DOES REMOTELY-SENSED AEROSOL HELP PREDICT GROUND-LEVEL $PM_{2.5}$ CONCENTRATIONS IN THE EASTERN U.S.?

Christopher J. Paciorek[1] and Yang Liu[2]

[1]Department of Biostatistics, Harvard School of Public Health, Boston, MA; [2]Department of Environmental Health, Harvard School of Public Health, Boston, MA

www.biostat.harvard.edu/~paciorek/research/presentations/presentations.html

INTRODUCTION

- Remote sensing observations of aerosol hold promise for adding information about $PM_{2.5}$ concentrations beyond that from monitors, particularly in suburban and rural areas with limited monitoring.
- AOD (aerosol optical depth) observations are frequently missing, and noisy and biased relative to $PM_{2.5}$.
- Bayesian statistical modeling holds promise for integrating AOD, $PM_{2.5}$, and GIS and weather information to predict monthly $PM_{2.5}$ concentrations on a fine grid (4 km).
- Here we assess the potential of AOD to improve predictions of $PM_{2.5}$ concentrations.
- This work is part of a larger project seeking to estimate monthly $PM_{2.5}$ concentrations at fine spatial scale as a data product for use in chronic health studies.

KEY QUESTIONS

1). Do spatial patterns in AOD reflect true patterns in $PM_{2.5}$?
- Statistical formulation: Should we model spatially-correlated (systematic) bias in AOD as a proxy for $PM_{2.5}$?

2.) Does AOD help to improve monthly and longer-term predictions of $PM_{2.5}$?
- Statistical formulation: Does including AOD in a statistical prediction model improve $PM_{2.5}$ concentrations when other information is included in the model?
- GIS-based covariates
- Meteorological covariates
- Spatial smoothing to estimate large-scale variation

SUMMARY OF RESULTS

1). Systematic, spatially-correlated error in AOD as a proxy for $PM_{2.5}$ suggests that spatial patterns in AOD do not reflect patterns in $PM_{2.5}$.
- Differences may be due to missing retrievals as well as errors from retrieval algorithm that are spatially-correlated.

2.) Inclusion of AOD in statistical prediction models for $PM_{2.5}$ has little impact on $PM_{2.5}$ predictions.
- Raw correlation between AOD and $PM_{2.5}$ reflects patterns in $PM_{2.5}$ that can be predicted using other sources of information.

ONGOING WORK

- Full cross-validatory assessment of statistical models with and without AOD to quantify effect of AOD on $PM_{2.5}$ prediction.
- Full development of monthly-scale Bayesian statistical models for $PM_{2.5}$ prediction, accounting for spatio-temporal structure.
- Monthly model predictions of $PM_{2.5}$ on a four-km grid over the eastern U.S. for 2000-2006.

DATA SOURCES

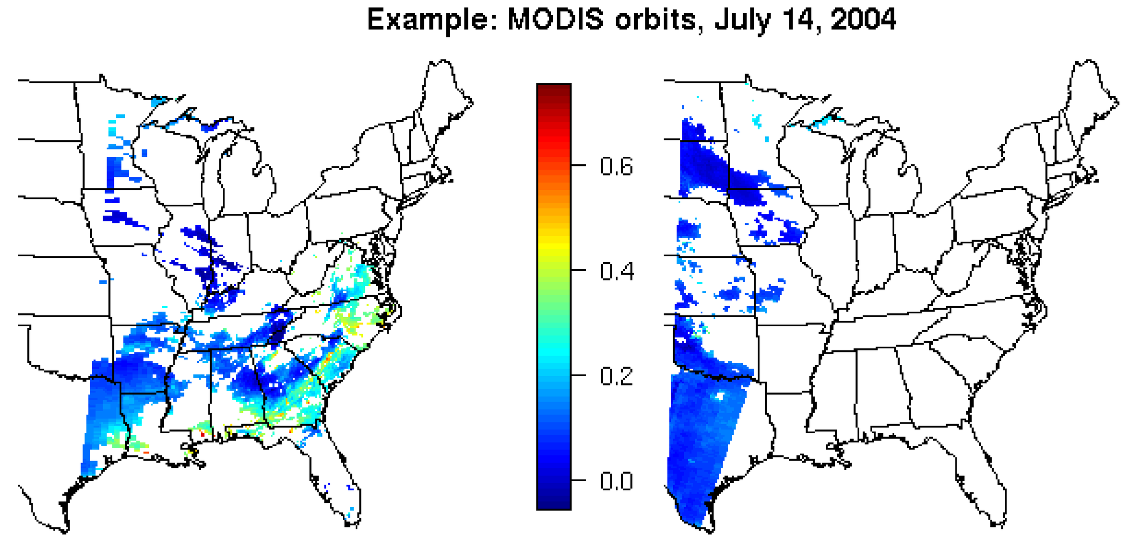
Remote Sensing Observations

- MISR AOD: 16 day orbit repeat, observations every 4-7 days at 10:30 am for a given location, 17.6 km nominal resolution
- MODIS AOD: 16 day orbit repeat, observations every 1-2 days at 10:30 am for a given location, 10 km nominal resolution
- GOES AOD: observations every half hour, 4 km nominal resolution

$PM_{2.5}$ and Covariate Information

- $PM_{2.5}$ measurements from AQS and IMPROVE: daily average, every 1, 3, or 6 days
- Weather data at 32 km, 3 hour resolution from North American Regional Reanalysis
- GIS-derived information: distance to roads (and road density) by road class, population density, land use
- NEI point source and county-level area emissions

Example: MODIS orbits, July 14, 2004



White indicates areas outside swath or with missing retrievals.

Key Question: Are such spatial patterns reflective of patterns in $PM_{2.5}$?

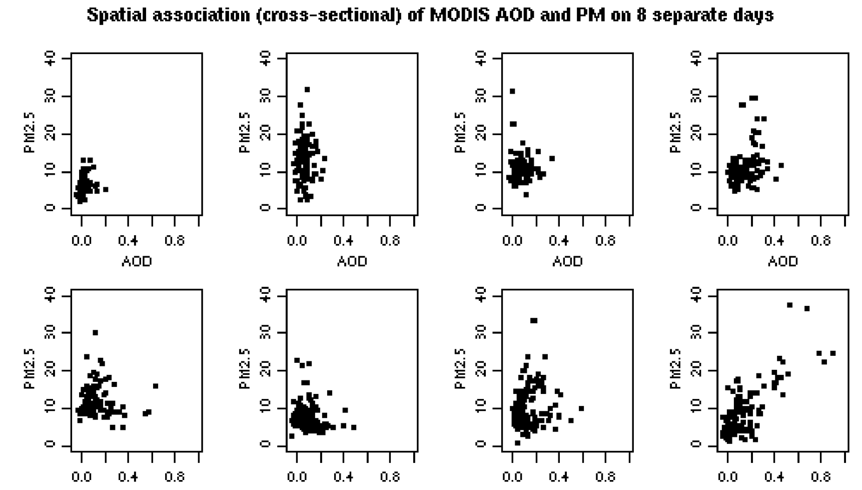
RAW ASSOCIATIONS OF AOD AND $PM_{2.5}$, EASTERN U.S.

DAILY ASSOCIATION

Associations of matched (time and space) daily AOD retrievals and $PM_{2.5}$ concentrations

	Raw AOD			Calibrated AOD		
	MODIS	MISR	GOES	MODIS	MISR	GOES
Overall correlation (longitudinal plus cross-sectional)	0.60	0.50	0.38	0.64	0.57	0.40
Average of daily (cross-sectional) correlations	0.35	0.30	0.23	0.45	0.32	0.29
Average of daily, April-October only	0.42	0.34	0.30	0.50	0.38	0.40

Spatial association (cross-sectional) of MODIS AOD and PM on 8 separate days



CALIBRATION

Involves adjusting AOD for relative humidity, planetary boundary layer, day of year, and regional spatial effects based on regression model of AOD on $PM_{2.5}$ for matched daily data from 2004.

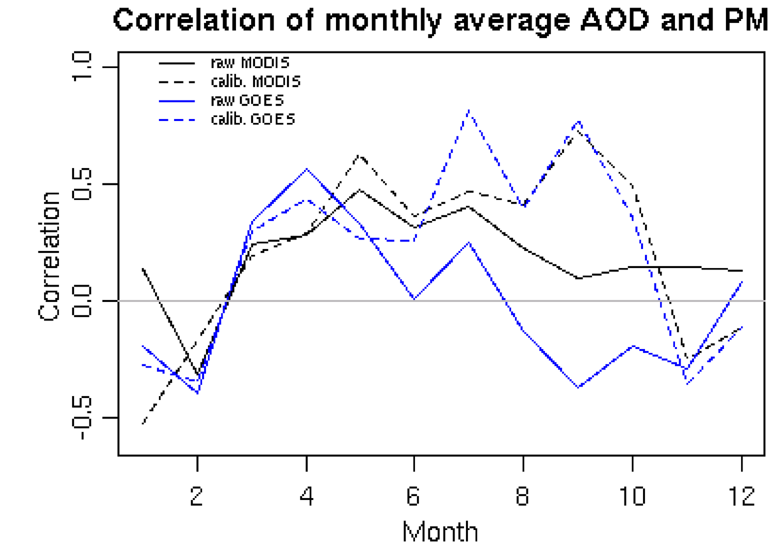
LONG-TERM ASSOCIATION

Associations (cross-sectional) of yearly average $PM_{2.5}$ with average of available AOD retrievals (not matched by day, matched in space)

	Raw AOD			Calibrated AOD		
	MODIS	MISR	GOES	MODIS	MISR	GOES
Overall correlation	0.09	0.25	-0.06	0.49	0.22	0.57
April-October only	-0.11	0.17	-0.24	0.41	0.13	0.64

Spatial (cross-sectional) associations of monthly average $PM_{2.5}$ with smoothed average of available AOD retrievals (not matched by day, matched in space)

Correlation of monthly average AOD and PM



Note: spatial smoothing of AOD done to interpolate areas with no retrievals in a month (this is primarily for MODIS and primarily in the winter).

SUMMARY

Moderate correlations are seen when comparing daily AOD and $PM_{2.5}$ matched in time and space, which includes both longitudinal and cross-sectional (spatial) association. Adjustment (calibration) for meteorology and regional spatial variation in the relationship improves correlations. However, when considering only cross-sectional association and comparing long-term average $PM_{2.5}$ concentrations with the average of available AOD retrievals, correlations are not strong. In interpreting the correlations that do exist, one possibility is that these represent large-scale associations such that AOD included in a full statistical model for $PM_{2.5}$ may not help improve $PM_{2.5}$ predictions because large-scale patterns can be estimated solely from spatial smoothing of the available $PM_{2.5}$ data.

This research is supported by HEI 4746-RFA05-2/06-7.

STATISTICAL MODELLING: MID-ATLANTIC CASE STUDY

Question 1: Do AOD patterns reflect $PM_{2.5}$ patterns?

MODEL FRAMEWORK: USE AOD AS DATA

- AOD and $PM_{2.5}$ treated as two separate datasets and both related to latent 'true' $PM_{2.5}$.
- Naturally deals with missing AOD – doesn't require imputation.
- Model structure for bias of AOD is critical – must model any systematic error (spatially-correlated bias) in AOD as a proxy for $PM_{2.5}$.

MODEL

Likelihoods for monthly average data:

$$\mathrm{PM}_i = y_i \sim \mathcal{N}\left(\mu + P(s(i)) + \sum_k f_k(z_{k,i}), \sigma^2_{y,i}\right)$$

$$\mathrm{AOD}_m = a_m \sim \mathcal{N}\left(\beta_0 + \phi(s_m) + \beta_1(\mu + P(s_m)), \sigma^2_{a,m}\right)$$

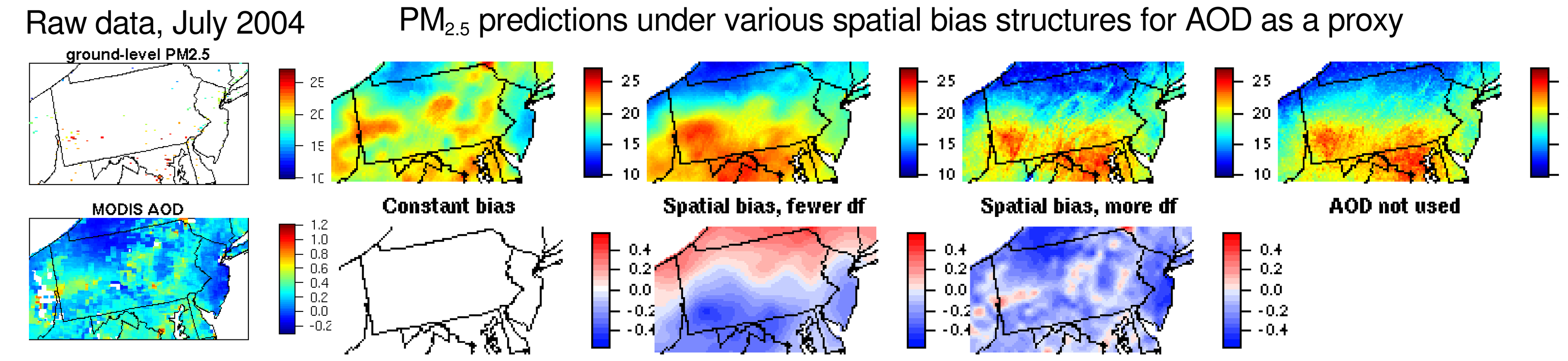
- $f_k(\cdot)$, $k = 1, \ldots, K_f$ are nonparametric regression functions of within-grid cell covariates.
- $\phi(s)$ is spatially-correlated additive bias.

Latent $PM_{2.5}$ process, $P(s)$, on 4 km grid:

$$P(s_m) = \sum_k h_k(w_k(s_m)) + g(s_m)$$

- $h_k(\cdot)$, $k = 1, \ldots, K_h$ are nonparametric regression functions of grid cell-scale covariates.
- $g(s)$ is Gaussian spatial process.

Raw data, July 2004 | $PM_{2.5}$ predictions under various spatial bias structures for AOD as a proxy



Estimated spatial bias (systematic error) of AOD as a proxy for $PM_{2.5}$

SUMMARY

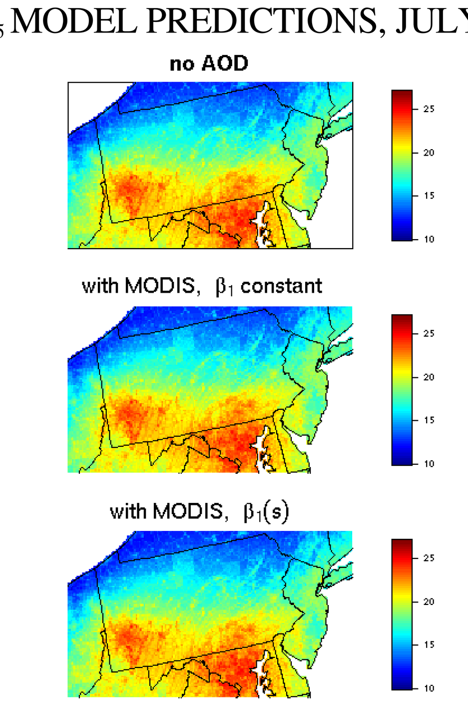
Model predictions are very sensitive to the assumptions about the nature of the bias in AOD as a proxy for $PM_{2.5}$. If no spatial (i.e., no systematic) bias is assumed, spatial patterns in AOD are reflected in $PM_{2.5}$ predictions. If systematic bias is parameterized in the model, the model estimates that the spatial pattern in AOD is distinct from spatial pattern in PM, discounting AOD as a proxy for $PM_{2.5}$. This suggests caution in interpreting spatial patterns in AOD as patterns in $PM_{2.5}$.

Question 2: Does AOD improve $PM_{2.5}$ predictions?

MODEL FRAMEWORK: USE AOD AS A PREDICTOR

- AOD treated as a regression predictor.
- Treats PM observations as gold standard with inherent calibration of AOD to $PM_{2.5}$.
- Missing AOD needs to be imputed.

$PM_{2.5}$ MODEL PREDICTIONS, JULY 2004



Note: Predictions based on GOES AOD as predictor are similar.

MODEL

Likelihood for monthly average $PM_{2.5}$:

$$\mathrm{PM}_i = y_i \sim \mathcal{N}\left(\mu + P(s(i)) + \sum_k f_k(z_{k,i}), \sigma^2_{y,i}\right)$$

Latent $PM_{2.5}$ process, $P(s)$, on 4 km grid:

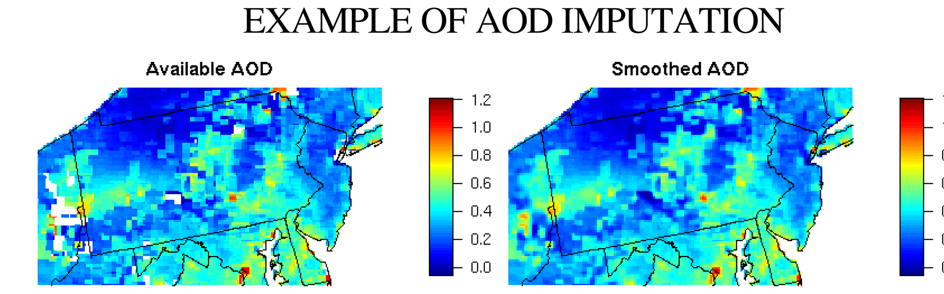
$$P(s_m) = \beta_1(s_m)A(s_m) + \sum_k h_k(w_k(s_m)) + g(s_m)$$

Imputation of AOD via spatial smoothing model, using a thin-plate spline-based GMRF model:

$$a_m \sim \mathcal{N}(\gamma_0 + A(s_m), \sigma^2_{a,m})$$
$$A(s) \sim \mathrm{GMRF}(\tau^2)$$

where $A(s)$ is smooth AOD process estimated everywhere.

EXAMPLE OF AOD IMPUTATION



SUMMARY

Estimate of beta1 (the AOD regression coefficient) is small. Including AOD as a predictor has little impact on model predictions of $PM_{2.5}$.