Spatiotemporal Modelling of Springwater Contaminants in the Presence of Complex Trends

Roussos Dimitrakopoulos[1] and Xiaochun Luo[2]

[1] W.H.Bryan Mining Geology Research Centre, The University of Queensland, Brisbane, Qld, Australia 4072

[2] Dept. of Mining & Metallurgical Engineering, McGill University, Montreal, Qc, Canada H3A 2A7

Abstract: Modelling and predicting groundwater quality from sparse data sets is increasingly being recognised as a critical and challenging issue. Data on groundwater quality vary simultaneously in time and space, reflecting complex chemical and physical/geological conditions. This paper deals with the spatiotemporal (s/t) stochastic estimation of water quality parameters when complex joint space-time trends are present. To account for the s/t nature of water contaminants, the estimation framework of universal kriging (UK) is used in a generalised s/t form. Several practical forms of s/t trends are examined including polynomials, Fourier functions and their combinations. Nitrate ion concentrations from natural springs are used to show the inference of s/t trends typical of water quality data and demonstrate the application of the s/t UK estimation process.

1. INTRODUCTION

The deterioration of groundwater quality is a critical worldwide issue whose recognition has prompted both interest and action to monitor water quality. Monitoring attempts to detect and predict contaminant levels, which may assist with the early detection of contaminant release from resource extraction and utilisation. Groundwater quality may be monitored from the analyses of water samples for dissolved ions. The results are subsequently used to model and predict contaminant levels and related variations. Groundwater chemical composition varies simultaneously in time and space due to complex interactions of aquifer characteristics, climatic fluctuations and human activities. The complexity of these interactions is expressed in data sets with distinct space-time trends that provide a challenge to modelling and prediction techniques.

Various space-time estimation techniques have been proposed, particularly in hydrology, atmospheric and environmental science. Earlier approaches include the use of time-averaged spatial covariances [Bilonick, 1983; Egbert & Lettenmaier, 1986], s/t separable covariance models [Rodriguez-Iturbe &Mejia, 1974; Rodriguez-Iturbe & Eagleson, 1987], or zonal anisotropy covariance models [Bilonick, 1985; Buxton & Pate, 1994]. Other approaches consider that a spatiotemporal process is a multivariate field constructed from spatial variables or temporal variables, space-time estimation is then based on co-kriging [Rouhani & Wackernagel, 1990; Goovaerts *et al.*, 1993; Goovaerts & Sonnet, 1993].

Modelling techniques accounting for complex space-time trends are somewhat limited in the literature. They include "moving space" models [Bras & Rodriguez-Iturbe, 1976], trend polynomial surface analysis [Berkowitz *et al.*, 1992], geostatistical models using generalised covariances with additive spatial and temporal components [Rouhani & Hall, 1989] and space-time separable polynomial generalised covariances [Christakos & Bogaert, 1996].

In this study, the step-by-step joint space-time modelling and prediction of springwater ion concentrations measured in springs over several years, is presented in the context of water quality monitoring. The modelling framework is first overviewed based on spatiotemporal random fields and the decomposition of a random field in a trend plus a random component [Dimitrakopoulos & Luo, 1997]. The trend models used are combinations of both polynomials and Fourier functions of various orders. The practical criteria for the fit of complex spatiotemporal trends are presented as part of the application. Next, the inference of spatiotemporal continuity measures [Dimitrakopoulos & Luo, 1994] for the contaminant residuals is demonstrated. Finally, the spatiotemporal inference of NO_3^- ion concentrations at unsampled locations in space-time is presented and the results discussed.

2. SPATIOTEMPORAL STOCHASTIC MODELS

2.1 Basic Definitions

The basic definitions of non-stationary spatiotemporal random fields are summarised in this section, as required for the present study. A space-time non-stationary random function, S/TRF, $Z(\mathbf{s},t)$, where $\mathbf{s} \in R^n$ and $t \in T$. $Z(\mathbf{s},t)$ is decomposed to $Z(\mathbf{s},t)=Y(\mathbf{s},t)+m(\mathbf{s},t)$, where $Y(\mathbf{s},t)$ is a stationary and ergodic random function with zero mean, covariance $C(\mathbf{h},\tau)$ and variogram $\gamma(\mathbf{h},\tau)$ with $\mathbf{h}=\mathbf{s}-\mathbf{s}'$, and $\tau = t-t'$. $m(\mathbf{s},t)$ is a joint space-time trend represented by

$$m(\mathbf{s},t) = \sum_{j=1}^{L} \alpha_j f_j(\mathbf{s}, t) = \mathbf{f}^T \alpha \qquad (1)$$

where $\mathbf{f}^T = [f_1(\mathbf{s},t), \ldots, f_L(\mathbf{s},t)]$ is a vector of known functions, and $\alpha = [\alpha_1,\ldots,\alpha_L]$ are unknown coefficients.

2.2 Spatiotemporal Estimation

Spatiotemporal estimation based on the above definitions, is formulated as follows. For a set of N data $\{z(\mathbf{s}_i,t_i), i=1, \ldots, N\}$, a kriging estimator at unknown location $(\mathbf{s}_0,t_0)$ is

$$z^*(\mathbf{s}_0,t_0) = \sum_{i=1}^{N} \lambda_i z(\mathbf{s}_i,t_i) \qquad (2)$$

where the weights λ_i are derived from the traditional universal kriging system (UK) system extended into space-time,

$$\begin{aligned} \mathbf{C}\lambda - \mathbf{F}\mu &= \mathbf{C}_0 \\ \mathbf{F}^T \lambda &= \mathbf{f}_0 \end{aligned} \qquad (3)$$

with vectors $\lambda^T=[\lambda_1, \ldots,\lambda_N]$, $\mu^T=[\mu_1, \ldots,\mu_L]$, $\mathbf{C}_0^T=[C(\mathbf{h}_{01},\tau_{01}), \ldots,C(\mathbf{h}_{kN}, \tau_{kN})]$, $\mathbf{f}_0^T = [f_1(\mathbf{s}_0,t_0), \ldots,$

$f_L(s_0,t_0)]$, matrix $\mathbf{C}=[C_{ij}]=[C(\mathbf{h}_{ij},\tau_{ij})]$, and the N×L matrix $\mathbf{F}=[F_{ij}]=[f_j(s_i,t_i)]$. The minimised UK estimation variance is

$$\sigma^2_{UK}(s_0, t_0) = -\mathbf{C}_0^T \lambda + \mathbf{f}_0^T \mu + C(\mathbf{h}_{00},\tau_{00}) \qquad (4)$$

2.3 Forms of Complex Spatiotemporal Trends

The s/t UK system in (3) requires that the s/t trend models meet several requirements discussed explicitly by Dimitrakopoulos & Luo [1997]. The same authors suggest three general types of s/t trend forms, polynomial, Fourier and mixed, summarised below for the common R^2 and T case.

The polynomial form of a s/t trend, following the notation in (1) is

$$f^T = [1, s_x, s_y, t, \ldots, s_x^{\xi}, s_y^{\xi}, t^{\zeta}] \qquad (5)$$

where ξ and ζ are orders in space and time, respectively.

The general Fourier form is

$$f^T = [1, \sin\omega_x s_x \cdot \sin\omega_y s_y \cdot \sin\omega_t t, \sin\omega_x s_x \cdot \sin\omega_y s_y \cdot \cos\omega_t t, \ldots, \cos\omega_x i s_x \cdot \cos\omega_y i s_x \cdot \sin\omega_t i t, \cos\omega_x i s_x \cdot \cos\omega_y i s_y \cdot \cos\omega_t i t] \qquad (6)$$

where ω_x, ω_y and ω_t are frequencies in directions s_x, s_y, and t, respectively; i denotes the order of the Fourier series.

The mixed form includes combinations of the above. The mixed forms can be generated from m(s,t)=m(s)m(t), where m(s) is the spatial trend and m(t) a temporal trend. An example is

$$f^T = [1, s_x, s_y, \sin\omega_t t, \cos\omega_t t, s_x\sin\omega_t t, s_x\cos\omega_t t, s_y\sin\omega_t t, s_y\cos\omega_t t, \ldots] \qquad (7)$$

2.4 Practical Spatiotemporal Trend Models

The spatiotemporal trend forms given in the previous section can adequately model complex joint s/t trends in most practical cases, using low orders.

When polynomial trends are considered, the space and time orders ξ and ζ up to 2, generate four combinations, sufficient to cover most practical cases. A s/t Fourier trend form of order i set to 1 generates eight terms which are also quite sufficient. A larger order either in time or space may generate an unnecessarily large number of terms without any practical significance. The orders that seem appropriate to the data set to be modelled are selected in practice.

Mixed forms are the most flexible. For instance, a mixed form consisting of polynomials of order up to 2 and Fourier series of order 1 generate several useful combinations:

a) for $\xi=1$, $i=1$:
$$f^T = (1, s_x, s_y, \sin t, \cos t, s_x\sin t, s_x x\cos t, s_y\sin t, s_y\cos t)$$

b) for $\xi=2$, $i=1$:
$$f^T = (1, s_x, s_y, s_x^2, s_x s_y, s_y^2, \sin\omega_t t, \cos\omega_t t, s_x\sin\omega_t t, s_x\cos\omega_t t, s_y\sin\omega_t t, s_y\cos\omega_t t, s_x^2\sin\omega_t t, s_x^2\cos\omega_t t, s_x s_y\sin\omega_t t, s_x s_y\cos\omega_t t, s_x s_y^2\sin\omega_t t, s_x s_y^2\cos\omega_t t)$$

c) for $i=1$, $\zeta=1$:
$$f^T = (1, \sin\omega_x s_x, \cos\omega_x s_x, \sin\omega_y s_y, \cos\omega_y s_y, t, t\sin\omega_x s_x, t\cos\omega_x s_x, t\sin\omega_y s_y, t\cos\omega_y s_y)$$

d) for $i=1$, $\zeta=1$:
$$f^T = (1, \sin\omega_x s_x, \cos\omega_x s_x, \sin\omega_y s_y, \cos\omega_y s_y, t, t^2, t\sin\omega_x s_x, t\cos\omega_x s_x, t\sin\omega_y s_y, t\cos\omega_y s_y, t^2\sin\omega_x s_y, t^2\cos\omega_x s_x, t^2\sin\omega_y s_y, t^2\cos\omega_y s_y)$$

Given their possible forms, space-time trend models can be fitted using trend surface analysis (Ripley, 1981; Agterberg, 1974) extended in space-time. This is the approach followed in the present study.

5. SPATIOTEMPORAL ANALYSIS OF NITRATE ION CONCENTRATIONS IN SPRING WATER

5.1 Study Area and Data Statistics

High nitrate ion concentrations usually reflect human activities, which include the use of fertilisers and waste from sewage systems. Variation in intensity of human activities combined with climatic and geohydrological result in space-time variations of nitrate concentrations in groundwater which may be sampled at natural spring locations. The applied aspects of modelling and prediction of nitrate levels in unsampled locations in space and instances in time is shown in this section, based on the developments outlined in section 2.

In the present study, the springwater nitrate (NO_3^-) ion concentrations, in milligrams per litre (mg/l) from 68 springs are analysed (Figure 1). The data accuracy is estimated at 5 mg/l. In total, 408 data were collected over six years (1975 to 1983). The statistics of the data are shown in Figure 2.

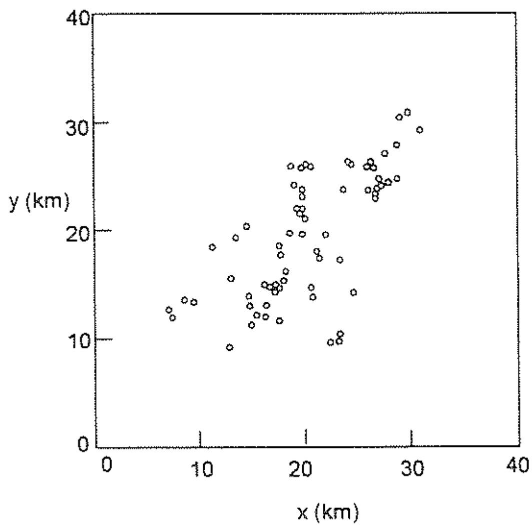


Figure: Spring locations

The trend analysis of the data set is presented in the next section.

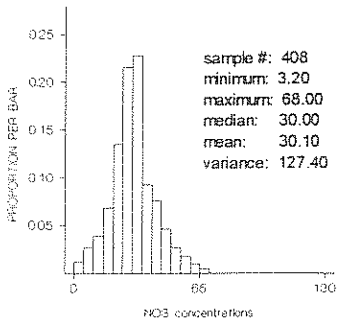


Figure 2: Data statistics

5.2 Trend Inference

Trend surface analysis is used in this section in space-time to identify the presence of trends and fit appropriate models to the NO_3^- ion concentration data set. To accommodate a local trend fitting, the study area is used to infer the appropriate form and order of the trend. The trend coefficients are inferred locally, to reflect local s/t conditions. Four spatially defined sub- areas are used to infer local trend coefficients. These correspond to the four quarters evenly dividing the study area. The subdivision was based on visual inspection of the data and is further verified by the presence of trends, well fitting trend models and the generation of acceptable residuals, as required by the decomposition in section 2.1.

A total of 16 trend models based on (5), (6), (7) are included in the trend surface analysis. The trend models used here are specified by the order of spatial component denoted with ξ, the order ζ of the temporal component and the order i of the Fourier terms. The notion $\xi=\xi+i$ indicates a mixed polynomial-Fourier space component with polynomial order ξ and Fourier order i. The period of the Fourier terms is derived from the continuity measures (omnidirectional experimental s/t variograms) and tested.

Table 1: Trend models and results of trend surface analysis. The spatial and temporal periods are 2.0 and 3.0 respectively.

Trend model (s/t)	Mean absolute error	Mean square error	Mean error	Goodness of fit
$\xi = 0/\ \zeta = 1$	8.436	127.39	.00000035	.0231
$\xi = 0/\ \zeta = 2$	8.293	124.27	-.00000559	.0516
$\xi = 1/\ \zeta = 0$	7.901	107.73	-.00000006	.1667
$\xi = 1/\ \zeta = 1$	7.789	105.01	.00000096	.1899
$\xi = 1/\ \zeta = 2$	7.660	101.89	.00000593	.2184
$\xi = 2/\ \zeta = 0$	7.251	83.306	.00002093	.3304
$\xi = 2/\ \zeta = 1$	7.122	80.579	-.00000799	.3535
$\xi = 2/\ \zeta = 2$	6.936	76.875	-.00002518	.3871
$\xi = 0/\ \zeta = i$	8.357	125.14	-.00000081	.0439
$\xi = 1/\ \zeta = i$	7.551	97.714	.00000145	.2531
$\xi = 2/\ \zeta = i$	8.249	110.70	-.00000019	.1317
$\xi = i/\ \zeta = 0$	5.542	48.530	-.00000003	.5742
$\xi = i/\ \zeta = 1$	8.470	128.25	.00000005	.0128
$\xi = i/\ \zeta = i$	7.504	94.344	.00000005	.2578
$\xi = 1+i/\ \zeta = 0$	7.504	94.344	.00000005	.2578
$\xi = 1+i/\ \zeta = 1$	**4.978**	**41.534**	**-.00001967**	**.6265**
$\xi = 1+i/\ \zeta = i$	7.824	105.72	.00000032	.1809

Table 1 summarises the 16 trend models used in the study and the results of trend surface analysis. The table includes the four goodness of fit criteria used to select the best s/t trend model. The periods for the Fourier terms are 2 in space and 3 in time and were retained due to the better fit of the trend models. Note that sensitivity to the periods of space and time trends was carried out. The best fitting trend model, bolded in Table 1, is linear plus harmonic in space ($\xi=1+i$) and linear in time ($\zeta=1$). Table 2 shows the trend coefficients for each of the sub areas used.

Table 2: Coefficients of trend model fitted to the nitrate ion concentration data. The spatial period is 2.0.

s/t terms	Trend Coefficients			
	Subarea 1	Subarea 2	Subarea 3	Subarea 4
1	36.16709	1388.283	-77.37017	91.21171
sinxsiny	14.66309	6.93164	5.94253	1.73782
sinxcosy	-11.24191	11.39687	-13.22187	-2.95470
cosxsiny	28.48949	2.38497	-9.09622	1.61511
cosxcosy	3.58883	1.56553	14.53264	-2.84192
x	3.17122	-9.25165	4.85633	-11.85963
y	-5.72212	-11.34528	1.84959	4.91687
t	-.37057	-16.95049	-2.19038	-2.30121
tsinxsiny	-.27221	-.08980	-.10192	.06021
tsinxcosy	-.00568	.16165	.12205	.01690
tcosxsiny	-.06222	-.12204	-.10081	.00007
tcosxcosy	.45124	.16563	-.25864	.08260
xt	-.02094	.74913	.19562	.15677
yt	.07927	.13082	-.07948	-.06043

5.3 Spatiotemporal Continuity of Residual Nitrate Ion Concentration

Following the UK decomposition in section 2.1, residuals of the data are generated based on the inferred s/t trends. In addition to the fit criteria of the trend models in the previous section, two requirements for accepting trend model are considered here. First, the residuals are normally distributed and, secondly, the spatiotemporal variograms or covariances of the residuals are isotropic. Both these requirements are met here, as shown in a later paragraph.

Figure 3 shows the histogram and statistics of the nitrate ion concentration residuals. The experimental directional s/t variograms are shown in Figure 4a to 4b and are clearly isotropic with a space range at about 900m and time range at possibly 6 years. Note that the calculation of the spatiotemporal experimental variograms is based on separate 2D space and time lags.

The model fitted to quantify the joint space-time continuity of the nitrate residuals is isotropic, exponential and given by:

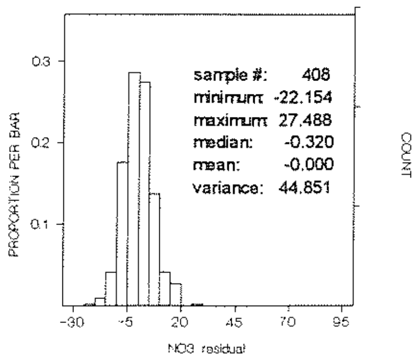
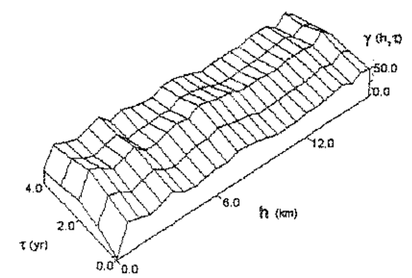


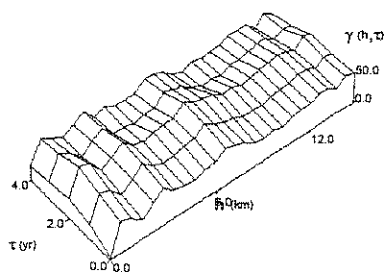
Figure 3: Statistics of nitrate residuals.

$$\gamma(h, \tau) = 2.98 + 41.5 \cdot \{1 - \exp[-((\frac{h}{0.3})^2 + (\frac{\tau}{2.0})^2)^{1/2}]\}$$

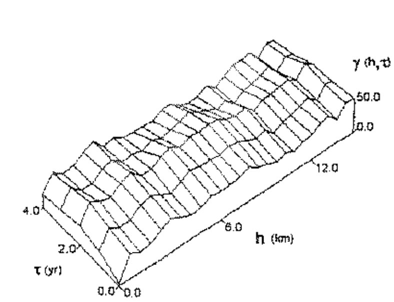
Note that issues relevant to definitions and permissibility criteria for variograms and covariance functions of stationary and ergodic spatiotemporal random fields are presented elsewhere [Dimitrakopoulos and Luo, 1994].



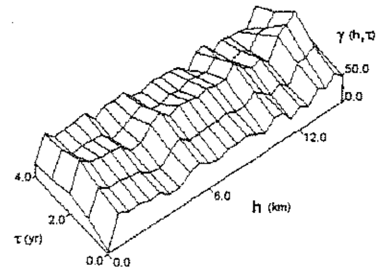
(a)



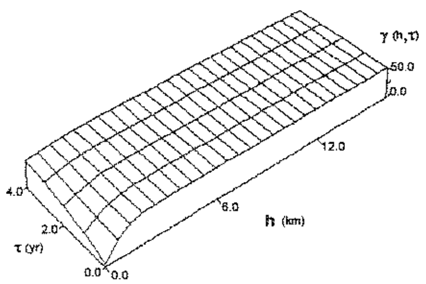
(b)



(c)



(d)



(e)

Figure 4: Spatiotemporal experimental variograms of nitrate ion concentration residuals. The spatial directions are (a) E-W, (b) NE, (c) N-S, (d) NW; (e) fitted variogram model.

6. ESTIMATION OF SPRINGWATER CONTENTS

Following the trend analysis and model fitting, the space-time UK system in (3) is used to interpolate at unsampled locations in space and instances in time. The estimation uses a grid size of 300 meters by 4 months. It is conjectured that the use of local trend coefficients will enhance the local prediction in the produced maps.

Figure 5 shows the spatiotemporal mapping of the nitrate in concentrations for 1979, 1980 and 1981. High NO_3^- concentrations are observed at the centre, southern central part and to a lesser extent the eastern part. The high-value zone in the central area expand with time while there are no distinct changes for other high-value zones. Further linking of the produce maps with information on human activated and hydrologic conditions in the study area could provide valuable interpretations.

7. CONCLUSIONS

A spatiotemporal modelling of water quality data for predictive and monitoring reasons is a critical and technically challenging issue. Since water quality data have a distinct spatiotemporal nature, it is necessary to develop and use appropriate modelling techniques. This study presents a technique which captures and quantifies joint spatiotemporal characteristics, including local trends and spatiotemporal continuity.

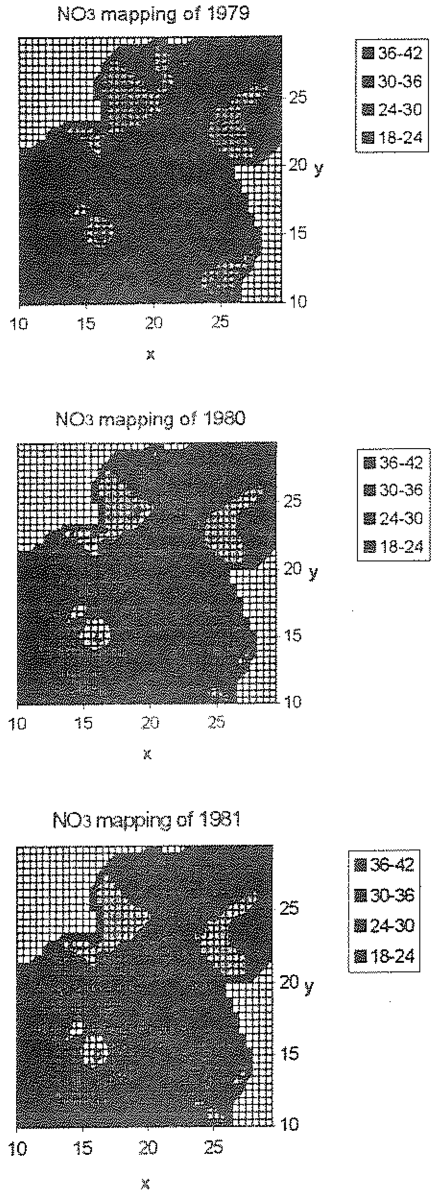


Figure 5 Spatiotemporal predictions of NO_3^- ion concentrations (in mg/l) for 1979, 1980 and 1981.

This study demonstrates that spatiotemporal universal kriging is a suitable stochastic modelling framework for modelling spatiotemporal groundwater data such as nitrate ion concentrations. Major advantage of the UK framework is its flexibility in describing spatiotemporal trends by capturing global characteristics as well as local details. In addition, the use of complex joint space-time trend models combining polynomial with harmonic terms adds the flexibility needed to adequately model data periodicities.

8. ACKNOWLEDGMENTS

Acknowledgments are in order to Pierre Goovaerts for providing the data set used. Funding was provided from the Natural Sciences and Engineering Research Council of Canada, Research Grant OGP0105803 to RD.

9. REFERENCES

Agterberg, F., *Geomathematics*, Elsevier, Amsterdam, The Netherlands, 1974.

Berkowitz, B., Ben-Zvi and Berkowitz, J., A spatial, time-dependent approach to estimation of hydrologic data, *J. Hydrology*, 135, 133-142, 1992.

Bilonick, R.A., Risk-qualified maps of hydrogen ion concentration for the New York state area for 1966-1978, *Atmospheric Environment*, 17, 2513-2524, 1983.

Bilonick, R.A., The space-time distribution of sulfate deposition in the northeastern U. S., *Atmospheric Environment*, 19, 1829-1845, 1985.

Bras, R.L. and Rodriguez-Iturbe, I., Rainfall generation: A non-stationary time-varying multi-dimensional problem, *Water Resources Research*, 12, 450-456, 1976.

Buxton, B.E. and Pate, A.D., Joint spatial-temporal modeling of concentrations of hazardous pollutants in urban air, in, R. Dimitrakopoulos (ed.), *Geostatistics for the Next Century*, Kluwer Academic Publishers, Dordrecht, 75-87, 1994.

Christakos, G. and Bogaert, P., Spatiotemporal analysis of spring water ion process derived from measurements at the Dybe basin in Belgium, *IEEE Trans. Geoscience and Remote Sensing*, 34, 1-17, 1996.

Dimitrakopoulos, R. and Luo, X., Spatiotemporal modelling: Covariances and ordinary kriging systems, in *Geostatistics for the Next Century*, R. Dimitrakopoulos (ed.), Kluwer Academic Publishers, Dordrecht, 88-93, 1994.

Dimitrakopoulos, R. and Luo, X., Joint space-time modelling in the presence of trends, in *Geostatistics-Woolongong'97*, E. Baafi and N. Scofield (eds.), Kluwer Academic Publ., Dordrecht, 138-149, 1997.

Egbert G. and Lettenmaier, D.P., Stochastic modeling of the space-time structure of atmospheric chemical deposition, *Water Resources Research*, 22, 165-179, 1986.

Goovaerts, P. and Sonnet, P., Study of spatial and temporal variations of hydrogeochemical variables using factorial kriging analysis, in *Geostatistics Troia '92*, A. Soares (ed.), Kluwer Academic Publishers, Dordrecht, 745-756, 1993.

Goovaerts, P., Sonnet, P. and Navarre, A., Factorial kriging analysis of spring water components in the Dyle River basin, *Water Resources Research*, 29, 2115-2125, 1993.

Ripley, B.D., *Spatial Statistics*, John Wiley & Sons, Inc., New York, NY, 1981.

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

Rodriguez-Iturbe, I. and Eagleson, P., Mathematical models of rainstorm events in space and time, *Water Resources Research*, 23, 181- 190, 1987.

Rouhani, S. and Hall, T. J., Space-time kriging of groundwater data, in *Geostatistics*, M. Armstrong (ed.), Kluwer Academic Publishers, 639-651, 1989.

Rouhani, S. and Myers, D., Problems in space-time kriging of geohydrological data. *Mathematical Geology*, 22, 611-623, 1990.

Rouhani, S. and Wackernagel, H., Multivariate geostatistical approach to space-time data analysis, *Water Resources Research*, 26, 585-591, 1990.