds, $N = 50$ (dashed) and $N = 100$ (dotted), for the probability of the maximum in the Los Angeles region to exceed the standard 65 $\mu g/m^3$ as the overall mean is reduced to a percentage of its current value.

6.2.5 Parameter uncertainty

So far nothing has been said on the sensitivity of this methodology to uncertainties in the parameter estimates. In this case the model for $PM_{2.5}$ concentration is given in a Bayesian framework, so that the parameters are estimated by sampling from a posterior distribution. Hence we could pick out individual samples from the posterior distribution and compute the distribution for the maximum for each of these samples and look at the spread of these. In Figure 9 this has been done for 10 samples from the posterior distribution, both for the upper bound and the lower bound based on $N = 100$ points. The figure shows that for lower levels the spread is quite large but that it gets smaller as higher levels are considered.

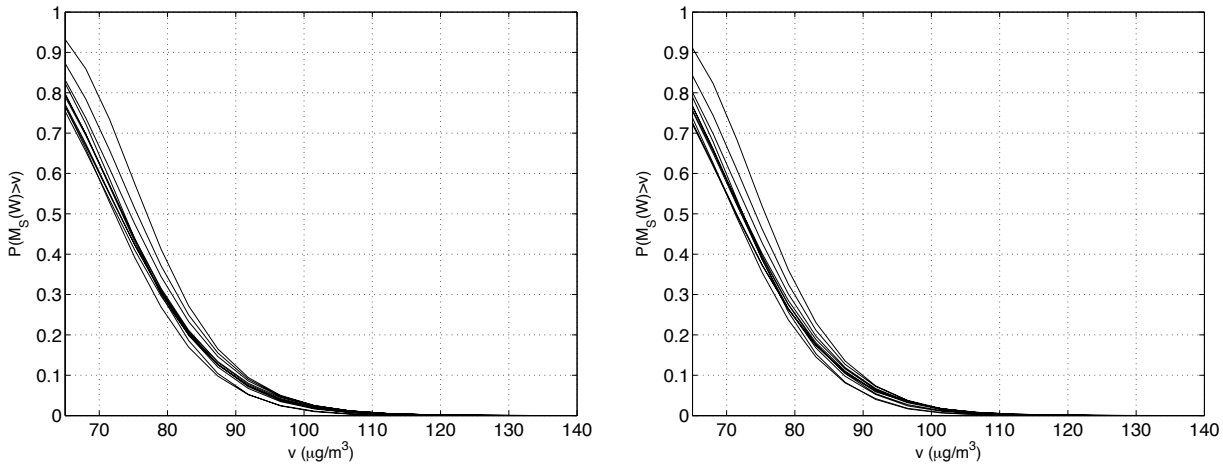

Figure 9: Upper (left) and lower bounds (right) for the survival function for 10 samples from the posterior distribution.

7 Summary and discussion

The purpose of this paper is to increase the understanding of the performance of environmental standards. In particular we want to illuminate what happens if the spatial variability of the pollution field away from the monitoring sites is taken into account. To do so we have been looking at the distribution of the maximum in a region that in some sense is representative of the monitoring site, conditional on a measurement being at the level prescribed by the standard. The methodology that we propose uses some recent results from the probability litterature and can be used in combination with a fitted statistical model and multivariate normal software.

The current implementations of the air quality standards are based on measurements from a network of monitoring stations. However, these implementations do not take into account that people do not live at, or sometimes even close, to the monitoring sites. In our case studies we show that with the current standards, the concentrations are with high

probability much higher than the intended exposure limits, in regions where people actually live. Thus we believe that there is reason to be concerned about the standards being set far too low.

In the case studies only a fixed time point is considered. However, it would be interesting to be able to compute a true spatio-temporal maximum so that one could say something about the maximum of the pollution levels during for example a year. That problem is of course much more complicated than the one considered in this work, and presumably it would require both heavier mathematical as well as computational tools.

8 Acknowledgements

Thanks to Professor I. Rychlik, Lund University, for useful dicussions and proofreading. Also we are indebted to Professor P.D. Sampson, University of Washington, for fitting the model in the second example. The first author was founded by the Swedish Foundation for Strategic Research, project Spatial statistics and image analysis for environment and medicine, A3 02:125. We are also grateful to the Swedish Organization of Graduate Engineers for supporting the second author as their Environmental Research Professor during part of his work.

A Mean and covariance

In this appendix, in order to make our presentation complete, we give full expressions for the means and covariances that are needed to compute the upper bound of the distribution of the maximum (17). First the means and covariances connected to the boundary term (15) are given and then those for the region term (16). Finally the expressions are written down in the special case of isotropy.

To compute the means and covariances frequent use is made of the differentiation rules for random fields and their relation to derivatives of the mean- and covariance functions. In general, if $W(\mathbf{x})$ is a random field with mean $m(\mathbf{x})$ and covariance function $r(\mathbf{s}, \mathbf{x})$ the following differentiation rules hold

$$\mathsf{E}[W_{ij}(\mathbf{x})] = \frac{\partial^{i+j} m(\mathbf{x})}{\partial x_1^i \partial x_2^j} \tag{19}$$

and

$$\mathrm{Cov}\left(W_{ij}(\mathbf{s}), W_{kl}(\mathbf{x})\right) = \frac{\partial^{i+j+k+l} r(\mathbf{s}, \mathbf{x})}{\partial s_1^i \partial s_2^j \partial x_1^k \partial x_2^l}. \tag{20}$$

A.1 On the boundary

Let $m(\mathbf{x}) = \mathsf{E}[W(\mathbf{x})]$ and $r(\mathbf{s}, \mathbf{x}) = \mathrm{Cov}(W(\mathbf{s}), W(\mathbf{x}))$ be the mean and covariance function of $W(\mathbf{x})$ and let $\gamma(\theta)$, $\theta \in [0, \Theta]$ be a parametrisation of the boundary of the region $\mathbf{S}$

satisfying $||\gamma'(\theta)||_2 = 1$. Moreover let $X(\theta)$ be the process $W(\mathbf{x})$ restricted to the boundary, i.e. $X(\theta) = W(\gamma(\theta))$, and let $\{\theta_1, \ldots, \theta_k\}$ be a subdivision of the interval $[0, \Theta]$. Writing $\gamma(\theta) = (x_1(\theta), x_2(\theta))$, the derivative $X'(\theta)$ may be expressed as

$$X'(\theta) = x_1'(\theta) W_{10}(\gamma(\theta)) + x_2'(\theta) W_{01}(\gamma(\theta)).$$

To be able to compute the upper bound (15) the joint distribution of the vector

$$X(\theta_1), \ldots, X(\theta_k), X'(\theta), X(\theta)$$

is needed. Since $W(\mathbf{x})$ is Gaussian the distribution will be Gaussian and by frequent use of (19) and (20) one can conclude that it has mean

$$\begin{aligned}
&\mathsf{E}\left[(X(\theta_1), \ldots, X(\theta_k), X'(\theta), X(\theta))\right] \\
&\quad = \left(m\left(\gamma(\theta_1)\right), \ldots, m\left(\gamma(\theta_k)\right), \left(x_1'(\theta) \frac{\partial m(\mathbf{x})}{\partial x_1} + x_2'(\theta) \frac{\partial m(\mathbf{x})}{\partial x_2} \right) \Big|_{\mathbf{x}=\gamma(\theta)}, m\left(\gamma(\theta)\right) \right),
\end{aligned}$$

and covariances

$$\begin{aligned}
\mathrm{Cov}(X(\theta_i), X(\theta_j)) &= r\left(\gamma(\theta_i), \gamma(\theta_j)\right), \ i = 1, \ldots, k, \ j = 1, \ldots, k \\
\mathrm{Cov}(X(\theta_i), X'(\theta)) &= \left(x_1'(\theta) \frac{\partial r(\mathbf{s}, \mathbf{x})}{\partial x_1} + x_2'(\theta) \frac{\partial r(\mathbf{s}, \mathbf{x})}{\partial x_2} \right) \Big|_{\mathbf{s}=\gamma(\theta_i), \mathbf{x}=\gamma(\theta)}, \ i = 1, \ldots, k \\
\mathrm{Cov}\left(X(\theta_i), X(\theta)\right) &= r\left(\gamma(\theta_i), \gamma(\theta)\right), \ i = 1, \ldots, k
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\left(X'(\theta)\right) &= \left(x_1'(\theta)^2 \frac{\partial^2 r(\mathbf{s}, \mathbf{x})}{\partial s_1 \partial x_1} + 2 x_1'(\theta) x_2'(\theta) \frac{\partial^2 r(\mathbf{s}, \mathbf{x})}{\partial s_1 \partial x_2} + x_2'(\theta)^2 \frac{\partial^2 r(\mathbf{s}, \mathbf{x})}{\partial s_2 \partial x_2} \right) \Big|_{\mathbf{s}=\mathbf{x}=\gamma(\theta)} \\
\mathrm{Cov}\left(X'(\theta), X(\theta)\right) &= \left(x_1'(\theta) \frac{\partial r(\mathbf{s}, \mathbf{x})}{\partial s_1} + x_2'(\theta) \frac{\partial r(\mathbf{s}, \mathbf{x})}{\partial s_2} \right) \Big|_{\mathbf{s}=\mathbf{x}=\gamma(\theta)} \\
\mathrm{Var}\left(X(\theta)\right) &= r\left(\gamma(\theta), \gamma(\theta)\right).
\end{aligned}$$

A.2 In the region

Let $\{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$ be a subdivision of the boundary $\partial\mathbf{S}$ and define $\mathbf{X}_n = (W(\mathbf{s}_1), \ldots, W(\mathbf{s}_n))$. Moreover let $Y(\mathbf{x}) = (W_{10}(\mathbf{x}), W_{02}(\mathbf{x}), W_{01}(\mathbf{x}), W(\mathbf{x}))$. Then the distribution of the vector $(\mathbf{X}_n, Y(\mathbf{x}))$ is multivariate Gaussian with mean

$$\mathsf{E}[(\mathbf{X}_n, Y(\mathbf{x}))] = (m_{\mathbf{X}_n}, m_Y(\mathbf{x})) = \left(m(\mathbf{s}_1), \ldots, m(\mathbf{s}_n), \frac{\partial m(\mathbf{x})}{\partial x_1}, \frac{\partial^2 m(\mathbf{x})}{\partial x_2^2}, \frac{\partial m(\mathbf{x})}{\partial x_2}, m(\mathbf{x}) \right),$$

and covariance matrix partitioned as

$$\Sigma(\mathbf{x}) = \begin{pmatrix} \Sigma_{\mathbf{X}_n} & \Sigma_{\mathbf{X}_n, Y}(\mathbf{x}) \\ & \Sigma_Y(\mathbf{x}) \end{pmatrix},$$

where it should be understood that the lower left part is the transpose of the upper right part. The different parts of the covariance matrix are

$$\Sigma_{\mathbf{X}_n} = (r(\mathbf{s}_i, \mathbf{s}_j)),\ i = 1, \ldots, n,\ j = 1, \ldots, n,$$

$\Sigma_{\mathbf{X}_n, Y}(\mathbf{x})$ is the matrix with rows

$$(\Sigma_{\mathbf{X}_n, Y}(\mathbf{x}))_{(i,\cdot)} = \left(\frac{\partial r(\mathbf{s},\mathbf{x})}{\partial x_1} \quad \frac{\partial^2 r(\mathbf{s},\mathbf{x})}{\partial x_2^2} \quad \frac{\partial r(\mathbf{s},\mathbf{x})}{\partial x_2} \quad r(\mathbf{s},\mathbf{x}) \right) \Big|_{\mathbf{s}=\mathbf{s}_i},\ i = 1, \ldots, n,$$

and

$$\Sigma_Y(\mathbf{x}) = \left(\begin{array}{cccc} \frac{\partial^2 r(\mathbf{s},\mathbf{x})}{\partial s_1 \partial x_1} & \frac{\partial^3 r(\mathbf{s},\mathbf{x})}{\partial s_1 \partial x_2^2} & \frac{\partial^2 r(\mathbf{s},\mathbf{x})}{\partial s_1 \partial x_2} & \frac{\partial r(\mathbf{s},\mathbf{x})}{\partial s_1} \\ & \frac{\partial^4 r(\mathbf{s},\mathbf{x})}{\partial s_2^2 \partial x_2^2} & \frac{\partial^3 r(\mathbf{s},\mathbf{x})}{\partial s_2^2 \partial x_2} & \frac{\partial^2 r(\mathbf{s},\mathbf{x})}{\partial s_2^2} \\ & & \frac{\partial^2 r(\mathbf{s},\mathbf{x})}{\partial s_2 \partial x_2} & \frac{\partial r(\mathbf{s},\mathbf{x})}{\partial s_2} \\ & & & r(\mathbf{s},\mathbf{x}) \end{array} \right) \Bigg|_{\mathbf{s}=\mathbf{x}}.$$

A.3 Isotropic case

In the case when the underlying field $\xi(\mathbf{x})$ is isotropic one can form general expressions for the means and covariances. Isotropy means that the covariance is just a function of distance $||\mathbf{s} - \mathbf{x}||$ and not on the exact locations of $(\mathbf{s}, \mathbf{x})$. Alternatively one can write the covariance as a function of the squared distance. Define

$$\hat{r}_\xi(h) = r_\xi(\mathbf{s}, \mathbf{x}),\ \text{where } h = ||\mathbf{s} - \mathbf{x}||^2 = (s_1 - x_1)^2 + (s_2 - x_2)^2,$$

and denote its derivatives by $\hat{r}'_\xi(h)$, $\hat{r}''_\xi(h)$, etc. To simplify notation introduce the following vectors:

$$\begin{aligned} r_{\mathcal{X}}(\mathbf{x}) &= (r_\xi(\chi_k, \mathbf{x})) = \left(\hat{r}_\xi(||\mathbf{x} - \chi_k||^2)\right) \\ r_{\mathcal{X}}^{(i)}(\mathbf{x}) &= \frac{\partial r_{\mathcal{X}}(\mathbf{x})}{\partial x_i} = \left(2(x_i - \chi_{k_i})\hat{r}'_\xi(||\mathbf{x} - \chi_k||^2)\right) \\ r_{\mathcal{X}}^{(i,i)}(\mathbf{x}) &= \frac{\partial^2 r_{\mathcal{X}}(\mathbf{x})}{\partial x_i^2} = \left(2\hat{r}'_\xi(||\mathbf{x} - \chi_k||^2) + 4(x_i - \chi_{k_i})^2 \hat{r}''_\xi(||\mathbf{x} - \chi_k||^2)\right). \end{aligned}$$

Under the isotropic assumption the mean and covariances from section A.1 for the process restricted to the boundary are given by

$$\mathsf{E}[X(\theta_i)] = m_\xi(\gamma(\theta_i)) + r_{\mathcal{X}}(\gamma(\theta_i))^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right),\ i = 1, \ldots, k \tag{21}$$

$$\begin{aligned} \mathsf{E}[X'(\theta)] = x_1'(\theta) &\left(\frac{\partial m_\xi(\mathbf{x})}{\partial x_1} + r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right) \right) \\ &+ x_2'(\theta) \left(\frac{\partial m_\xi(\mathbf{x})}{\partial x_2} + r_{\mathcal{X}}^{(2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right) \right), \end{aligned} \tag{22}$$

$$\mathsf{E}[X(\theta)] = m_\xi(\gamma(\theta)) + r_{\mathcal{X}}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right), \tag{23}$$

and

$$\begin{aligned}
\mathrm{Cov}\left(X(\theta_i), X(\theta_j)\right) &= \hat{r}_\xi(||\gamma(\theta_i) - \gamma(\theta_j)||^2) - r_{\mathcal{X}}(\gamma(\theta_i))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\gamma(\theta_j)),\ i,j = 1,\ldots,k \\
\mathrm{Cov}(X(\theta_i), X'(\theta)) &= x_1'(\theta) \left(-2(s_1 - x_1)\hat{r}_\xi'(||\mathbf{s}-\mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(1)}(\mathbf{x})\right)\Big|_{\mathbf{s}=\gamma(\theta_i),\mathbf{x}=\gamma(\theta)} \\
&\quad + x_2'(\theta) \left(-2(s_2 - x_2)\hat{r}_\xi'(||\mathbf{s}-\mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\mathbf{x})\right)\Big|_{\mathbf{s}=\gamma(\theta_i),\mathbf{x}=\gamma(\theta)} \\
\mathrm{Cov}\left(X(\theta_i), X(\theta)\right) &= \hat{r}_\xi(||\gamma(\theta_i) - \gamma(\theta)||^2) - r_{\mathcal{X}}(\gamma(\theta_i))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\gamma(\theta)) \\
\mathrm{Var}\left(X'(\theta)\right) &= -2\hat{r}_\xi'(0) - x_1'(\theta)^2 r_{\mathcal{X}}^{(1)}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(1)}(\gamma(\theta)) \\
&\quad - 2x_1'(\theta)x_2'(\theta) r_{\mathcal{X}}^{(1)}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\gamma(\theta)) - x_2'(\theta)^2 r_{\mathcal{X}}^{(2)}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\gamma(\theta)) \\
\mathrm{Cov}\left(X'(\theta), X(\theta)\right) &= -x_1'(\theta) r_{\mathcal{X}}^{(1)}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\gamma(\theta)) - x_2'(\theta) r_{\mathcal{X}}^{(2)}(\gamma(\theta))^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\gamma(\theta)) \\
\mathrm{Var}\left(X(\theta)\right) &= \hat{r}_\xi(0) - r_{\mathcal{X}}\left(\gamma(\theta)\right)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}\left(\gamma(\theta)\right).
\end{aligned}$$

The mean and covariances of section A.2 that are needed to evaluate the contribution from the region, that is the mean and covariance of the vectors $\mathbf{X}_n = (W(\mathbf{s}_1), \ldots, W(\mathbf{s}_n))$ and $Y(\mathbf{x}) = (W_{10}(\mathbf{x}), W_{02}(\mathbf{x}), W_{01}(\mathbf{x}), W(\mathbf{x}))$ become under isotropy

$$m_{\mathbf{X}_n} = \left(m_\xi(\mathbf{s}_1) + r_{\mathcal{X}}(\mathbf{s}_1)^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right), \ldots, m_\xi(\mathbf{s}_n) + r_{\mathcal{X}}(\mathbf{s}_n)^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right)\right) \tag{24}$$

and

$$\begin{aligned}
&m_Y(\mathbf{x}) \\
&= \left(\frac{\partial m_\xi(\mathbf{x})}{\partial x_1} + r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right),\ \frac{\partial^2 m_\xi(\mathbf{x})}{\partial x_2^2} + r_{\mathcal{X}}^{(2,2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right),\right. \\
&\quad \left.\frac{\partial m_\xi(\mathbf{x})}{\partial x_2} + r_{\mathcal{X}}^{(2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right),\ m_\xi(\mathbf{x}) + r_{\mathcal{X}}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} \left(z(\mathcal{X}) - m_\xi(\mathcal{X})\right)\right).
\end{aligned} \tag{25}$$

For the covariances one gets

$$\Sigma_{\mathbf{X}_n} = \left(\hat{r}_\xi(||\mathbf{s}_i - \mathbf{s}_j||^2) - r_{\mathcal{X}}(\mathbf{s}_i)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{s}_j)\right),\ i = 1,\ldots,n,\ j = 1,\ldots,n,$$

$$\begin{aligned}
(\Sigma_{\mathbf{X}_n,Y}(\mathbf{x}))_{(i,1)} &= -2(\mathbf{s}_{i_1} - x_1)\hat{r}_\xi'(||\mathbf{s}_i - \mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s}_i)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(1)}(\mathbf{x}) \\
(\Sigma_{\mathbf{X}_n,Y}(\mathbf{x}))_{(i,2)} &= 2\hat{r}_\xi'(||\mathbf{s}_i - \mathbf{x}||^2) + 4(\mathbf{s}_{i_2} - x_2)^2 \hat{r}_\xi''(||\mathbf{s}_i - \mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s}_i)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2,2)}(\mathbf{x}) \\
(\Sigma_{\mathbf{X}_n,Y}(\mathbf{x}))_{(i,3)} &= -2(\mathbf{s}_{i_2} - x_2)\hat{r}_\xi'(||\mathbf{s}_i - \mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s}_i)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\mathbf{x}) \\
(\Sigma_{\mathbf{X}_n,Y}(\mathbf{x}))_{(i,4)} &= \hat{r}_\xi(||\mathbf{s}_i - \mathbf{x}||^2) - r_{\mathcal{X}}(\mathbf{s}_i)^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{x}),\ i = 1,\ldots,n
\end{aligned}$$

and

$$\Sigma_Y(\mathbf{x}) = \begin{pmatrix} -2\hat{r}'_\xi(0) & 0 & 0 & 0 \\ & 12\hat{r}''_\xi(0) & 0 & 2\hat{r}'_\xi(0) \\ & & -2\hat{r}'_\xi(0) & 0 \\ & & & \hat{r}_\xi(0) \end{pmatrix}$$

$$- \begin{pmatrix} r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(1)}(\mathbf{x}) & r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2,2)}(\mathbf{x}) & r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\mathbf{x}) & r_{\mathcal{X}}^{(1)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{x}) \\ & r_{\mathcal{X}}^{(2,2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2,2)}(\mathbf{x}) & r_{\mathcal{X}}^{(2,2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\mathbf{x}) & r_{\mathcal{X}}^{(2,2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{x}) \\ & & r_{\mathcal{X}}^{(2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}^{(2)}(\mathbf{x}) & r_{\mathcal{X}}^{(2)}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{x}) \\ & & & r_{\mathcal{X}}(\mathbf{x})^T \Sigma_{\mathcal{X}}^{-1} r_{\mathcal{X}}(\mathbf{x}) \end{pmatrix}.$$

References

Adler, R.J. (1981). *The geometry of random fields.* Wiley, New York.

Azaïs, J.-M., and Delmas, C. (2002). Asymtotic expansions for the distribution of the maximum of Gaussian random fields. *Extremes*, **5**(2), pp. 181-212.

Azaïs, J.-M., and Wschebor, M. (2005). On the distribution of the maximum of a gaussian field with d parameters. *Annals of Applied Probability*. **15**(1), pp. 254-278.

Banerjee, S., and Gelfand, A.E. (2003). On smoothness properties of spatial processes. *J. Multivar. Anal.* , **84**, pp. 85-100.

Barnett V., and O'Hagan, A. (1997). *Setting environmental standards: the statistical approach to handling uncertainty and variation.* Chapman & Hall, London.

Bertollini, R., Dora, C., Krzyzanowski, M., and Stanners, D. (eds.) (1996). *Environment and Health 1–Overview and Main European Issues.* WHO Regional publications, European Series No. 68.

Bookstein, F.L. (1989). Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, **11**(6), pp. 567-585.

Brodtkorb, P.A., Johannesson, P., Lindgren, G., Rychlik, I., Rydén, J., and Sjö, E. (2000). WAFO–a Matlab toolbox for analysis of random waves and loads. In *Proceedings of the 10'th (2000) International Offshore and Polar Engineering Conference*, Vol. III.

Bulinskaya, E.V. (1961). On the mean number of crossings of a level by a stationary Gaussian process. *Teoriya Veroyatnostei i ee Primeeneiya*, **6**, pp. 474-477.

Carroll, R.J., Chen, R., George, E.I. (1997). Trends in ozone exposure in Harris County, Texas. *Journal of the American Statistical Association.*, **92**, pp. 392-415.

COM (2000). *Common Position paper of the Council of the European Union*, 10-10-2000.

Cox, L.H., Guttorp, P., Sampson, P.D., Caccia, D.C., and Thompson, M.L. (1999). A preliminary statistical examination of the effects of uncertainty and variability on environmental regulatory criteria for ozone. In *Novartis Foundation Symposium 220, Environmental statistics: analysing data for environmental policy.* Wiley, Chichester. pp. 122-143

Cressie, N. (1993). *Statistics for Spatial Data*, Wiley, New York.

Damian, D., Sampson, P.D., and Guttorp, P. (2001). Bayesian estimation of semi-parametric non-stationary spatial covariance structures, *Environmetrics*, **12**, pp. 161-178.

Guttorp, P. (2006). Setting environmental standards: A statisticians's perspective. To appear in *Environmental Geosciences*.

Guttorp, P., Fuentes, M., and Sampson, P.D. (2006). Using transforms to analyze space-time processes. Submitted to *SEMSTAT Proceedings of a workshop on space-time analysis*. Also available as NRCSE TRS 80 at *http://www.nrcse.washington.edu/pdf/trs80.pdf*

Kac, M. (1943). On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.* , **49**, pp. 314-320.

Kent, J.T. (1989). Continuity properties for random fields. *Annals of Probability*, **17**, pp. 1432-1440.

Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*, Springer Verlag, New York.

Mercadier C. (2005). Numerical bounds for the distribution of the maximum of one- and two-parameter processes. To appear in *Advances in Applied Probability*.

Piterbarg, V.I. (1996) *Asymptotic methods in the theory of Gaussian processes and fields*. American Mathematical Society, Providence, Rhode Island.

Rice, S.O. (1944). Mathematical analysis of random noise, *Bell Syst. Tech. J.* , **23**, pp. 292–332.

Rice, S.O. (1945). Mathematical analysis of random noise, *Bell Syst. Tech. J.* , **24**, pp. 46–156.

Rychlik, I. (1987). A note on Durbin's formula for the first passage density. *Statistics and Probability Letters*, **5**, pp. 425-428.

Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields. *Ann. Probab.*, **21**(1), pp. 34-71.

Takemura, A., and Kuriki, S. (2002). On the equivalence of the tube of Euler characteristic methods for the distribution of the maximum of Gaussian fields over piecewise smooth do-

mains. *Ann. Probab.*, **12**(2), pp. 768-796.

Thompson, M.L., Cox, L.H., Sampson, P.D., and Caccia, D.C. (2002). Statistical hypothesis testing formulations for U.S. environmental regulatory standards for ozone. *Environmental and Ecological Statistics*, **9**(4), pp. 321-339.

US Clean Air Act (1990). Available at *http://www.epa.gov/oar/caa/caa.txt*

US EPA (2005). Review of the national ambient air quality standards for particulate matter: policy assessment of scientific and technical information. Available at *http://www.epa.gov/ttn/naaqs/standards/pm/data/pmstaffpaper_20050630.pdf*

US EPA AIRS database. Available at *http://www.epa.gov/ttn/airs/airsaqs/index.htm*.