Bayesian Analysis (2012) **7**, Number 2, pp. 259–262

Comment on Article by Lum and Gelfand

Rajarshi Guhaniyogi* and Sudipto Banerjee†

We congratulate the authors for a well-written article on a problem of clear and increasing scientific relevance. Quantile regression is being widely deployed in a number of disciplines to harness additional information about the relationship between the outcome and covariates at the extremes of the outcome's distribution. Modeling this in the context of spatially referenced data is challenging. We have a few comments on some aspects of the model.

The authors have argued the importance of conditional spatial quantile models to account for varying effects of covariates across quantiles. To model these conditional spatial quantiles of the response, they deploy the asymmetric Laplace process (ALP). An elegant characterization of this distribution in terms of Gaussian and Gamma random variables constitutes the premise of constructing spatial ALP's. A particularly attractive feature of this approach is the ease with which the Markov chain Monte Carlo (MCMC) samplers can be designed, not only to update model parameters but also to carry out spatial interpolation at arbitrary locations.

A point worth noting is that the ALP process can induce high correlations between two outcomes at extreme quantiles ($p \rightarrow 0$ and $p \rightarrow 1$), even if the corresponding locations are distant. In other words, outcomes arising from the tails of the distribution are assumed to be highly correlated, irrespective of how far away in space they have been collected. This is perhaps why the authors restricted their inference to quantiles between 0.2 and 0.8 in the Baton Rouge real estate data example. We wonder how serious this issue is in practice and whether their methodology is rendered invalid for applications desiring more accurate inference on higher quantiles (see Reich et al. 2011).

We, however, recognize the flexibility of the ALP process to adapt to a more generalized set up. In particular, the Gaussian-Gamma representation can be easily extended to arrive at multivariate asymmetric Laplace processes (see Kotz et al. 2001, references therein). The multivariate ALP is characterized by a multivariate normal random variable and a gamma random variable. Analogous to the univariate setting, the multivariate normal random variable can be replaced by a multivariate Gaussian process. It is not hard to envision rich multivariate spatial quantile regression models for settings where interest lies in jointly modeling a quantile for multiple spatially referenced outcomes. A natural choice for the error process is the multivariate ALP. Inference will then boil down to modeling the matrix-valued cross-covariance function of the multivariate Gaussian process.

We also appreciate the attention to large spatial datasets for which the authors have devised the asymmetric Laplace predictive process (ALPP). The Gaussian predictive process, $\tilde{w}_m(\boldsymbol{s})$ in Section 7.1, emerges as a conditional expectation of the full-rank

*Department of Statistical Science, Duke University, Durham NC, rajarshign84@gmail.com
†Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis MN, sudiptob@biostat.umn.edu

 DOI:10.1214/12-BA708A

parent spatial process (denoted $w_m(\boldsymbol{s})$ in Section 7.1). This renders an analytically tractable "residual" process $(w_m(\boldsymbol{s}) - \tilde{w}_m(\boldsymbol{s}))$ that is independent of the predictive process. We do not know of any other low-rank process that offers this luxury. This allows the authors to replace the residual process with an independent process, $\tilde{\eta}_m(\boldsymbol{s})$ in Section 7.1 or $\eta(\boldsymbol{s})$ in Section 7.2, whose variance structure agrees with that of the residual process. The structured variance adjustment to the predictive process is now a well-established method to adjust for excessive smoothness and results in more accurate estimates for the spatial (and "nugget", when it is present) variance as well as improved predictive performance.

It is, however, possible to go further and rather than simply adjust for the variances in the residual process, Sang and Huang (2012) propose using a "tapered" adjustment to the residual process. More precisely, suppose $\eta(\boldsymbol{s}) \sim GP(0, C_\nu(\cdot,\cdot))$, where $C_\nu(\cdot,\cdot)$ is a compactly supported correlation function. Using the notation used in the paper let,

$$\ddot{\epsilon}_p(\boldsymbol{s}) = \sqrt{\frac{2\xi(\boldsymbol{s})}{\tau p(1-p)}}\ddot{Z}(\boldsymbol{s}) + \frac{1-2p}{p(1-p)}\xi(\boldsymbol{s}), \tag{1}$$

where $\ddot{Z}(\boldsymbol{s}) = (1-\alpha)\left[\tilde{w}(\boldsymbol{s}) + (w(\boldsymbol{s}) - \tilde{w}(\boldsymbol{s}))\eta(\boldsymbol{s})\right] + \delta(\boldsymbol{s})$. This is likely to approximate the residual process better, especially on a local scale, although the inferential improvements in practice, at least in our experience with mean regression, are often not that substantial. It will be interesting to see its effects on quantile regression.

From a more theoretical standpoint, tapering does have some advantages in retaining the smoothness of the parent process. For example, even if the parent process is not mean square differentiable (say with an exponential covariance function), the predictive process is still mean square infinitely differentiable (see Guhaniyogi 2012). Thus, the predictive process suffers from excessive smoothness. The structured variance adjustment over-compensates for this by making the resulting process not even mean square continuous. Using an appropriately chosen tapering function, one can ensure that the resulting modified predictive process retains the same degree of smoothness as the parent process. While the structured variance adjustments yield a diagonal matrix, tapering will yield a sparse matrix corresponding to the disperison of $(w(\boldsymbol{s}) - \tilde{w}(\boldsymbol{s}))\tilde{\eta}(\boldsymbol{s})$. This makes the tapered models computationally accessible using sparse matrix computations.

Flexible nonparametric modeling of the error distribution, for example, the dependent Dirichlet process outlined in Kottas and Krnjajic (2009) can also constitute another route of exploration. Note that the shape of the asymmetric Laplace distribution is perhaps restrictive with p determining skewness of the density, hence limiting its flexibility in modeling skewness and tail behavior. In fact, the AL distribution is skewed for $p \neq 0.5$ and is symmetric for the median regression case. This may be deemed undesirable because median regression seeks to account for the skewness of the error distribution. It will be interesting to explore formulations of spatial quantile regression using the spatially dependent Dirchlet process mixture (Gelfand et al. 2005), perhaps embedded within the AL distribution. Exploration along this line might result in more flexible classes of models suitable for handling complex spatial dependencies at the tails of the response distribution.

References

Gelfand, A. E., Kottas, A., and MacEachern, S. N. (2005). "Bayesian Nonparametric Spatial Modeling with Dirichlet Process Mixing." *Journal of the American Statistical Association*, 100: 1021–1035.

Guhaniyogi, R. (2012). "On Bayesian Hierarchical Modeling for Large Spatial Datasets." *Unpublished Ph.D. Thesis-University of Minnesota.*

Kottas, A. and Krnjajic, M. (2009). "Bayesian Semiparametric modeling in Quantile Regression." *Scandanavian Journal of Statistics*, 36: 297–319.

Kotz, S., Kozubowski, T. J., and Podgorski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance*. Boston, New York: Birkhauser.

Reich, B. J., Fuentes, M., and Dunson, D. B. (2011). "Bayesian Spatial Quantile Regression." *Journal of the American Statistical Association*, 106: 6–20.

Sang, H. and Huang, J. Z. (2012). "A Full-scale Approximation of the Covariance Functions." *Journal of the Royal Statistical Society - Series B*, 74: 111–132.