Supplemental Material for "An analysis of asthma hospitalizations, air pollution, and weather in Los Angeles County, California"

MCMC sampler and prediction algorithms

Using notations for distributions analogous to Gelman et al. (2004), the posterior distribution for model (1) is

$$p(\boldsymbol{\beta}, \boldsymbol{u}, \sigma^2, \tau^2, \phi \,|\, \boldsymbol{y}) \propto Unif(\phi \,|\, a_\phi, b_\phi) \times IG(\tau^2 \,|\, a_\tau, b_\tau) \times IG(\sigma^2 \,|\, a_\sigma, b_\sigma) \times N(\boldsymbol{\beta} \,|\, \boldsymbol{\mu}_\beta, \Sigma_\beta) \times N(\boldsymbol{u} \,|\, \mathbf{0}, \sigma^2 R(\boldsymbol{\phi})) \times \prod_{i=1}^{n} N(y(t_i) \,|\, \boldsymbol{x}(t_i)'\boldsymbol{\beta} + u(t_i), \tau^2). \tag{1}$$

Estimation of (1) customarily proceeds using a Markov chain Monte Carlo (MCMC) algorithm. Often a marginalized likelihood, obtained by integrating out the temporal effects $\boldsymbol{u}$, is used to allow the algorithm to operate on a much lower-dimensional parameter space. The posterior distribution $p(\boldsymbol{\beta}, \sigma^2, \tau^2, \phi \,|\, \boldsymbol{y})$ is proportional to

$$Unif(\phi \,|\, a_\phi, b_\phi) \times IG(\tau^2 \,|\, a_\tau, b_\tau) \times IG(\sigma^2 \,|\, a_\sigma, b_\sigma) \times N(\boldsymbol{\beta} \,|\, \boldsymbol{\mu}_\beta, \Sigma_\beta) \times N\left(\boldsymbol{y} \,|\, X\boldsymbol{\beta}, \sigma^2 R(\boldsymbol{\phi}) + \tau^2 I_n\right), \tag{2}$$

where $\boldsymbol{y} = (y(t_1) \ldots y(t_n))'$, X is the matrix of regressors whose i-th row is $\boldsymbol{x}(t_i)'$, and $\boldsymbol{I}_n$ is the $n \times n$ identity matrix. We update $\boldsymbol{\beta}$ from its full conditional distribution $N(\boldsymbol{\mu}_{\boldsymbol{\beta}|\cdot}, \Sigma_{\boldsymbol{\beta}|\cdot})$, where

$$\Sigma_{\boldsymbol{\beta}|\cdot} = [\Sigma_{\boldsymbol{\beta}}^{-1} + X'\Sigma_{\boldsymbol{y}}^{-1}X]^{-1} \text{ and } \boldsymbol{\mu}_{\boldsymbol{\beta}|\cdot} = \Sigma_{\boldsymbol{\beta}|\cdot} X'\Sigma_{\boldsymbol{y}}^{-1}\boldsymbol{y}\ , \tag{3}$$

where $\Sigma_{\boldsymbol{y}} = \sigma^2 R(\phi) + \tau^2 I_n$. The remaining parameters can be updated using Metropolis steps, where the target distribution for any (set of) parameter(s) is proportional to the product of the terms in (2) that involve that (those) parameter(s). These will yield posterior samples of $\Omega = \{\boldsymbol{\beta}, \sigma^2, \phi, \tau^2\}$.

In this sampling scheme, the temporal random effects $\boldsymbol{u}$ are not sampled directly. This reduces the parameter space, which results in a more efficient MCMC algorithm. A key advantage of the first stage Gaussian model (as in model 1) is that samples from the posterior distribution of $\boldsymbol{u}$ can be recovered in a posterior predictive fashion. More precisely, we seek to evaluate

$$p(\boldsymbol{u} \,|\, \boldsymbol{y}) \propto \int p(\boldsymbol{u} \,|\, \Omega, \boldsymbol{y}) p(\Omega \,|\, \boldsymbol{y})\, d\Omega\ . \tag{4}$$

Because the full conditional distribution of $\boldsymbol{u}$ in (4) is again multivariate normal, (4) is easily evaluated using *composition sampling*. To be precise, for each sample $\{\boldsymbol{\beta}^{(l)}, \sigma^{2(l)}, \tau^{2(l)}, \phi^{(l)}\}$, we draw $\boldsymbol{u}^{(l)}$ from $N(\boldsymbol{\mu}_{\boldsymbol{u}|\cdot}, \Sigma_{\boldsymbol{u}|\cdot})$, where

$$\Sigma_{\boldsymbol{u}|\cdot} = \left[\frac{R\left(\phi^{(l)}\right)^{-1}}{\sigma^{2(l)}} + \frac{I_n}{\tau^{2(l)}}\right]^{-1} \text{ and } \boldsymbol{\mu}_{\boldsymbol{u}|\cdot} = \Sigma_{\boldsymbol{u}|\cdot} \frac{\left(\boldsymbol{y} - X\boldsymbol{\beta}^{(l)}\right)}{\tau^{2(l)}}. \tag{5}$$

For predictions, if $\mathcal{T}_0 = \{t_{0,1}, t_{0,2}, \ldots, t_{0,n_0}\}$ is a collection of n_0 *new* time points, assumed to be on the same time interval as the observed data, the posterior predictive distribution of $p(\boldsymbol{u}_0 \,|\, \boldsymbol{y})$, where $\boldsymbol{u}_0 = (u(t_{0,1}), u(t_{0,2}), \ldots, u(t_{0,n_0}))'$, is

$$p(\boldsymbol{u}_0 \,|\, \boldsymbol{y}) \propto \int p(\boldsymbol{u}_0 \,|\, \boldsymbol{u}, \Omega, \boldsymbol{y}) p(\boldsymbol{u} \,|\, \Omega, \boldsymbol{y}) p(\Omega \,|\, \boldsymbol{y}) d\Omega \boldsymbol{u}. \tag{6}$$

Given posterior samples, $\{\Omega^{(l)}\}_{l=1}^{L}$, this distribution can again be obtained via composition sampling: we first draw $\boldsymbol{u}^{(l)}$ for each l as described in (4) and then draw $\boldsymbol{u}_0^{(l)}$ from $p(\boldsymbol{u}_0 \,|\, \boldsymbol{u}^{(l)}, \Omega^{(l)}, \boldsymbol{y})$, where this last distribution is derived as a conditional distribution from a multivariate normal and, hence, is again multivariate normal. More precisely, the process realizations over the new time points are conditionally independent of the observed outcomes given the realizations over the observed time points and the process parameters. In other words, $p(\boldsymbol{u}_0 \,|\, \boldsymbol{u}, \Omega, \boldsymbol{y}) = p(\boldsymbol{u}_0 \,|\, \boldsymbol{u}, \Omega)$, which is a multivariate normal distribution with mean and variance given by

$$\begin{aligned} \mathrm{E}[\boldsymbol{u}_0 \,|\, \boldsymbol{u}, \Omega] &= \mathrm{Cov}(\boldsymbol{u}_0, \boldsymbol{u})\mathrm{Var}(\boldsymbol{u})^{-1}\boldsymbol{u} = R_0(\phi)' R(\phi)^{-1}\boldsymbol{u} \\ \text{and } \mathrm{Var}[\boldsymbol{u}_0 \,|\, \boldsymbol{u}, \Omega] &= \sigma^2 \left\{ R(\boldsymbol{\phi}) - R_0(\phi)' R(\phi)^{-1} R_0(\phi) \right\} \;, \end{aligned}$$

where $R_0(\phi)'$ is the $n_0 \times n$ matrix with (i,j)-th element given by $\rho(t_{0,i}, t_j; \phi)$. Finally, given a set of covariates at a new time point t_0, samples from the posterior predictive distribution of the outcome variable, $y(t_0)^{(l)}$, are drawn from $N(\boldsymbol{x}(t_0)'\boldsymbol{\beta}^{(l)} + \boldsymbol{u}_0^{(l)}, \tau^{2(l)})$ for $l = 1, 2, \ldots, L$.

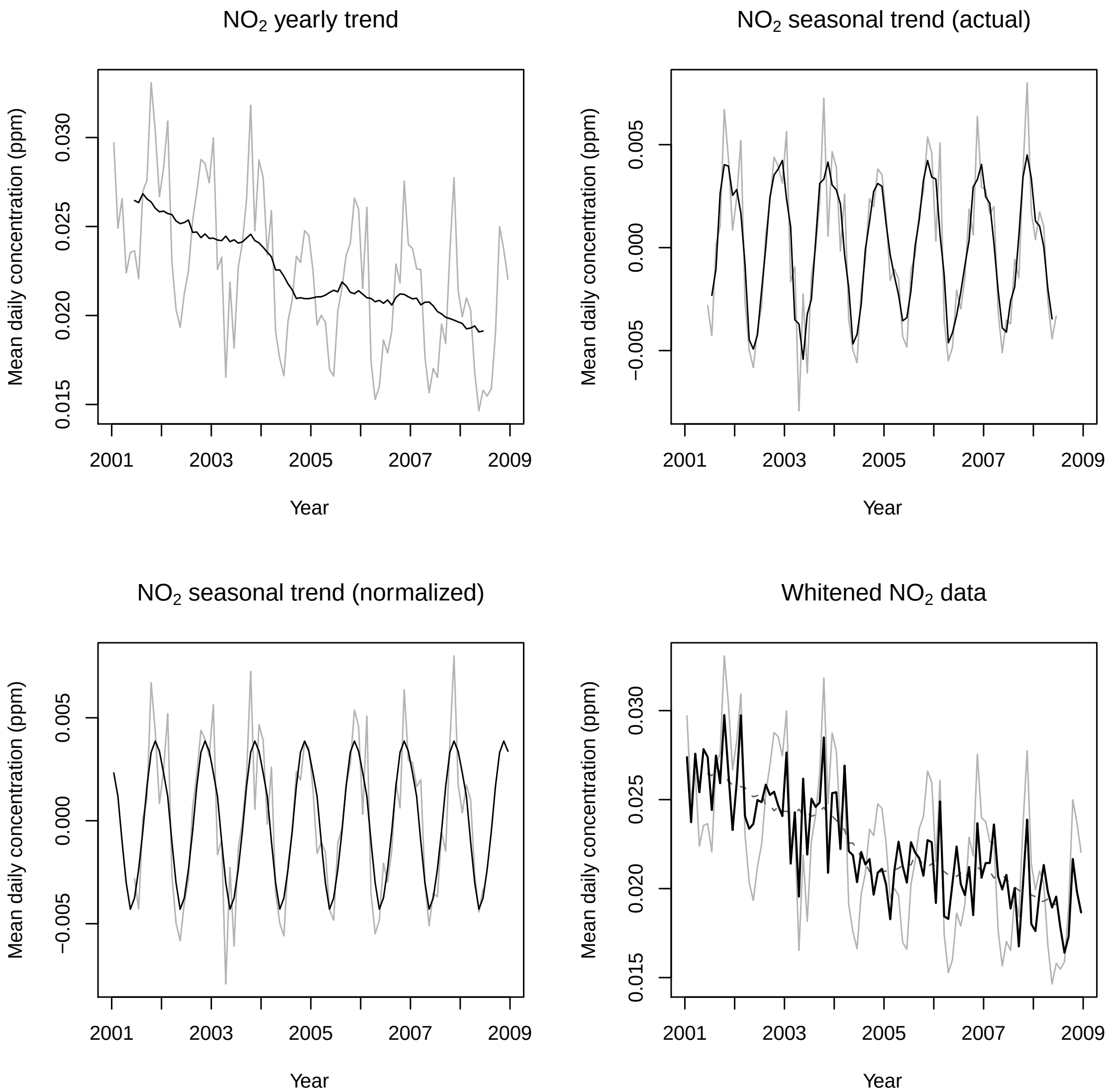

Figure 1: Example of seasonal de-trending process (NO_2) showing the filtered data (black lines, see individual plot titles) and actual data values (gray lines).

Table 1: Correlation coefficients for predictor variables prior to seasonal de-trending.

	CO	NO_2	O_3	PM_{10}	$PM_{2.5}$	Max Temp	RH
CO	1	**0.88**	**-0.85**	-0.11	0.42	**-0.62**	0.28
NO_2		1	**-0.63**	0.20	**0.60**	-0.28	0.12
O_3			1	0.34	-0.08	**0.81**	-0.18
PM_{10}				1	**0.61**	**0.66**	-0.44
$PM_{2.5}$					1	0.25	0.05
Max Temp						1	-0.41
RH							1

Table 2: Correlation coefficients for predictor variables after seasonal de-trending.

	CO	NO_2	O_3	PM_{10}	$PM_{2.5}$	Max Temp	RH
CO	1	**0.84**	-0.33	0.34	**0.68**	0.10	0.31
NO_2		1	-0.21	0.49	**0.76**	0.29	0.19
O_3			1	0.08	-0.03	0.29	-0.25
PM_{10}				1	**0.57**	**0.55**	-0.33
$PM_{2.5}$					1	0.34	0.2
Max Temp						1	**-0.51**
RH							1

Table 3: GPD scores for all models evaluated.

	Temporal effects			*Non-temporal effects*		
Model	G	P	D	G	P	D
$CO + O_3 + PM_{10}$	0.176	0.740	0.916	1.214	1.278	2.492
$NO_2 + O_3$	0.189	0.744	0.932	1.196	1.228	2.424
CO + RH	0.199	0.755	0.954	1.181	1.213	2.394
$CO + PM_{10}$	0.203	0.770	0.973	1.211	1.249	2.461
$CO + O_3$ + Tmax	0.201	0.775	0.976	1.201	1.267	2.469
CO	0.224	0.790	1.014	1.215	1.224	2.439
NO_2	0.224	0.790	1.015	1.198	1.209	2.407
$CO + O_3$	0.233	0.817	1.050	1.214	1.250	2.465
$NO_2 + O_3 + PM_{10}$ + RH	0.239	0.823	1.062	1.139	1.222	2.360
O_3 + Tmax	0.231	0.831	1.062	1.459	1.502	2.961
$CO + PM_{10}$ + RH	0.241	0.823	1.064	1.177	1.237	2.413
$CO + O_3$ + RH	0.241	0.823	1.064	1.179	1.237	2.416
$O_3 + PM_{2.5}$	0.233	0.839	1.073	1.327	1.356	2.683
$NO_2 + O_3 + PM_{10}$	0.250	0.831	1.080	1.178	1.234	2.412
NO_2 + RH	0.262	0.823	1.085	1.141	1.171	2.312
$NO_2 + O_3$ + RH	0.263	0.834	1.096	1.140	1.197	2.336
$NO_2 + O_3$ + Tmax	0.262	0.836	1.098	1.154	1.216	2.370
$NO_2 + PM_{10}$ + RH	0.268	0.842	1.109	1.137	1.196	2.333
O_3 + RH	0.254	0.864	1.118	1.369	1.408	2.777
Tmax	0.260	0.866	1.126	1.483	1.492	2.976
$O_3 + PM_{10}$	0.265	0.871	1.136	1.438	1.476	2.914
$PM_{2.5}$ + RH	0.267	0.880	1.147	1.271	1.308	2.579
O_3	0.278	0.874	1.151	1.460	1.473	2.933
$CO + O_3 + PM_{10}$ + RH	0.278	0.875	1.154	1.177	1.262	2.439
PM_{10} + RH	0.276	0.898	1.174	1.298	1.339	2.638
$O_3 + PM_{2.5}$ + Tmax	0.277	0.902	1.179	1.304	1.366	2.670
RH	0.281	0.902	1.183	1.372	1.384	2.756
CO + Tmax	0.298	0.888	1.186	1.208	1.236	2.444
$NO_2 + PM_{10}$	0.308	0.888	1.196	1.177	1.215	2.392
$O_3 + PM_{2.5}$ + RH	0.298	0.919	1.218	1.264	1.325	2.589
NO_2 + Tmax	0.332	0.894	1.226	1.158	1.192	2.350
PM_{10}	0.316	0.921	1.238	1.464	1.479	2.944
$PM_{2.5}$	0.316	0.927	1.243	1.340	1.355	2.695
$PM_{2.5}$ + Tmax	0.323	0.947	1.269	1.312	1.349	2.661
$O_3 + PM_{10}$ + RH	0.327	0.950	1.277	1.289	1.357	2.646

References

Gelman A, Carlin JB, Stern HS, Rubin DB. 2004. Bayesian Data Analysis. Second Edition. Boca Raton, FL: Chapman and Hall/CRC Press.