

8.6.3 The wet sulfate deposition data

Exploratory data analysis scattered throughout Chapter 3 suggested that it would be wise to transform the wet sulfate deposition data to the square-root scale and to set aside three outliers, yielding a modified dataset that was called the clean wet sulfate deposition data. Polynomials of orders zero through five were fitted to these modified data by OLS in Section 4.5.3, and summary statistics from these fits suggested that polynomials up to third-order (and possibly higher) were worthy of further consideration. The analyses presented in Section 4.5.3 also suggested that spatial autocorrelation exists and is isotropic in the residuals from the quadratic fit, but perhaps does not exist in the residuals from the cubic and higher-order fits.

Here, we fit polynomial models to the clean wet sulfate deposition data once again, but this time using likelihood-based methods. Models having polynomial mean structures up to (and including) order five and having several geostatistical covariance functions were estimated by ML and REML. Custom R code was written showing how to using the `optim` function to minimize two times the negative loglikelihood to obtain ML and REML estimates. The function `optim` contains multiple minimization algorithms, and we used Nelder-Mead, as mentioned earlier. This also returns the minimized loglikelihood value, which can be used to compute AIC and other diagnostics. For example, Figure 1 shows AIC computed for all polynomial models on the two spatial coordinates from the constant mean model up to a fifth-order polynomial for both independent error models and spatially-autocorrelated error model, where the spatial autocorrelation is an exponential model. Note that AIC generally decreases with the higher polynomial orders for the independent error models, while for the spatial models AIC decreases slightly to the 4th order polynomial, and starts to increase again with the fifth order polynomial. We should point out here that AIC should be based on a likelihood minimized using ML rather than REML estimation if the column space of the fixed-effects design matrix changes among models. REML estimation can be used in AIC to compare spatial models for a fixed-effects design matrix that does not change among models, for example, when comparing spatial autocorrelation models for an unchanging set of covariates. AIC has a reputation for favoring overly complex models, so we also use leave-one-out cross-validation (LOOCV) and n-old cross-validation.

LOOCV eliminates one datum at a time, using all of the rest of the data to predict the one that was removed. There is a fast and a slow way to do this. The slow way is to remove a datum and then re-estimate all of the parameters using ML or REML estimation each time. However, with the removal of but a single datum, the parameter estimates change very little. A fast way to achieve LOOCV is based on holding all parameters at their values as estimated by using all of the data, and then using results from partitioned matrices so that we only have to invert the covariance matrix once (which can be saved from the ML or REML estimation, so there is in fact no additional inverses required. Note that if a matrix is partitioned as,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix},$$

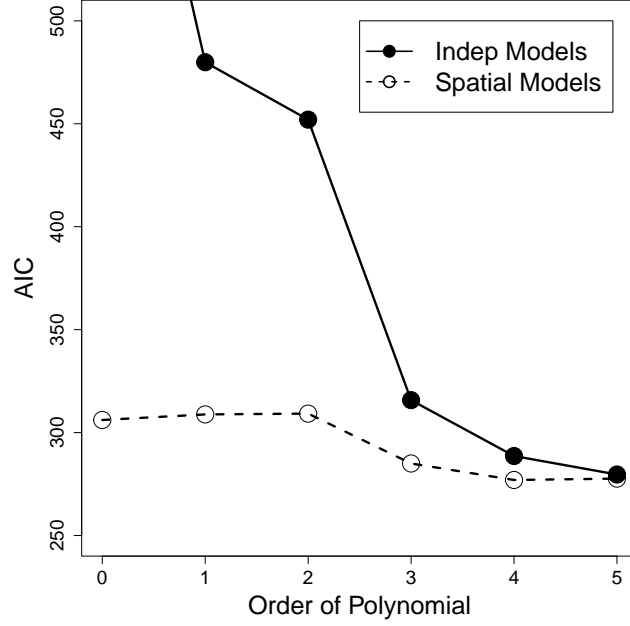


Figure 1: AIC for polynomial models on the spatial coordinates, for both independent error models (solid line and circle) and spatial models (dashed line and open circles).

then

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}\mathbf{S}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}\mathbf{S}^{-1} \\ -\mathbf{S}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & \mathbf{S}^{-1} \end{bmatrix},$$

where $\mathbf{S} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ is known as the Schur complement of block Σ_{22} of matrix Σ . Now write Σ^{-1} in block form as

$$\Sigma^{-1} = \begin{bmatrix} \tilde{\Sigma}_{11} & \tilde{\Sigma}_{12} \\ \tilde{\Sigma}_{21} & \tilde{\Sigma}_{22} \end{bmatrix},$$

where the dimensions of the blocks in Σ^{-1} follow those from the dimensions of Σ . Note that the inverse of Σ_{11}^{-1} can be obtained from Σ^{-1} by

$$\Sigma_{11}^{-1} = \tilde{\Sigma}_{11} - \tilde{\Sigma}_{12}\tilde{\Sigma}_{22}^{-1}\tilde{\Sigma}_{21}.$$

Moreover, let us order the data such that the datum to be removed is last, so that \mathbf{S} is a scalar, then the inverse of \mathbf{S} is trivial and Σ_{11}^{-1} can be computed rapidly. The main computational expense of kriging predictions rely on the inverse covariance matrix for the observed data, but in LOOCV that is given by Σ_{11}^{-1} , which is computed rapidly if we already have Σ^{-1} .

LOOCV results for the clean wet sulfate deposition data are presented in Figure 2.

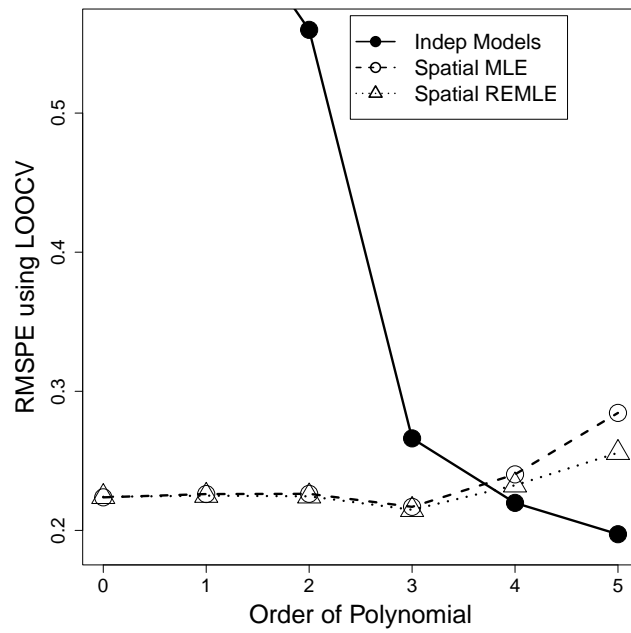


Figure 2: LOOCV for polynomial models on the spatial coordinates, for independent error models (solid line and circle) and spatial models with ML estimation (dashed line and open circles) and REML estimation (dotted line with open triangles).