Mapping the annual exceedance frequencies of the PM10 air quality standard

Comparing kriging to a generalized linear spatial model

Khuluse S.* *(skhuluse@csir.co.za)* and Stein A.

Abstract

An objective in mapping air quality attributes such as concentrations of airborne particles (particulate matter – PM) is to determine those areas which can be considered as hotspots and determine factors that contribute to their formation. Classical kriging has been applied extensively in mapping air quality variables, with most applications focusing on average pollutant concentrations. Generalized linear spatial process models are being applied as an alternative to classical kriging.

In this paper we compare ordinary and regression kriging models to the Poisson log-linear spatial model (Diggle et al. 1998, Diggle et al. 2007) with and without covariate information in mapping annual average exceedance frequencies of the South African PM10 air quality standard of 120 $\mu g/m^3$ (RSA Govt. Gazette 2009, 2012). We use daily PM10 data from 36 air quality monitoring sites in the Highveld (Gauteng and western Mpumalanga provinces) for the 48 months period from September 2009 to August 2012. Higher concentrations are observed in high density residential areas, with high proportion of informal and mixed types of dwellings. Therefore, significance of household energy use, number of households and settlement type as explanatory variables in mapping the yearly exceedance rates are explored.

Materials and Methods

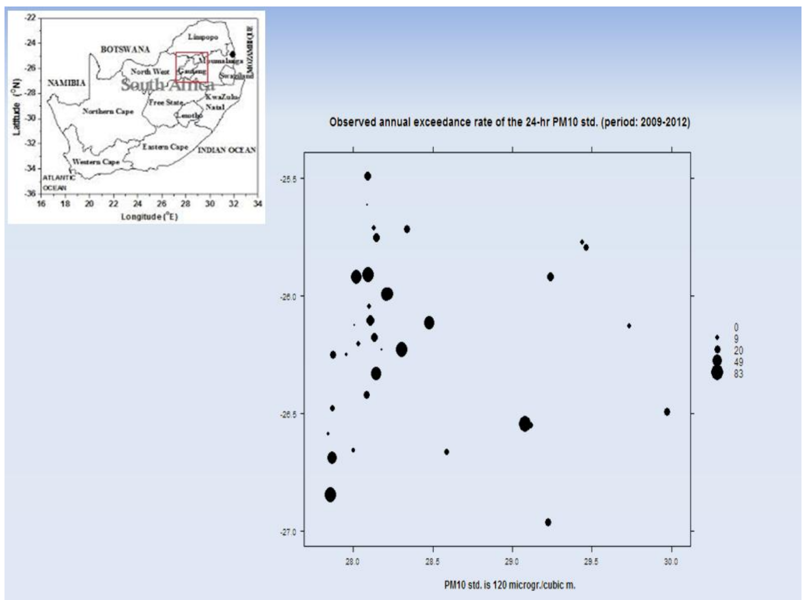

Figure 1 Study area and obversed average number of exceedances per year

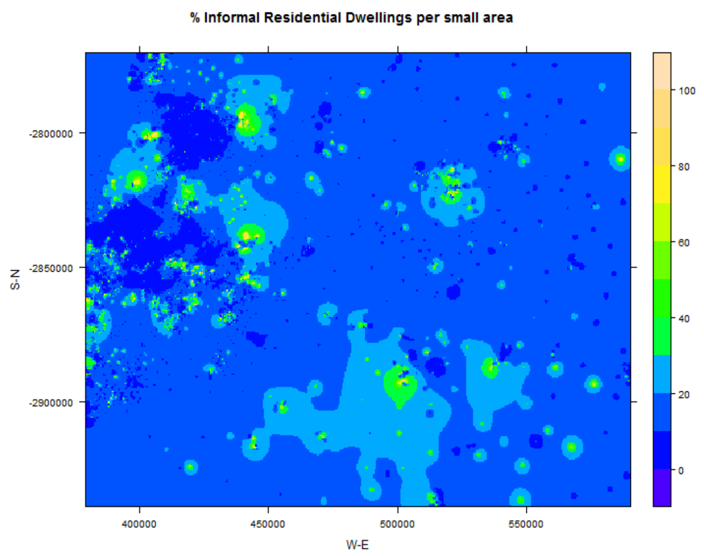

Figure 2 Map of percentage of dwellings in a small area classified as informal obtained using inverse distance weighting.

Consider n spatial observations $y(s_i)$, in this case these are average number of exceedances of the national 24-hour regulatory level of PM10 per year obtained from daily data observed during the period September 2008 to August 2012. These can be considered as realizations of a random field $\mathbf{Y(s)}$, a spatially correlated Poisson process. Spatial correlation is invoked thorough the Poisson intensity $\boldsymbol{\lambda}\mathbf{(s)}$ which is considered a function of an unobserved spatial process that accounts for large scale spatial trend through covariates $\boldsymbol{\mu}\mathbf{(s)}=\mathbf{X(s)}^T\boldsymbol{\beta}$ and local spatial variation through a latent zero-mean Gaussian process $\mathbf{W(s)}$ with covariance matrix $\mathbf{V} = \sigma_s^2\mathbf{R}$. Elements of matrix $\mathbf{R}$ describe correlation between observations that are distance $h=|s_j-s_i|$ apart. In this case an exponential model is considered for the correlation between pairs of observations. The Poisson log-linear spatial model (abbreviated as GLSM) is set up as follows:

$$\mathbf{Y(s)}|\mathbf{W(s)} \sim \text{Poisson}(\boldsymbol{\lambda}\mathbf{(s)})$$

$$\text{Log}(\boldsymbol{\lambda}\mathbf{(s)}) = \boldsymbol{\mu}\mathbf{(s)} + \mathbf{W(s)}$$

Results

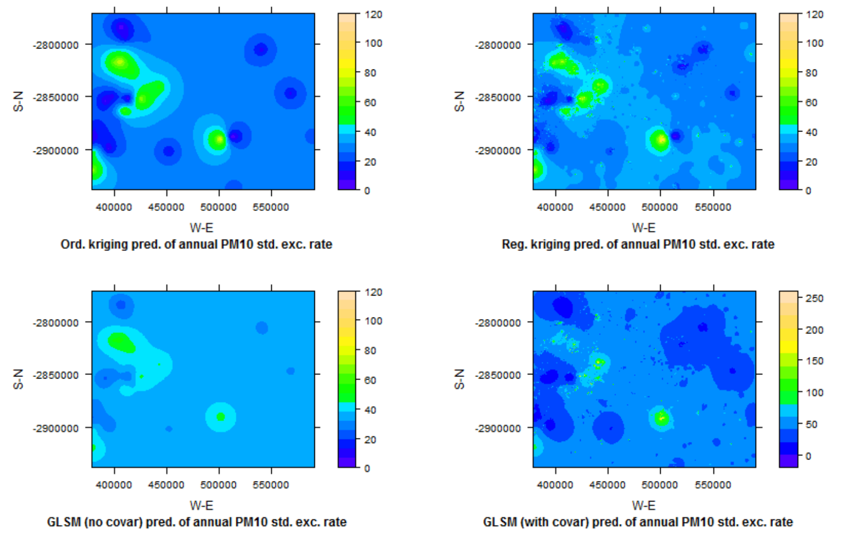

Figure 3 Maps of the number of exceedances of the PM10 standard per year. The covariate in regression kriging and the Poisson log-linear spatial model (GLSM2) is the spatial layear giving the % percentage of informal dwelling type in each small area

Table 1 Prediction results - 4 of the 36 stations were used as validation sites.

		Ord Krig	**Reg Krig (Covariate: % Informal dwell/SAL)**	**Poisson log-linear spatial model**	**Poisson log-linear spatial model (Covariate: % Informal dwell/SAL)**	
Validation sites	Obs	σ2=712.7, φ=8.5 km, WLS=0.014	β1=0.3, σ2=553.7, φ=6.7 km, WLS=0.0116	β0 = 3, σ2=1.3, φ=8 km, -log L = -137	β0=2.9, β1=0.01, σ2=1.3, φ=8 km, -log L= -135	Sites close to obs sites
Delta (Urban Bground)	0.5	26.0	23.8	33.9	29.7	Alex (20)
Buccleuch (Traffic)	8.1	43.5	32.8	57.8	42.1	Alex (20)
Tembisa (Residential)	51.3	52.7	42.3	59.9	41.8	Ivory (54)
Langerverwacht (Industrial)	11.1	57.6	80.5	74.3	207.4	Club (0); Secunda (82)
ME {E(ME)=0}		*27.2*	*27.1*	*38.8*	*62.5*	
RMSPE {E(RMSPE)=Var(Y)}		*32.2*	*28.8*	*42.6*	*86.4*	
RMNSPE {E(RMNSPE)=1}		*1.8*	*2.0*	*0.8*	*1.9*	

Concluding Remarks

The spatial pattern of exceedance rate – in terms of location of hot-spots - is consistent for both kriging and the Poisson log-linear model.

The advantage of having additional explanatory spatial variables is diminished if it's correlation with the response surface is weak. The precision of both regression kriging and the Poisson log-linear model in predicting exceedannce rate in unsampled areas including the validation sites is very low.

Number of households (dwelling total per SAL) and houseld use of biofuels (wood, coal and dung) for cooking and heating were not significant as explanatory variables; percentage of dwellings classified as informal per small area from Census 2011 was the only significant explanatory variable.

References

1. Bivand R. S., Pebesma E. J., Gomez-Rubio V. (2008). *Applied Spatial Data Analysis with R.* Springer.
2. Diggle, P. J., Tawn, J. A. and Moyeed, R. A. (1998). *Model-based geostatistics (with Discussion).* Applied Statistics, 47, 299–350.
3. Diggle, P. J. and Ribeiro Jr., P. J. (2007). *Model-based geostatistics.* Springer series in statistics.
4. Monestiez P., Dubroca L., Bonnin E., Durbec J.-P., Guinet C. (2004). *Comparison of model based geostatistical methods in ecology: Application to fin whale distribution in northwestern Mediterranean sea.* In proceedings of Geostatistics Banff, Leuangthong O. and Deutch C. V. (eds.), Springer.
5. RSA Government Gazette (2009). *National ambient air quality standards*. National environmental management: Air quality act 2004 (Act No. 39 of 2004), Vol. 534, No. 32816.
6. Van de Kassteele J. (2006). Statistical air quality mapping. PhD Dissertation. (Promoters Stein A. and Dekkers A. L. M.)

Acknowledgements: South African Weather Service, Stats SA for the data. Pierre Du Plessis, Nontembeko Dudeni-Tlhone, Peter Schmitz and Johan Maritz for their technical inputs. CSIR and Nuffic for funding.

ITC CSIR our future through science