SPATIAL INTERPOLATION OF OZONE EXPOSURE IN NORWAY FROM SPACE-TIME DATA

GUDMUND HØST AND TURID FOLLESTAD
Norwegian Computing Center
Box 114 Blindern, N-0314 Oslo, Norway

Abstract. We propose a statistical method for spatial interpolation of ozone exposure in Norway, as measured by the accumulated dose above a threshold of 40 *ppb* (AOT40). Space-time ozone data from 10 stations in Norway during 1992 were modeled as the sum of a space-time mean field and a space-time residual field. The mean field was taken as a combination of temporal functions and spatial trend fields. We suggest a reduced rank method for spatial weighting of the temporal functions. After fitting variogram functions to component fields, the spatial AOT40 field was inferred by conditional simulation. Finally, the proposed model was checked by crossvalidation.

1. Introduction

During Summer, European emissions of nitrogen oxides and volatile organic compounds give rise to episodes of elevated ozone concentrations, which may prove harmful to humans and vegetation. Short-term peak concentrations are most harmful for humans, but vegetation is also damaged by long-term accumulated exposure at moderate ozone concentrations. We focus on the critical levels for forests, as measured by the exposure accumulated over a threshold of 40 *ppb* (AOT40). The recommended critical level given by UN-ECE (1994) to apply for forests is an AOT40 value of 10000 *ppb h*, accumulated during 1/4–30/9. In this study, we analyze space-time ozone data from Norway during 1/4–30/9 1992 to interpolate the spatial variable AOT40.

In the past years, space-time models have been fitted to atmospheric variables by several authors. To some extent, the approach taken depends on the objective of the study. For some alternative models, the reader may

consult Høst, Omre and Switzer (1995), Haas (1995), Sølna and Switzer (1996), Carroll, Chen, Li, Newton, Schmiediche and Wang (1997) and Wikle, Berliner and Cressie (1997).

The distribution of ozone in space and time is complex. Important sources of variability are emission and weather patterns, North–South trends in daylight, time of the day, local NO_x sources, and local vegetation. A common method for interpolating spatial data to unmonitored locations is Kriging (Cressie 1991). This approach was taken by Tørseth, Mortensen and Hjellbrekke (1996), who aggregated space-time ozone data to spatial AOT40-values and interpolated these values by Kriging. On the other hand, Anderson and Smith (1997) describe a normal random effects model for year to year values of AOT40 based on data from Great Britain and Ireland. They also state conditions under which AOT40 may be regarded as approximately normally distributed.

The interpolated values and associated interpolation errors obtained from Kriging may be unsatisfactory. At each spatial location AOT40 is a non-negative random variable, therefore it will have a skewed probability distribution. Kriging interpolation and Kriging interpolation errors may not be realistic for skewed data. Specifically, Kriging interpolation intervals are symmetric and could produce negative AOT40 quantiles. In contrast, we present a richer analysis, which specifically takes into account observed space-time features of the raw ozone data and also allows for skewness of AOT40. By using a space-time model, we hope to translate structure from the densely sampled time-domain to unobserved spatial locations. We propose to do this by using a flexible parametric model for the time-trend at the monitoring stations. This parametric time-trend model is selected by rank regression (Davies and Tso 1982). The time-trend coefficients at each monitoring station are interpreted as a realization of a set of spatial random fields and the residual data are modeled as a realization of a a space-time separable random field. Inference of ozone concentrations at for unobserved locations in space-time are done by condotional simulation in space-time. A conditional simulation of the spatial AOT40-field is obtained by time-integration of the simulated space-time ozone field, as will be elaborated in Section 3.2.

2. Norwegian Ozone Data

In this work we have analyzed Norwegian data for the period 1/4–30/9 1992, which is a period with unusually high ozone exposure in Southern Norway. Our data consist of ozone measurements from 10 monitoring stations, each with 2805–4392 hourly values.

Figure 1 shows time series for each monitoring station for the period.

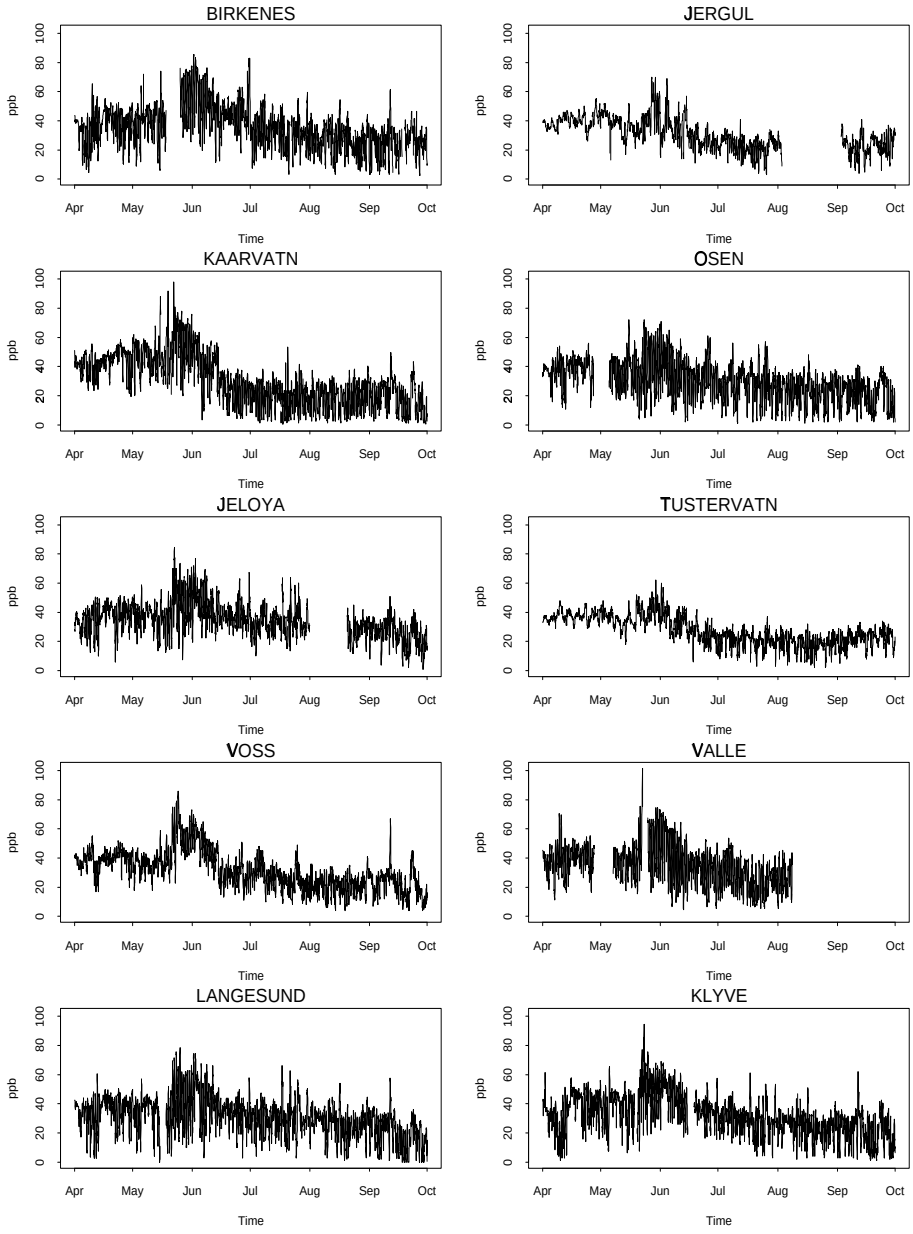

Figure 1. Observations at the monitoring stations for the study period of 1. April – 30. September, 1992.

Characteristic features are the generally high concentrations in April–June, and an episode of elevated concentrations in late May and early June.

All time series exhibit diurnal cycles in the ozone concentration, but the amplitude of these cycles varies considerably among the sites and through the year.

In contrast to Carroll et al. (1997), we did not see obvious signs of skewness in the raw ozone data. We will see later that skewness in AOT40 may still be significant.

3. Method

Denote by $y(\boldsymbol{x}, t)$ the ozone concentration at spatial location $\boldsymbol{x}$ and time t. Let $y(\boldsymbol{x}, t)$ have the decomposition

$$y(\boldsymbol{x}, t) = \boldsymbol{g}'(\boldsymbol{x})\, \boldsymbol{h}(t) \; + \; v(\boldsymbol{x}, t). \tag{1}$$

Here, $\boldsymbol{g}(\boldsymbol{x})$ is a p-vector of spatial trend fields, $\boldsymbol{h}(t)$ is a p-vector of temporal structure functions and $v(\boldsymbol{x}, t)$ is a zero-mean space-time residual field, independent of $\boldsymbol{g}$. The vector $\boldsymbol{h}(t)$ spans a wide class of temporal structure functions, chosen to reflect trend, seasonal and diurnal effects. Therefore, (1) describes a class of models, within which we will select a particular model for the present study. In general, one might include other covariates in the $\boldsymbol{g}$- or $\boldsymbol{h}$-functions.

The motivation behind the decomposition (1) is that systematic (trend) and non-systematic (residual) temporal fluctuations may have different spatial persistence. In particular, residual fluctuations in time may have only local influence, while systematic temporal fluctuations may vary slowly with geographic location. Thus, by modeling trend and residual separately we may obtain more precise inference for ozone concentrations and AOT40 at unobserved locations. Each $g_k(\boldsymbol{x})$; $k = 1, \ldots, p$ is taken as a realization of a second-order stationary random field. A feature of this approach is that uncertainty about the trend model will be specifically incorporated in the error estimate when interpolating AOT40 to unobserved locations. The second-order properties of the residual $v(\boldsymbol{x}, t)$ is specified by the space-time covariance function

$$\mathrm{Cov}\{v(\boldsymbol{x}', t'), v(\boldsymbol{x}'', t'')\} = \sigma^2\, \rho_x(\|\boldsymbol{x}' - \boldsymbol{x}''\|; d_x)\; \rho_t(|t' - t''|; d_t). \tag{2}$$

Here, σ^2 is the residual variance, while $\rho_x(\cdot; d_x)$ and $\rho_t(\cdot; d_t)$ are correlation functions with range parameters d_x and d_t, respectively. The separable form of the space-time correlation function was chosen for convenience, but is commonly used in space-time analysis, see Rodriguez-Iturbe and Mejia (1974), Sølna and Switzer (1996). We will see later that this particular choice allows for efficient simulation.

3.1. MODEL SELECTION AND PARAMETER ESTIMATION

The problem of model selection resembles the choice of the smoothing parameter in non-parametric or multivariate regression. Within these branches of statistics, it is well-known that restricted estimates may reduce parameter uncertainty and give better predictions at the cost of some bias (Frank and Friedman 1993). In our case, we select a rank parameter $K < p$, where $K + 1$ is the efficient number of spatial trend fields we will use. Given K, we have a specific model, and we obtain values of the $\boldsymbol{g}$-fields and v-field at the spatial and space-time locations, respectively. Now, the covariance parameters σ^2, d_x, and d_t may be estimated for the space-time residual field, and similar parameters are estimated to facilitate interpolation of the spatial trend fields. Variability not accounted for in the estimated trend will be specifically incorporated in the estimated space-time residual field.

We introduce matrix notation for the space time data

$$\boldsymbol{Y} = \boldsymbol{H}\boldsymbol{G} + \boldsymbol{V}.$$

The elements of these matrices are $(\boldsymbol{Y})_{ji} = y(\boldsymbol{x}_i, t_j)$; $i = 1, \ldots, n$; $j = 1, \ldots, m$, $(\boldsymbol{H})_{jk} = h_k(t_j)$; $j = 1, \ldots, m$; $k = 1, \ldots, p$, $(\boldsymbol{G})_{ki} = g_k(\boldsymbol{x}_i)$; $k = 1, \ldots, p$; $i = 1, \ldots, n$ and $(\boldsymbol{V})_{ji} = v(\boldsymbol{x}_i, t_j)$; $i = 1, \ldots, n$; $j = 1, \ldots, m$. We select the rank K by a procedure borrowed from reduced rank regression, as described by Davies and Tso (1982) and Aldrin (1996). The idea is to estimate $\boldsymbol{G}$ by constrained least squares, the constraint being that the estimate should have rank K. The usual procedure for selecting K when the residuals are uncorrelated is cross-validation, as described by Aldrin (1996). However, this may not work well if the residuals are correlated (Diggle and Hutchinson 1989, Hart and Wehrly 1986). A reasonable procedure in the present case is to choose the smallest possible K which allows for elimination of the main temporal structure in the residuals. Optimal procedures for model selection when the residuals are correlated in space-time is a topic for future research. The main ingredient of reduced rank regression is the singular value decomposition. Following Aldrin (1996), the reduced rank estimate of $\boldsymbol{G}$ may be expressed

$$\widetilde{\boldsymbol{G}} = \sum_{k=1}^{K} \boldsymbol{b}_k \boldsymbol{a}_k',$$

where $\boldsymbol{a}_k$ is given by singular value decomposition

$$(n\boldsymbol{H}'\boldsymbol{H})^{-1/2}\boldsymbol{H}'\boldsymbol{Y} = \sum_{k=1}^{r} \lambda_k \boldsymbol{u}_k \boldsymbol{a}_k'.$$

Here, r is the rank of $(n\boldsymbol{H}'\boldsymbol{H})^{-1/2}\boldsymbol{H}'\boldsymbol{Y}$, and $\boldsymbol{b}_k$ is given by

$$\boldsymbol{b}_k = \lambda_k(\frac{1}{n}\boldsymbol{H}'\boldsymbol{H})^{-1/2}\boldsymbol{u}_k.$$

Now, the rank K estimate of $\boldsymbol{g}$ at location $\boldsymbol{x}_i$ may be expressed

$$\widetilde{\boldsymbol{g}}(\boldsymbol{x}_i) = \sum_{k=1}^{K} a_k(\boldsymbol{x}_i)\boldsymbol{b}_k. \tag{3}$$

Here, $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K$ are ranked vectors of weights for the temporal structure functions. For each k we take the coefficients $a_k(\boldsymbol{x}_i); i = 1, \ldots, n$ as realizations of second order stationary spatial random fields with constant mean and we fit variograms to these fields. In geneal, the a_k-fields may be cross-correlated. In addition comes an intercept term, which we denote by $a_0(\boldsymbol{x})$. The advantage of (3) is that we can now fit spatial models to the $(K+1)$ a_k- fields instead of the p g_k-fields. This may be of great help if K can be chosen small. In the present case, we follow Aldrin (1996) and replace (3) by a slightly more general expression that allows for missing data.

The main focus of this study is not to study elaborate methods for parameter estimation, therefore we have used the following simple procedure. First, empirical residuals were calculated by subtracting estimated structure components $\boldsymbol{H}\widetilde{\boldsymbol{G}}$ from data values $\boldsymbol{Y}$. Then the spatial correlation range parameter d_x was estimated by fitting an exponential correlation function by the method of Cressie (1991, pp. 74–75). A similar procedure was used in the time domain to estimate d_t. To reduce effects of missing data, σ^2 was estimated as follows. For each observation time t_j, the usual bias-corrected maximum likelihood (Cressie 1991, pp. 91–92) estimate $\widehat{\sigma}_j^2$ was found using the available observations and plugging in the estimated spatial correlation matrix. Then $\widehat{\sigma}^2$ was taken as the weighted average of all the $\widehat{\sigma}_j^2$'s, the weights being proportional to the number of observations used at each time t_j. For the $a_0, \ldots, a_K$ coefficient fields a correlation range was fitted graphically to the empirical variogram. Then each variance parameter was fitted as explained above.

3.2. INTERPOLATION OF AOT40

The accumulated exposure above $40ppb$ $(80\mu g/m^3)$ is defined as

$$z(\boldsymbol{x}) = \int_T y(\boldsymbol{x}, t)\, I[y(\boldsymbol{x}, t) \geq 40ppb]\, dt,$$

where T is the growth period 1/4–30/9.

For the time period T, we generate realizations of the spatial AOT40-field by conditional simulation from the estimated statistical model given

the data $\boldsymbol{Y}$. After generating 1000 simulations, we infer the relevant statistics such as the spatial mean- and median-fields, confidence levels, exceedance probabilities, or excursions above high levels.

Since $v(\boldsymbol{x}, t)$ is a zero-mean space-time separable random field (Cressie 1991), we construct a realization from a spatial field having correlation function ρ_x and a temporal process having correlation function ρ_t. It is easily verified that the product of two such independent (zero mean) processes has the correlation structure given by (2). Thus, we obtain a great reduction of computational effort, because we do unconditional space-time simulation by simulating separately in the spatial and temporal dimensions. In this application, we have used two independent Gaussian processes, giving a space-time field which is second-order stationary.

The exponential correlation functions we used are well-suited for unconditional simulation by the screening sequential algorithm, see Omre, Sølna and Tjelmeland (1992). From the unconditional simulations, conditional simulations are generated as described in Cressie (1991, pp. 207–209). Each simulation consists of one spatial simulation for each of the a_k-fields and one space-time simulation of the residual field. The a_k-fields are conditioned on the estimated coefficients at the data locations, while the residual field is conditioned on estimated space-time residuals.

4. Results

The temporal basis functions $\boldsymbol{h}(t) = (h_1(t), \ldots, h_p(t))$ must be chosen to capture the main temporal structure of the ozone data. We have used $p = 85$, with temporal functions as follows. To capture diurnal variation, one indicator function for each hour of the day was introduced. Furthermore, a linear function was used to capture trend, trigonometric functions with long period were chosen to capture long-term variability, and sinusoids was used to account for the early Summer episode in Southern Norway. Finally, cross-terms of long-term and diurnal functions were used to reflect long-term fluctuations in the diurnal variation.

After deciding on the basis functions $\boldsymbol{h}(t)$, the reduced rank procedure was used to fit a weighted combination of basis functions to the data. We used rank $K = 2$, giving 3 spatial trend functions to fit. This choice seems to remove most of the temporal structure in the data, and it explains 46–79% of the total variance at the monitoring stations. Even with the complex trend model the space-time residuals are correlated, but this correlation is mainly positive and of short range. The main temporal structure component $\boldsymbol{b}_1'\boldsymbol{h}(t)$ (+constant) is shown in Figure 3, decomposed into diurnal and seasonal effects. We see that the main seasonal component (Figure 3a) has a maximum in early June, with decaying values towards the end of

the season. This maximum is a special feature of the year 1992 and is related to unusual meteorological conditions. The main diurnal component (Figure 3b), has a minimum in the early morning and a maximum in the afternoon. The seasonal diurnal amplitude (Figure 3c) has a maximum in late June.

An exponential correlation function $\rho(h;d) = \exp(-3h/d)$ was fitted to each of the a_0-, a_1- and a_2-fields. The fitted range parameter was $d = 750\,km$ for all three fields. The estimated space-time residual seemed to fit well to exponential covariance functions ρ_x and ρ_t. The fitted parameters were $\widehat{d}_x = 345\,km$, $\widehat{d}_t = 44\,hours$ and $\widehat{\sigma} = 9.5\,ppb$. Figure 2 (right panel) shows the estimated exceedance risk for the 10000 $ppb\,h$ critical level, based on 1000 simulations. The estimated exceedance risk is defined as $100 \times \widehat{p}$, where $\widehat{p}$ is the estimated exceedance probability. The map shows that exceedance/non-exceedance can only be determined accurately close to the monitoring stations, based on the present statistical analysis. On the other hand, the estimated risk of exceedance in Northern Norway is usually less than 25%, while for large areas in Southern Norway this risk is above 50%. The locations of the monitoring stations are shown in the left panel of Figure 2.

4.1. MODEL VALIDATION

First, we checked our interpolations against hourly concentrations of three monitoring stations not used in the analysis (shown by triangles in Figure 2). Interpolating the available observations at these stations by our model gave (true) root mean square interpolation error (RMSIE) of 20.5, 13.0 and 17.1 $\mu g/m^3$ for stations Nordmoen, Prestbakke and Svanvik, respectively. The estimated RMSIE from the statistical model were 21.2, 16.7 and 18.1 $\mu g/m^3$, which is very similar.

To further check the statistical model, the following cross-validation exercise was carried out. Each monitoring station was deleted from the data set and the AOT40 value was computed from 1000 conditional simulations using all other data. Figure 4 shows a fitted probability density of cross-validated AOT40-values at the monitoring stations. As expected, we see that the densities are skewed with heavy right tails. We also see that the observed AOT-values have fairly large estimated probability density, indicating that the model is reasonable for interpolating AOT40. The crude RMS error of the 10 cross-validated minus observed AOT40 values were 4586 $ppb\,h$. On the other hand, the empirical variance of the cross-validated AOT40 values were 8131 $ppb\,h$. This indicates that the true uncertainty intervals of the present interpolator may be smaller than the corresponding model-estimated intervals.

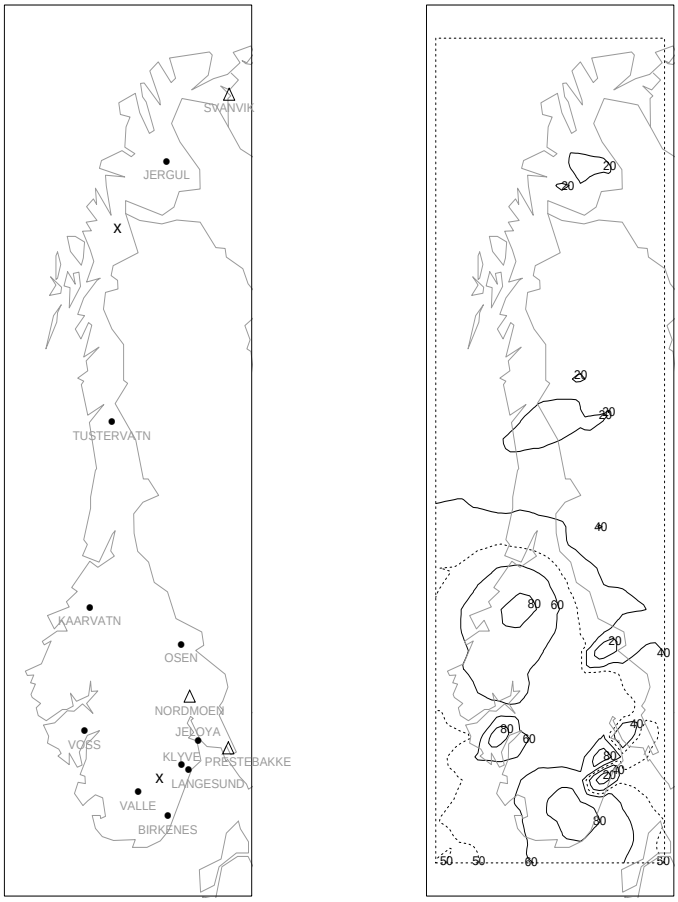

Figure 2. Left panel: Study area and location of monitoring stations. Filled circles are stations used in the analysis, triangles are validation stations. Right panel: Estimated risk $(100 \times \widehat{p})$ *of exceeding the critical level of 10000 ppb h AOT40.*

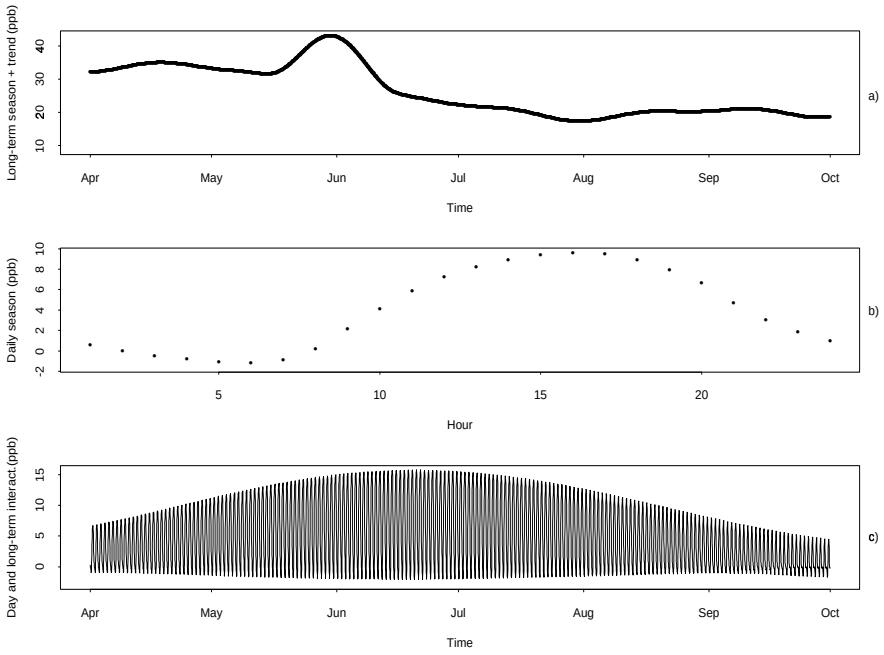

Figure 3. Main temporal structure components of ozone data.

5. Concluding Remarks

We have presented a general statistical model for fitting of geographically located time series of ozone concentrations with the purpose of spatially interpolating accumulated ozone concentration above a threshold. The method gives interpolated values and precision estimates for AOT40 exposure. Thus, we may give estimates of the quantities such as confidence intervals or the probability of exceeding critical levels. The method may also be used to interpolate hourly ozone concentration for any time and location within the domain of study. Furthermore, the method tolerates missing data in the observed data series and gives a consistent method for incorporating such data series in the analysis. Within the proposed framework, we may analyze in detail particular events of high concentrations, or generate time series of ozone for use in the study of effects on the environment.

The space-time separability has been used to obtain fast simulations of the residual process. Furthermore, the present exponential covariance function allows for efficient computations in the spatial domain due to screening effects (Cressie 1991). We hope that the limitations of these simplifying assumptions may be checked through future work.

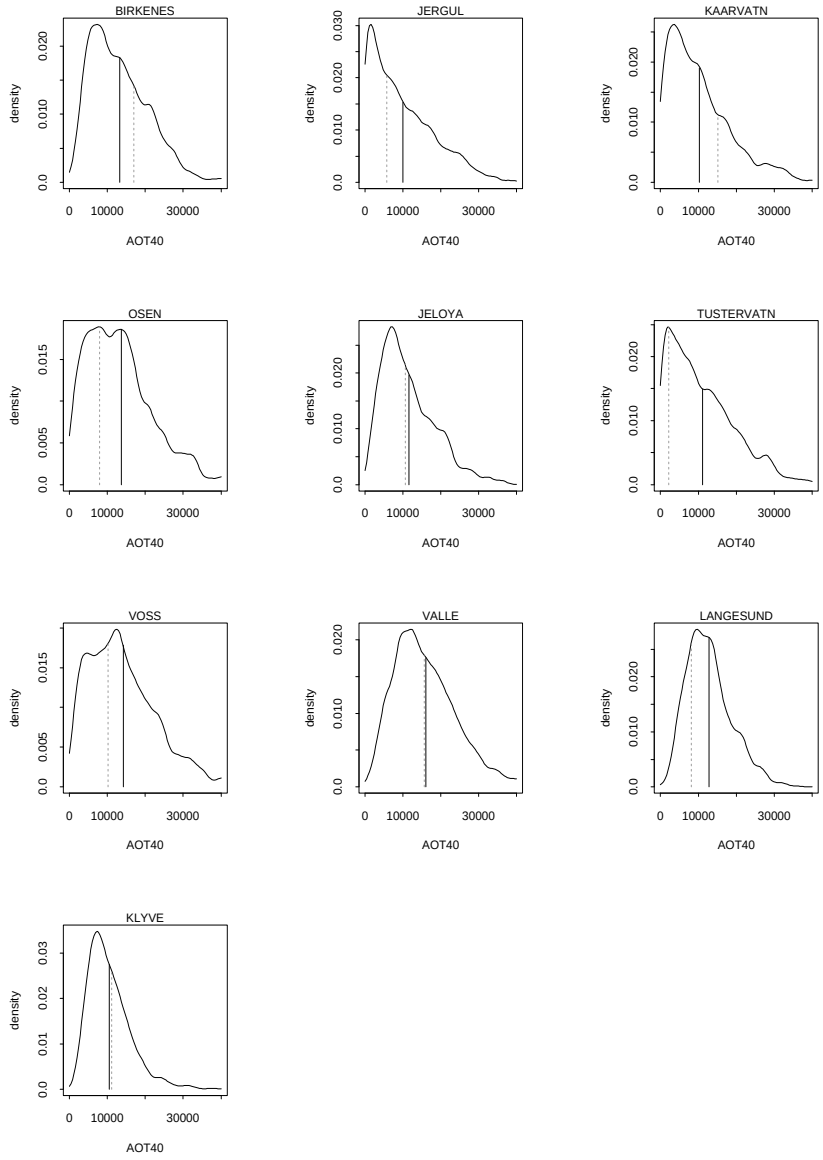

Figure 4. Density of cross-validated AOT40-values at the monitoring stations. Also shown is the observed AOT40-value (vertical full line) and the model estimate (broken vertical line).

Acknowledgements

This work was supported by the Norwegian State Pollution Control Authority (SFT). We thank Kjetil Tørseth and Sverre Solberg of NILU for critical comments and for providing the data. Magne Aldrin is thanked for helpful suggestions and for some software routines.

References

Aldrin, M.: 1996, Regression with few and noisy data, Dr.scient. thesis, Department of Mathematics, Statistics Division, University of Oslo.

Anderson, C. W. and Smith, R. I.: 1997, Environmental statistics and the mapping of ozone exposures, *Proceedings of the Workshop on Statistics at Universities*, Eötvös Press, Budapest, pp. 95–99.

Carroll, R. J., Chen, R., Li, T. H., Newton, H. J., Schmiediche, H. and Wang, N.: 1997, Ozone exposure and population density in Harris County, Texas (with discussion), *Journal of the American Statistical Association* **92**, 392–415.

Cressie, N.: 1991, *Statistics for Spatial Data*, Wiley, New York.

Davies, P. T. and Tso, M. K.-S.: 1982, Procedures for reduced-rank regression, *Applied Statistics* **31**, 244–255.

Diggle, P. J. and Hutchinson, M. F.: 1989, On spline smoothing with correlated errors, *Australian Journal of Statistics* **31**, 166–182.

Frank, I. E. and Friedman, J. H.: 1993, A statistical view of some chemometrics regression tools (with discussion), *Technometrics* **35**, 109–135.

Haas, T. C.: 1995, Local prediction of a spatio-temporal process with an application to wet sulphate deposition, *Journal of the American Statistical Association* **90**, 1189–1199.

Hart, J. D. and Wehrly, T. E.: 1986, Kernel regression estimation using repeated measurement data, *Journal of the American Statistical Association* **81**(396), 1080–1088.

Høst, G., Omre, H. and Switzer, P.: 1995, Spatial interpolation errors for monitoring data, *Journal of the American Statistical Association* **90**, 853–861.

UN-ECE: 1994, Critical levels for ozone in Europe, testing and finalizing the concepts, UN-ECE workshop report, Kuopio, Finland, 1996.

Omre, H., Sølna, K. and Tjelmeland, H.: 1992, Simulation of random functions on large lattices, *4th International Geostatistics Congress*, Troia, Portugal.

Rodriguez-Iturbe, I. and Mejia, J. M.: 1974, The design of rainfall networks in time and space, *Water Resources Research* **10**, 713–728.

Sølna, K. and Switzer, P.: 1996, Time trend estimation for a geographic region, *Journal of the American Statistical Association* **91**, 577–589.

Tørseth, K., Mortensen, L. and Hjellbrekke, A. G.: 1996, Mapping ground level ozone according to the critical level concept based on accumulated exposure over the threshold of 40 ppb (in Norwegian), NILU OR 12/96, Norwegian Institute of Air Research, P.O.Box 100, N-2007 Kjeller, Norway.

Wikle, C. K., Berliner, M. and Cressie, N.: 1997, Hierarchical Bayesian space-time models, Preprint Number 97-13, Iowa State University, Department of Statistics.