Technical Vignette 2: Smoothing characteristics of CAR models

Christopher Paciorek, Department of Biostatistics, Harvard School of Public Health

Version 1.0: January 2008

This vignette describes some troublesome features of smoothing with conditional auto-regressive models using standard 0-1 weights. One very nice solution is to use weights based on thin-plate spline smoothing as I discuss below.

Conditional auto-regressive models (a type of Markov random field specification) do spatial smoothing (and smoothing in other spaces) using a precision matrix that is often based on the neighborhood structure of the areal units under consideration. A standard neighborhood structure is to let all the neighbors of an areal unit have weight one and non-neighbors have weight zero. For gridded units, one might let all units whose centroids are within a certain distance have weight one and others have weight zero. This reduces to the nearest 8 and nearest 4 (i.e, cardinal direction neighbors) cells if one specifies the distance appropriately.

CAR models are generally used for data in areal units in which the areal units represent a partition of the domain and there are data in all the units. However, one might also consider using CAR models for gridded representations of space in which many of the grid cells have no observations. Rue and Held (2005) have convincingly argued for the computational advantages of CAR models through the use of sparse matrices and joint MCMC proposals for the full process (and also the full process and the process hyperparameter). Therefore, let's consider a latent process on a grid, g, with an CAR-derived precision matrix, κQ where κ is the precision parameter and Q contains the negative of the weights for all pairs of points and the diagonal of Q the sum of the weights for each location. I fit a Bayesian model with a simple likelihood based on the latent process evaluated at the appropriate cell, $y_i \sim \mathcal{N}(\mu + g_i, \sigma_y^2)$. Fig. 1 shows results for an example dataset. The middle row shows the smoothing for simple 0-1 weights for two neighborhoods. Note the strange surface prediction, with bulls-eyes (cusps) around the observations and, for points away from the observations, smoothing towards the overall mean. In contrast, Rue and Held (2005, p. 114), Thomas Kneib's PhD thesis (p.60), and Yue and Speckman (in preparation, U. Missouri Statistics) describe weights motivated by the thin-plate spline penalty, with weights of 8 for the four cardinal neighbors, -2 for the four diagonal neighbors and -1 for the 4 cardinal neighbors at distance of two units. This structure gives the smoothing seen in the lower left plot, which is very similar to a thin-plate spline smooth (lower right plot), and displays more attractive smoothing characteristics. In this case, the best fit to the data is a very smooth surface.

In Fig. 2, I show the smoothing kernels induced by the different neighborhood specifications, which help to explain the smoothing we see in Fig. 1. I calculate the smoothing kernel as follows. First, recall that the kriging smoothing matrix is $S = C(C + \sigma_y^2 I)^{-1} = (\kappa Q)^{-1}((\kappa Q)^{-1} + \sigma_y^2 I)^{-1}$

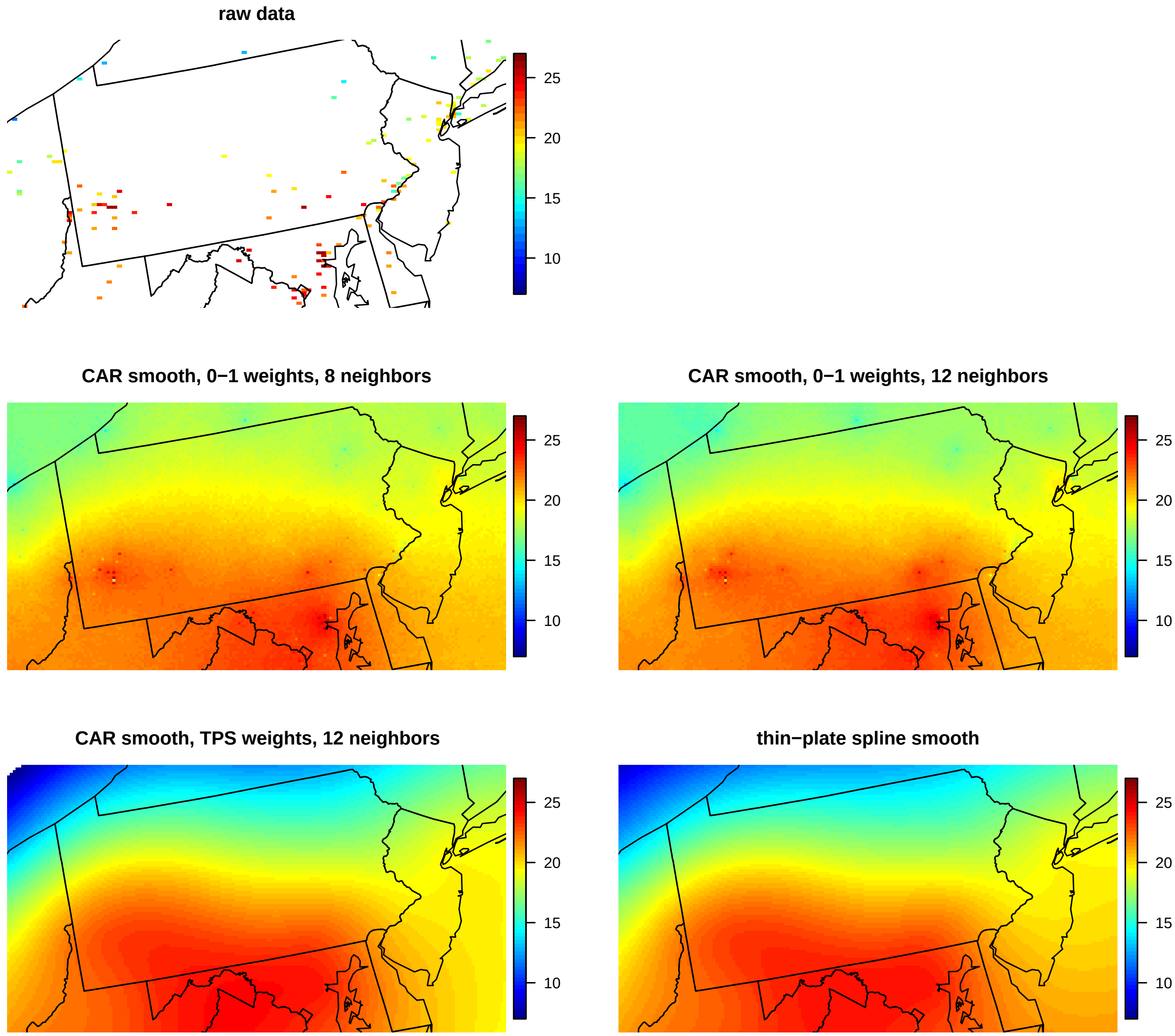

Figure 1. (top row) Raw data (average PM2.5 concentrations for July 2004 for Pennsylvania, USA). (middle row) CAR-based smooth of the gridded raw data using 0-1 weights and either 8 or 12 neighbors. (bottom row) on left, CAR-based smooth with the weights of Yue and Speckman. On right, a thin-plate spline smooth using gam() from the mgcv package in R.

where I substitute $(\kappa Q)^{-1} = C$ for the usual kriging covariance matrix. For example, for the weights of Yue and Speckman, on the domain of a torus to avoid boundary effects, the diagonal elements of Q are 20 and the off-diagonals are -8, 2, and 1 for the appropriate neighbor pairs. Of course we cannot invert Q, so re-express $S = (\kappa Q)^{-1}(\kappa Q)(\kappa Q + \sigma_y^{-2} I)\sigma_y^{-2} I = (\sigma_y^2 \kappa Q + I)^{-1}$. This allows us to consider the smoothing characteristics of different CAR neighborhood structures.

Using the expression, we look at the implied smoothing kernels for a variety of values of κ (i.e, of $\sigma_y^2 \kappa$, the effective smoothing parameter) for different neighborhood structures and weights (Fig. 2). We see that for the higher levels of smoothing (larger κ) the TPS-based CAR model of Held/Kneib/Yue has a kernel that drops off in a smooth fashion, with reasonably high weights on nearby grid cells. In contrast, the 0-1 CAR models have kernels that drop quickly from the focal point. This is what causes the bulls-eyes in the smoothing in Fig. 1 for the 0-1 CAR models. Also note that with the increasing number of neighbors for the 0-1 approach, the weights drop very slowly to zero, which would cause some smoothing to the global mean, not a feature that we probably want in a spatial smoother. The more natural kernels from the Held/Kneib/Yue approach make sense in light of results in Rue and Held (2005), pp. 192-193, in which they show that one gets positive and negative weights when one uses CAR models to try to approximate Gaussian processes based on standard geostatistical covariance functions. One issue with the TPS-based CAR model is that I don't know if anyone has shown posterior propriety for models that use this GMRF prior.

The implications are that one would not want to use the 0-1 approach to do smoothing as it gives a very strange smoother. While I have demonstrated this on sparse data, the results carry over to the situation with data in all the grid cells. As a result, I would not recommend use of the standard 0-1 CAR model in either the sparse or dense cases, including standard disease mapping settings. In contrast the thin-plate spline-based CAR model seems to have very nice features.

References

Rue, H. and Held, L. (2005), *Gaussian Markov Random Fields: Theory and Applications*, Boca Raton: Chapman & Hall.

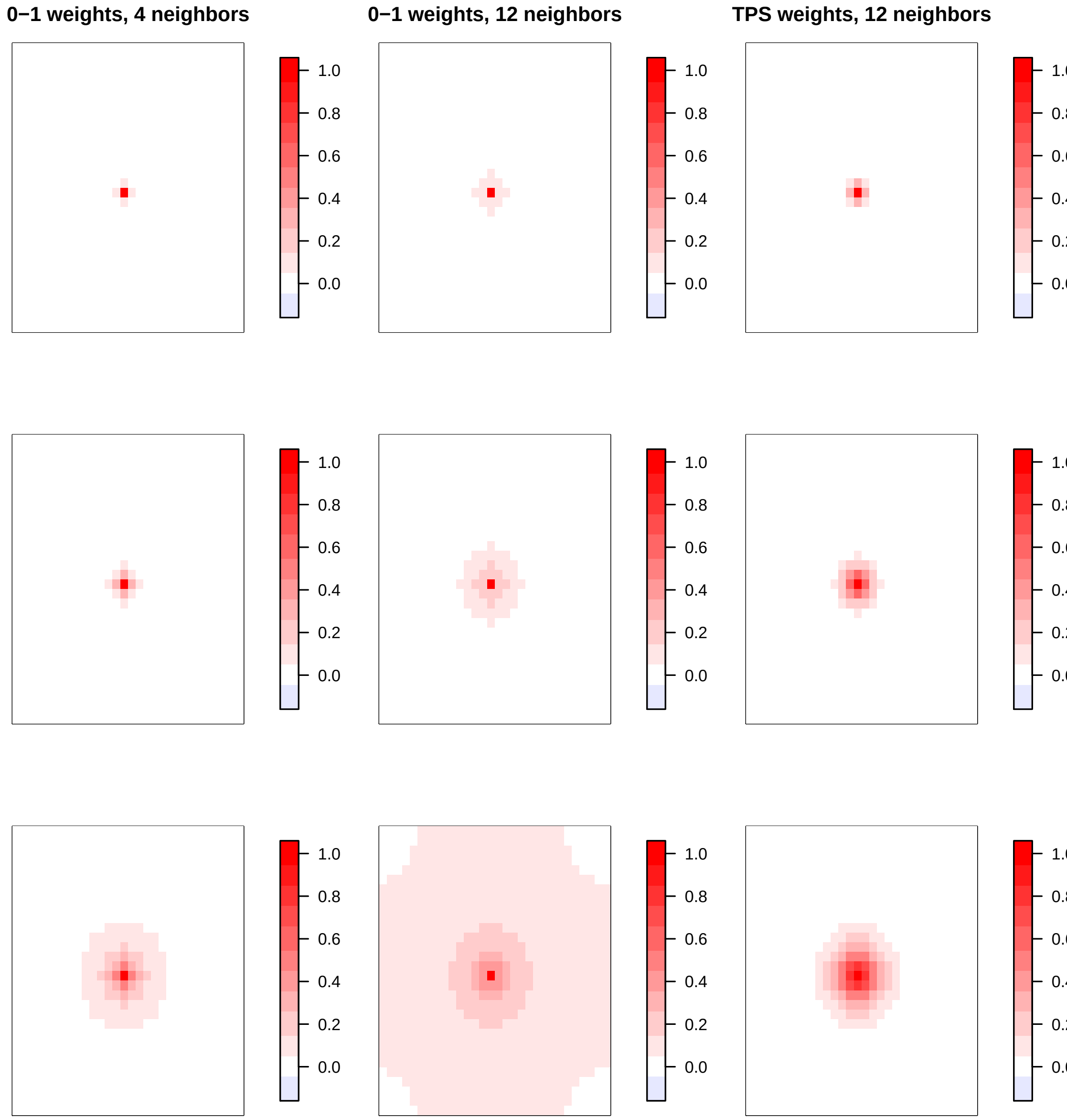

Figure 2. Smoothing kernels for three levels of smoothing: top row is minimal smoothing with $\kappa = 10$*, middle row is moderate smoothing with* $\kappa = 100$*, bottom row is more smoothing with* $\kappa = 1000$*. In all cases,* $\sigma_y^2 \equiv 0.1^2$*. Left column is based on simple 0-1 weights for four nearest neighbors, middle column on 0-1 weights using 12 neighbors, and right column the Yue and Speckman thin-plate-spline-motivated weights. Note that I use a torus as the domain to avoid boundary effects.*