Modeling thoughts for the Mexico City PM_{10} and O_3 data.

Sitting in Philadelphia, waiting for my plane to Madrid, here is what I came up with. I will first offer the continuous time version and then discretize it to hourly time with a reparametrization.

Let $Z_{1i}(t), Z_{2i}(t)$ denote the PM and ozone levels, respectively, at station i at time t. Then.

$$\begin{aligned} Z_{1i}(t) = & \quad \mathbf{X}^T(t)\boldsymbol{\beta}_{1i} + \psi_{1i} + \eta_{1i}(t) + \epsilon_{1i}(t) \\ Z_{2i}(t) = & \quad \mathbf{X}^T(t)\boldsymbol{\beta}_{2i} + \psi_{2i} + \eta_{2i}(t) + \epsilon_{2i}(t) \end{aligned}$$

Here, $\mathbf{X}(t)$ includes an intercept along with temperature and relative humidity at time t. (If it is regional, then we can add a regional subscript.) The $\boldsymbol{\beta}$'s are site level coefficients - total of 6 x 24, ψ_{1i}, ψ_{2i} is a bivariate CAR created by coregionalization of two independent CARs - 2 x 24 spatial effects plus 3 in A_ψ plus a τ_1^2 and a τ_2^2 for the independent CARs. $\eta_{1i}(t), \eta_{2i}(t)$ provides the dynamics at site i, i.e., a bivariate GP over $(0, T]$, again created by coregionalization from independent $\gamma_{1i}(t), \gamma_{2i}(t)$ each a GP with exponential covariance function, $e^{-\phi_{1i}|t-t'|}$ and $e^{-\phi_{2i}|t-t'|}$ respectively - common coregionalization matrix, A_η for all i but site level decay parameters so altogether 2 x 24 ϕ's and 3 in A_η. Finally, the $\epsilon_{1i}(t)$ and $\epsilon_{2i}(t)$ are pure Gaussian errors with variances σ_1^2 and σ_2^2, respectively - 2 more parameters.

The discretized version on hourly scale becomes:

$$\begin{aligned} Z_{1i}(t+1) = & \quad \mathbf{X}^T(t)\boldsymbol{\beta}_{1i} + \rho_{1i}Z_{1i}(t) + \psi_{1i} + \epsilon_{1i}(t+1) \\ Z_{1i}(t+1) = & \quad \mathbf{X}^T(t)\boldsymbol{\beta}_{1i} + \rho_{2i}Z_{21}(t) + \psi_{2i} + \epsilon_{1i}(t+1) \end{aligned}$$

where we have AR(1) structure with the ρ's replacing the ϕ's. The numbe rof parameters is the same, of course.

You might check my specifications. Moreover, not necessarily the best model but relatively low dimensional and certainly computationally straightforward. We can tune it better after we see what fitting it looks like. Thoughts?- Alan