Combining Satellite Imagery and Numerical Model Simulation to Estimate Ambient Air Pollution

Nancy Murray[1], Howard H. Chang[1], Yang Liu[2], Heather Holmes[3]

[1]Department of Biostatistics and Bioinformatics, Emory University; [2] Department of Environmental Health, Emory University; [3] Department of Physics, University of Nevada, Reno

The Importance of Modeling $PM_{2.5}$

- Ambient fine particulate matter less than 2.5 μm in aerodynamic diameter ($PM_{2.5}$) is a common air pollutant with global importance.
- $PM_{2.5}$ is linked to various adverse health outcomes such as asthma and cardiovascular events.
- Sources of $PM_{2.5}$ include power generation, industrial operations, automobiles, and wildfires.

Approaches to Modeling $PM_{2.5}$

- Typically, $PM_{2.5}$ is modeled from regression-type approaches that only allow for one input or do not allow for uncertainty measurements.
- Bayesian hierarchical models (BHM) can fuse ground-level monitoring data with chemical transport models (CTMs) or satellite-based data.
- *Our approach*: Combine bias-corrected CTMs and satellite-based data using a Bayesian Model Averaging (BMA) framework.

Our Approach: Ensemble Modeling

- We extend the BMA forecast framework from Raftery, et al. by considering the following model:

$$\widehat{PM}_{st} = w_s \mu_{st}^{(CMAQ)} + (1 - w_s)\mu_{st}^{(AOD)}$$

where $\widehat{PM}_{st}$ is the estimated $PM_{2.5}$ value; μ_{st} is the posterior predictive mean of $PM_{2.5}$ from the $PM_{2.5}$-CMAQ BHM and the $PM_{2.5}$-AOD BHM; and w_s is the weight for the $PM_{2.5}$-CMAQ BHM at location *s*. w_s is a spatial random process, where,

$$\text{logit}(w_s) \sim \text{Gaussian}(0,\ \tau^2 e^{-\|s-s'\|/\rho})$$

Application to the Southeastern US

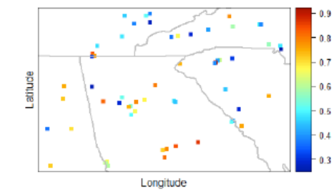

Figure 3. **Ensemble weights, w_s,** for predictions from the $PM_{2.5}$-Community Multiscale Air Quality (CMAQ) Bayesian hierarchical model at ground-monitoring locations.

Data Sources: CMAQ and AOD

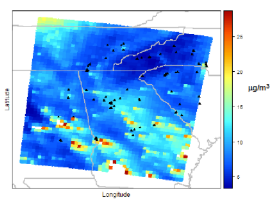

Figure 1. Simulation of $PM_{2.5}$ from the Community Multiscale Air Quality (CMAQ) model on March 17, 2005. Values at each 12 km x 12 km grid cell are linked and plotted at the closest 1 km x 1km grid cell. Black triangles indicate ground-monitoring locations.

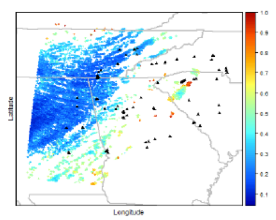

Figure 2. Satellite-derived aerosol optical depth (AOD) at 1 km x 1 km gridded resolution on March 17, 2005. Black triangles indicate ground-monitoring locations.

10-fold Cross-Validation Results

Table 1. Comparison of proposed method's performance against individual assessments in 10-fold cross-validation experiments.

Method	RMSE	Coverage of 95% PI	Average Posterior SD	R^2
$PM_{2.5}$-AOD BHM	3.40	94.07	3.30	0.78
$PM_{2.5}$-CMAQ BHM	3.14	95.05	3.28	0.81
Ensemble (CMAQ + AOD)	2.99	97.14	2.40	0.83

RMSE: root mean squared error; PI: prediction interval; SD: standard deviation; $PM_{2.5}$: particulate matter less than 2.5 $\mu g/m^3$; AOD: aerosol optical depth; BHM: Bayesian hierarchical model; CMAQ: Community Multiscale Air Quality

Conclusions and Future Work

- Spatially resolved estimates and their corresponding uncertainties are an important component of determining environmental health disparities.
- Extensions of the ensemble method: 1) allow weights to depend on spatial and temporal covariates (e.g. land use and meteorology) and 2) expand the model to allow for more than two inputs.

Acknowledgments and References

This research is supported by National Institutes of Health grant number R01ES027892.

- Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*. **133,** 1155-1174.

Corresponding author: Nancy Murray
Email: nancy.murray@emory.edu