High Resolution Space-Time Ozone Modeling for Assessing Trends

Sujit Sahu
Southampton Statistical Sciences Research Institute, S^{3}RI
University of Southampton
http://www.maths.soton.ac.uk/staff/Sahu/

32nd Spring Lecture Series, University of Arkansas, April 2007

Ozone pollution

- Ground level ozone: bad health effects: primarily respiratory, lung function, coughing, throat irritation, congestion, bronchitis, emphysema, asthma
- Ozone is a secondary pollutant.
- Sunlight + VOC + NOx = Ozone.
- Meteorology conditions - sunlight, high temperature (so primarily from April to September), wind direction and wind speed.
- VOC's (Volatile Organic Compounds) - organic gases but really "chemicals that participate in the formation of ozone"

Ozone pollution

- Ground level ozone: bad health effects: primarily respiratory, lung function, coughing, throat irritation, congestion, bronchitis, emphysema, asthma
- Ozone is a secondary pollutant.
- Sunlight + VOC + NOx = Ozone.
- Meteorology conditions - sunlight, high temperature (so primarily from April to September), wind direction and wind speed.
- VOC's (Volatile Organic Compounds) - organic gases but really "chemicals that participate in the formation of ozone"

Ozone pollution

- Ground level ozone: bad health effects: primarily respiratory, lung function, coughing, throat irritation, congestion, bronchitis, emphysema, asthma
- Ozone is a secondary pollutant.
- Sunlight + VOC + NOx = Ozone.
- Meteorology conditions - sunlight, high temperature (so primarily from April to September), wind direction and wind speed.
- VOC's (Volatile Organic Compounds) - organic gases but really "chemicals that participate in the formation of ozone"

Ozone pollution

- Ground level ozone: bad health effects: primarily respiratory, lung function, coughing, throat irritation, congestion, bronchitis, emphysema, asthma
- Ozone is a secondary pollutant.
- Sunlight + VOC + NOx = Ozone.
- Meteorology conditions - sunlight, high temperature (so primarily from April to September), wind direction and wind speed.
- VOC's (Volatile Organic Compounds) - organic gases but really "chemicals that participate in the formation of ozone"

Ozone pollution

- Ground level ozone: bad health effects: primarily respiratory, lung function, coughing, throat irritation, congestion, bronchitis, emphysema, asthma
- Ozone is a secondary pollutant.
- Sunlight + VOC + NOx = Ozone.
- Meteorology conditions - sunlight, high temperature (so primarily from April to September), wind direction and wind speed.
- VOC's (Volatile Organic Compounds) - organic gases but really "chemicals that participate in the formation of ozone"

EPA Standard

Use daily 8-hour maximum ozone levels.

Definition

- Daily 8-hour maximum = Maximum of averages formed by 8 successive hourly ozone levels in a day.

- National Ambient Air Quality Standards (NAAQS) require:

3-year rolling average of the annual 4th highest daily 8-hour maximum is less than 80 parts per billion (ppb).

- Compliance is $\leq 80 ppb$.

So, with regard to the EPA standard, high temporal resolution is daily 8-hour maximum ozone levels.

EPA Standard

Use daily 8-hour maximum ozone levels.

Definition

- Daily 8-hour maximum = Maximum of averages formed by 8 successive hourly ozone levels in a day.

- National Ambient Air Quality Standards (NAAQS) require:

3-year rolling average of the annual 4th highest daily 8-hour maximum is less than 80 parts per billion (ppb).

- Compliance is $\leq 80ppb$.

So, with regard to the EPA standard, high temporal resolution is daily 8-hour maximum ozone levels.

Objectives in Modeling

- Find the probability that a site is non-compliant.
- What is the overall trend?
- Any reduction in ozone levels?
- Trend in ozone levels after adjusting for met-variables?
- Spatial pattern in ozone summaries.
- Illustrative, not definitive, but EPA likes the work.

Our data set

Data

- Daily 8-hour maximum ozone data from 53 sites in Ohio (mix of urban, suburban, and rural monitoring sites, several large cities separated by large rural areas).
- 50 NAMS/SLAMS (mostly urban) and 3 CASTNET (rural).
- Meteorological data from 3 CASTNET and 9 weather stations around Ohio.
- For 8 years, 1997-2004, $T = 169$ days in each year from April 15 to September 30.
- About 10% missing observations.
- Set aside data from 15 sites for validation.
- About 142,543 data points!

Location of the Sites in Ohio

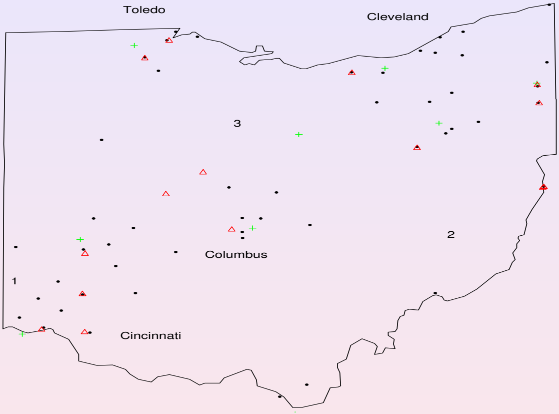


Figure: The 50 NAMS/SLAMS sites are plotted as points; three CASTNET (rural) sites are 1, 2, 3 and 9 met-sites (two outside Ohio) are $+$, and 15 validation sites are $\triangle$.

Overall Trends

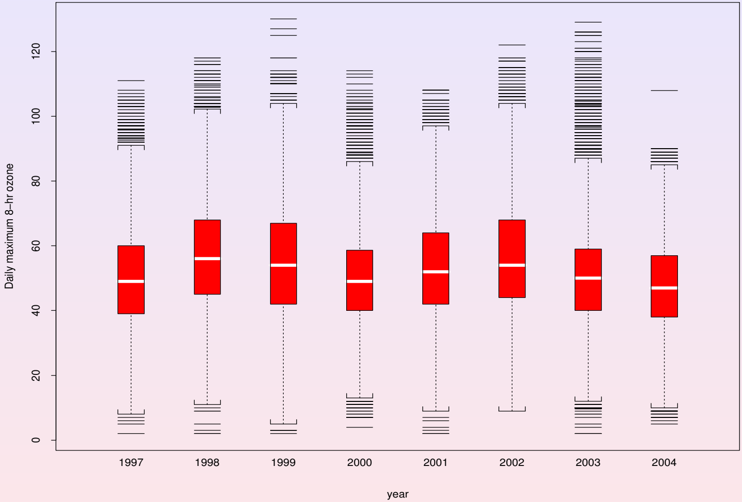
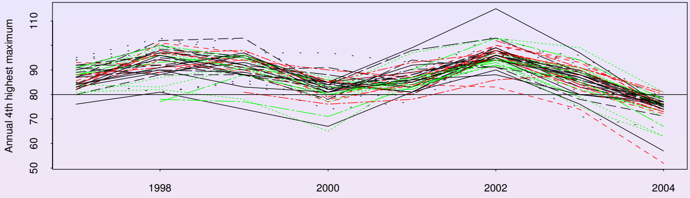


Figure: Boxplot of the daily 8-hour maximum ozone concentrations by years.

Summaries



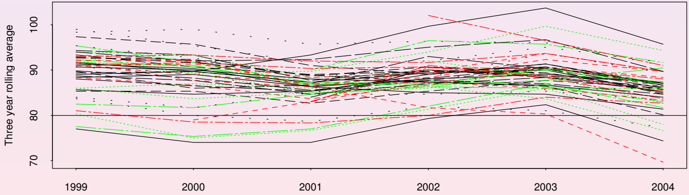


Figure: Annual 4th highest daily 8-hour maximums at 53 data sites in (a) and 3-yr rolling averages in (b).

Modeling Details

Square-Root ozone:

- observed $= Z_l(\mathbf{s}, t)$ (noisy data), $\mathbf{s} = 1, \ldots, 53$ sites, $l = 1, \ldots, 8$ years, $t = 1, \ldots, 169$ days.
- true $= O_l(\mathbf{s}, t)$ (required for inference).

Measurement error model:

$$Z_l(\mathbf{s}, t) = O_l(\mathbf{s}, t) + \epsilon_l(\mathbf{s}, t), \;\; \epsilon_l(\mathbf{s}, t) \sim N(0, \sigma_\epsilon^2).$$

- A key point: We do not want to build a dynamic model for $O_l(\mathbf{s}, t)$ given the meteorology, $x_l(\mathbf{s}, t)$, at $l, \mathbf{s}, t$. Neither do we want to build a transitional model for meteorology, i.e., a weather forecasting model.
- We (authors) would replace the weather forecasters!! (prediction will require future met.)

Modeling Details

Square-Root ozone:

- observed $=Z_l(\mathbf{s}, t)$ (noisy data), $\mathbf{s} = 1, \ldots, 53$ sites, $l = 1, \ldots, 8$ years, $t = 1, \ldots, 169$ days.
- true $= O_l(\mathbf{s}, t)$ (required for inference).

Measurement error model:

$$Z_l(\mathbf{s}, t) = O_l(\mathbf{s}, t) + \epsilon_l(\mathbf{s}, t), \ \ \epsilon_l(\mathbf{s}, t) \sim N(0, \sigma_\epsilon^2).$$

- A key point: We do not want to build a dynamic model for $O_l(\mathbf{s}, t)$ given the meteorology, $x_l(\mathbf{s}, t)$, at $l, \mathbf{s}, t$. Neither do we want to build a transitional model for meteorology, i.e., a weather forecasting model.
- We (authors) would replace the weather forecasters!! (prediction will require future met.)

Modeling Details

Square-Root ozone:

- observed $=Z_l(\mathbf{s}, t)$ (noisy data), $\mathbf{s} = 1, \ldots, 53$ sites, $l = 1, \ldots, 8$ years, $t = 1, \ldots, 169$ days.
- true $= O_l(\mathbf{s}, t)$ (required for inference).

Measurement error model:

$$Z_l(\mathbf{s}, t) = O_l(\mathbf{s}, t) + \epsilon_l(\mathbf{s}, t), \;\; \epsilon_l(\mathbf{s}, t) \sim N(0, \sigma_\epsilon^2).$$

- A key point: We do not want to build a dynamic model for $O_l(\mathbf{s}, t)$ given the meteorology, $x_l(\mathbf{s}, t)$, at $l, \mathbf{s}, t$. Neither do we want to build a transitional model for meteorology, i.e., a weather forecasting model.
- We (authors) would replace the weather forecasters!! (prediction will require future met.)

Problems in modeling

Model meteorology?

- Because modeling met is challenging!
- We shall model increment in met, δ-met, $\delta_l(\mathbf{s}, t) = \mathbf{x}_l(\mathbf{s}, t) - \mathbf{x}_l(\mathbf{s}, t-1)$.
- $p = 4$ met-variables: Daily max-temperature, relative humidity, wind speed in morning and afternoon.

Mis-alignment

- Have not observed met everywhere, only at a few data and met-stations, see Figure 1.
- No observed met in prediction sites.
- Independent (of ozone) spatial modeling of δ-met.

Problems in modeling

Model meteorology?

- Because modeling met is challenging!
- We shall model increment in met, δ-met, $\boldsymbol{\delta}_l(\mathbf{s}, t) = \mathbf{x}_l(\mathbf{s}, t) - \mathbf{x}_l(\mathbf{s}, t-1)$.
- $p = 4$ met-variables: Daily max-temperature, relative humidity, wind speed in morning and afternoon.

Mis-alignment

- Have not observed met everywhere, only at a few data and met-stations, see Figure 1.
- No observed met in prediction sites.
- Independent (of ozone) spatial modeling of δ-met.

Possible Models

Increment in ozone = function of δ-met.

Random-walk model:

$$\left[O_l(\mathbf{s},t) - \mathbf{x}_l'(\mathbf{s},t)\beta\right] = \left[O_l(\mathbf{s},t-1) - \mathbf{x}_l'(\mathbf{s},t-1)\beta\right] + \eta_l(\mathbf{s},t)$$

$$\text{i.e. } O_l(\mathbf{s},t) = O_l(\mathbf{s},t-1) + \boldsymbol{\delta}_l'(\mathbf{s},t-1)\beta + \eta_l(\mathbf{s},t).$$

Nice but, explosive!

Alternative autoregressive:

$$\left[O_l(\mathbf{s},t) - \mathbf{x}_l'(\mathbf{s},t)\beta\right] = \rho\left[O_l(\mathbf{s},t-1) - \mathbf{x}_l'(\mathbf{s},t-1)\beta\right] + \eta_l(\mathbf{s},t),$$

need to model $\mathbf{x}_l(\mathbf{s},t)$, which we seek to avoid!

Possible Models

Increment in ozone = function of δ-met.

Random-walk model:

$$\left[O_l(\mathbf{s},t) - \mathbf{x}_l'(\mathbf{s},t)\beta\right] = \left[O_l(\mathbf{s},t-1) - \mathbf{x}_l'(\mathbf{s},t-1)\beta\right] + \eta_l(\mathbf{s},t)$$

i.e. $O_l(\mathbf{s},t) = O_l(\mathbf{s},t-1) + \delta_l'(\mathbf{s},t-1)\beta + \eta_l(\mathbf{s},t).$

Nice but, explosive!

Alternative autoregressive:

$$\left[O_l(\mathbf{s},t) - \mathbf{x}_l'(\mathbf{s},t)\beta\right] = \rho\left[O_l(\mathbf{s},t-1) - \mathbf{x}_l'(\mathbf{s},t-1)\beta\right] + \eta_l(\mathbf{s},t),$$

need to model $\mathbf{x}_l(\mathbf{s},t)$, which we seek to avoid!

Our model

$$O_l(\mathbf{s}, t) = \rho O_l(\mathbf{s}, t-1) + \xi_l + \boldsymbol{\delta}_l(\mathbf{s}, t)'\beta + \eta_l(\mathbf{s}, t)$$

- $\rho O_l(\mathbf{s}, t-1)$: auto-regressive.
- ξ_l: annual intercept.
- $\boldsymbol{\delta}_l(\mathbf{s}, t)'\beta$: thought to be local adjustment for change in local met.
- $\eta_l(\mathbf{s}, t)$: spatially colored error, independent in time (l and t); $\boldsymbol{\eta}_{lt} \sim N(\mathbf{0}, \Sigma_\eta)$.
- $\Sigma_\eta(i, j) = \sigma^2_\eta \, \rho_\eta(\mathbf{s}_i - \mathbf{s}_j; \phi_\eta)$.

Need an initial condition when $l = 1$.

Our model

$$O_l(\mathbf{s}, t) = \rho O_l(\mathbf{s}, t-1) + \xi_l + \boldsymbol{\delta}_l(\mathbf{s}, t)'\beta + \eta_l(\mathbf{s}, t)$$

- $\rho O_l(\mathbf{s}, t-1)$: auto-regressive.
- ξ_l: annual intercept.
- $\boldsymbol{\delta}_l(\mathbf{s}, t)'\beta$: thought to be local adjustment for change in local met.
- $\eta_l(\mathbf{s}, t)$: spatially colored error, independent in time (l and t); $\boldsymbol{\eta}_{lt} \sim N(\mathbf{0}, \Sigma_\eta)$.
- $\Sigma_\eta(i, j) = \sigma^2_\eta \, \rho_\eta(\mathbf{s}_i - \mathbf{s}_j; \phi_\eta)$.

Need an initial condition when $l = 1$.

Initial condition

$$O_l(\mathbf{s}, 1) = \mu_l + \gamma_l(\mathbf{s})$$

- μ_l is the global level μ_l (free of $\mathbf{s}$),
- $\gamma_l(\mathbf{s})$ is the additional regional effect in year l at location $\mathbf{s}$.

Assume:

- $\boldsymbol{\gamma}_l \sim N(\mathbf{0},\ \Sigma_l)$ independently,
- $\Sigma_l = \sigma_l^2 \Sigma_\gamma$.
- $\Sigma_\gamma(i,j) = \rho_\gamma(\mathbf{s}_i - \mathbf{s}_j; \phi_\gamma)$.
- σ_l^2 is given inverse-gamma prior.

Modeling δ-met

$\delta_l(\mathbf{s}, t)$ is a p-dimensional spatial process:

- observed at 12 (3 data and 9 met sites)
- un-observed elsewhere (50 data and pred. sites).

Covariance of $\delta_l(\mathbf{s}, t) = AA' = \sum_{k=1}^{p} \mathbf{a}_k \mathbf{a}_k'$ where $A = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p)$ is lower-triangular and is called the co-regionalization matrix. Assume:

$$C(\mathbf{s}_i - \mathbf{s}_j) \equiv \mathrm{Cov}\left[\delta_l(\mathbf{s}_i, t),\ \delta_l(\mathbf{s}_j, t)\right] = \sum_{k=1}^{p} \rho_k(\mathbf{s}_i - \mathbf{s}_j; \phi_k) \mathbf{a}_k \mathbf{a}_k',$$

so that the kth component decay at rate ϕ_k spatially. In other words, $\delta_l(\mathbf{s}, t) = A\mathbf{v}_l(\mathbf{s}, t)$ where the components of $\mathbf{v}_l(\mathbf{s}, t)$ are p independent processes.

Modeling δ-met

$\delta_l(\mathbf{s}, t)$ is a p-dimensional spatial process:

- observed at 12 (3 data and 9 met sites)
- un-observed elsewhere (50 data and pred. sites).

Covariance of $\delta_l(\mathbf{s}, t) = AA' = \sum_{k=1}^{p} \mathbf{a}_k \mathbf{a}_k'$ where $A = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p)$ is lower-triangular and is called the co-regionalization matrix. Assume:

$$C(\mathbf{s}_i - \mathbf{s}_j) \equiv \text{Cov}\left[\delta_l(\mathbf{s}_i, t),\ \delta_l(\mathbf{s}_j, t)\right] = \sum_{k=1}^{p} \rho_k(\mathbf{s}_i - \mathbf{s}_j; \phi_k) \mathbf{a}_k \mathbf{a}_k',$$

so that the kth component decay at rate ϕ_k spatially.
In other words, $\delta_l(\mathbf{s}, t) = A\mathbf{v}_l(\mathbf{s}, t)$ where the components of $\mathbf{v}_l(\mathbf{s}, t)$ are p independent processes.

Modeling δ-met...

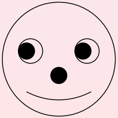
Denote the observed $\boldsymbol{\delta}_l(\mathbf{s}, t)$ by $\boldsymbol{\delta}_{lt}^{(2)}$.

$$\boldsymbol{\delta}_{lt}^{(2)} \sim N\left(\mathbf{0},\ \Sigma_{22}\right),$$

$l = 1, \ldots, 8$ and $t = 1, \ldots, 169$, independently, provides the likelihood for estimating A and $\phi_1, \ldots, \phi_p$. We use Metropolis-Hastings for this.

For the unobserved sites, use kriging to obtain:

$$\boldsymbol{\delta}_{lt}^{(1)} | \boldsymbol{\delta}_{lt}^{(2)} \sim N\left(\Sigma_{12}\Sigma_{22}^{-1}\, \boldsymbol{\delta}_{lt}^{(2)},\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right)$$



Predictions

At a new site $\mathbf{s}'$ and time t', model is:

$$Z_l(\mathbf{s}', t') \sim N\left(O_l(\mathbf{s}', t'), \sigma_\epsilon^2\right).$$

- Bayesian predictive distributions:

$$\pi(z_{\text{pred}}|z_{\text{obs}}) = \int \pi(z_{\text{pred}}|\text{par})\, \pi(\text{par}|z_{\text{obs}}) d\text{par}.$$

- Need to simulate from: $O_l(\mathbf{s}', t')$. But that needs $O_l(\mathbf{s}', t'-1)$, $O_l(\mathbf{s}', t'-2)$ and so on...
- These are simulated sequentially with $O_l(\mathbf{s}', 1)$ found using $\gamma_l(\mathbf{s}')$.
- To obtain $\gamma_l(\mathbf{s}')$ and $\delta_l(\mathbf{s}', t')$ we use kriging again

Predictions

At a new site $\mathbf{s}'$ and time t', model is:

$$Z_l(\mathbf{s}', t') \sim N\left(O_l(\mathbf{s}', t'), \sigma_\epsilon^2\right).$$

- Bayesian predictive distributions:

$$\pi(z_{\text{pred}}|z_{\text{obs}}) = \int \pi(z_{\text{pred}}|\text{par})\, \pi(\text{par}|z_{\text{obs}})\, d\text{par}.$$

- Need to simulate from: $O_l(\mathbf{s}', t')$. But that needs $O_l(\mathbf{s}', t'-1)$, $O_l(\mathbf{s}', t'-2)$ and so on...
- These are simulated sequentially with $O_l(\mathbf{s}', 1)$ found using $\gamma_l(\mathbf{s}')$.
- To obtain $\gamma_l(\mathbf{s}')$ and $\delta_l(\mathbf{s}', t')$ we use kriging again.

Predictions

At a new site $\mathbf{s}'$ and time t', model is:

$$Z_l(\mathbf{s}', t') \sim N\left(O_l(\mathbf{s}', t'), \sigma_\epsilon^2\right).$$

- Bayesian predictive distributions:

$$\pi(z_{\text{pred}} | z_{\text{obs}}) = \int \pi(z_{\text{pred}} | \text{par})\, \pi(\text{par} | z_{\text{obs}}) d\text{par}.$$

- Need to simulate from: $O_l(\mathbf{s}', t')$. But that needs $O_l(\mathbf{s}', t'-1)$, $O_l(\mathbf{s}', t'-2)$ and so on...
- These are simulated sequentially with $O_l(\mathbf{s}', 1)$ found using $\gamma_l(\mathbf{s}')$.
- To obtain $\gamma_l(\mathbf{s}')$ and $\delta_l(\mathbf{s}', t')$ we use kriging again

Predictions

At a new site $\mathbf{s}'$ and time t', model is:

$$Z_l(\mathbf{s}', t') \sim N\left(O_l(\mathbf{s}', t'), \sigma_\epsilon^2\right).$$

- Bayesian predictive distributions:

$$\pi(z_{\text{pred}}|z_{\text{obs}}) = \int \pi(z_{\text{pred}}|\text{par})\,\pi(\text{par}|z_{\text{obs}})d\text{par}.$$

- Need to simulate from: $O_l(\mathbf{s}', t')$. But that needs $O_l(\mathbf{s}', t'-1)$, $O_l(\mathbf{s}', t'-2)$ and so on...
- These are simulated sequentially with $O_l(\mathbf{s}', 1)$ found using $\gamma_l(\mathbf{s}')$.
- To obtain $\gamma_l(\mathbf{s}')$ and $\delta_l(\mathbf{s}', t')$ we use kriging again.

Predictions

At a new site $\mathbf{s}'$ and time t', model is:

$$Z_l(\mathbf{s}', t') \sim N\left(O_l(\mathbf{s}', t'), \sigma_\epsilon^2\right).$$

- Bayesian predictive distributions:

$$\pi(z_{\text{pred}}|z_{\text{obs}}) = \int \pi(z_{\text{pred}}|\text{par})\,\pi(\text{par}|z_{\text{obs}})d\text{par}.$$

- Need to simulate from: $O_l(\mathbf{s}', t')$. But that needs $O_l(\mathbf{s}', t'-1)$, $O_l(\mathbf{s}', t'-2)$ and so on...
- These are simulated sequentially with $O_l(\mathbf{s}', 1)$ found using $\gamma_l(\mathbf{s}')$.
- To obtain $\gamma_l(\mathbf{s}')$ and $\boldsymbol{\delta}_l(\mathbf{s}', t')$ we use kriging again.

Definition of Summaries

- **The annual 4*th* highest daily maximum** $f_l(\mathbf{s}) = 4th$ highest value of $\{O_l^2(\mathbf{s}, 1), \ldots, O_l^2(\mathbf{s}, T)\}$.
- **The 3-year rolling averages**

$$g_l(\mathbf{s}) = \frac{f_{l-2}(\mathbf{s}) + f_{l-1}(\mathbf{s}) + f_l(\mathbf{s})}{3}.$$

- **The metereology-adjusted summaries**

$$h_l(\mathbf{s}) = \frac{1}{T}\sum_{t=1}^{T}\left\{O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\beta\right\}^2.$$

- **The unadjusted levels**

$$u_l(\mathbf{s}) = \frac{1}{T}\sum_{t=1}^{T} O_l^2(\mathbf{s}, t).$$

Comment

- $O_l(\mathbf{s}, t) - \mathbf{x}_l'(\mathbf{s}, t)\beta$ is the natural definition of locally adjusted level though in the presence of $O_l(\mathbf{s}, t-1)$ perhaps a bit muddy.
- From a public health viewpoint all that matters is the realized O_3 levels.
- Still, it can be of value to study trends in the non-meteorology component.

Trend

Relative percentage trend values

For example,

$$c^{\text{adj}}_{97,04}(\mathbf{s}) = 100 \times \frac{h_{04}(\mathbf{s}) - h_{97}(\mathbf{s})}{h_{97}(\mathbf{s})},$$

$$c^{\text{un-adj}}_{97,04}(\mathbf{s}) = 100 \times \frac{u_{04}(\mathbf{s}) - u_{97}(\mathbf{s})}{u_{97}(\mathbf{s})}.$$

- All these summaries and trend values are simulated at each MCMC iteration for making inference.
- Using the MCMC iterates, we can assess if the relative percentage trend values are statistically significant.

Trend

Relative percentage trend values

For example,

$$c^{\text{adj}}_{97,04}(\mathbf{s}) = 100 \times \frac{h_{04}(\mathbf{s}) - h_{97}(\mathbf{s})}{h_{97}(\mathbf{s})},$$

$$c^{\text{un-adj}}_{97,04}(\mathbf{s}) = 100 \times \frac{u_{04}(\mathbf{s}) - u_{97}(\mathbf{s})}{u_{97}(\mathbf{s})}.$$

- All these summaries and trend values are simulated at each MCMC iteration for making inference.
- Using the MCMC iterates, we can assess if the relative percentage trend values are statistically significant.

Choosing ϕ_η and ϕ_γ

Use validation mean-square error

$$\text{VMSE} = \frac{1}{n_v} \sum_{i=1}^{15} \sum_{l=1}^{8} \sum_{t=1}^{169} \left(Z_l^2(\mathbf{s}_i', t) - \hat{Z}_l^2(\mathbf{s}_i', t) \right)^2 I(\text{obs})$$

where $I(\text{obs}) = 1$ if $Z_l(\mathbf{s}_i', t)$ has been observed and 0 otherwise, and

$$n_v = \sum_{i=1}^{15} \sum_{l=1}^{8} \sum_{t=1}^{169} I(\text{obs}).$$

For our data set $n_v = 5,991$.

Validation

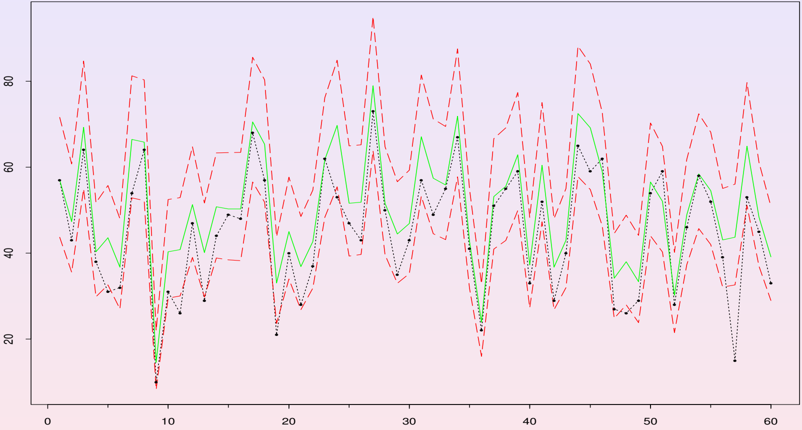


Figure: Observed data are points (black), validation predictions are solid lines and the 95% prediction intervals are broken lines.

Analysis: Parameter estimates

Table: The estimates of the parameters.

	mean	sd	95% interval
ρ	0.7783	0.0030	(0.7723, 0.7842)
β_1(temp)	0.1069	0.0021	(0.1029, 0.1109)
β_2(humidity)	–0.0126	0.0004	(–0.0134,–0.0118)
β_3(w. speed am)	–0.0025	0.0030	(–0.0083, 0.0033)
β_3(w. speed pm)	–0.0120	0.0025	(–0.0170,–0.0074)
σ_ϵ^2	0.0486	0.0007	(0.0460, 0.0487)
σ_η^2	0.3235	0.0046	(0.3149, 0.3326)

Analysis: Parameter estimates...

Table: $\mu_1, \ldots, \mu_8$.

mean	sd	95% interval
7.303	0.105	(7.098, 7.510)
6.281	0.145	(5.981, 6.555)
6.175	0.140	(5.882, 6.442)
7.171	0.094	(6.982, 7.354)
6.545	0.078	(6.388, 6.698)
7.353	0.093	(7.174, 7.542)
8.416	0.083	(8.257, 8.581)
7.091	0.112	(6.876, 7.316)

Table: $\sigma_1^2, \ldots, \sigma_8^2$.

mean	sd	95% interval
0.235	0.057	(0.146, 0.366)
0.425	0.098	(0.269, 0.649)
0.396	0.093	(0.248, 0.611)
0.196	0.047	(0.122, 0.305)
0.126	0.030	(0.078, 0.196)
0.197	0.046	(0.124, 0.302)
0.153	0.035	(0.097, 0.234)
0.278	0.063	(0.178, 0.424)

Analysis: Parameter estimates...

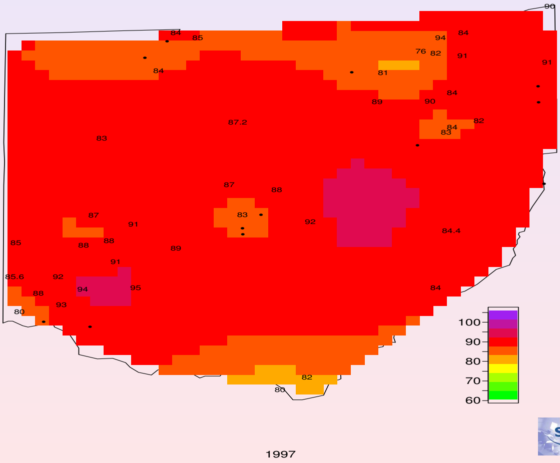
Table: $\mu_1, \ldots, \mu_8$.

mean	sd	95% interval
7.303	0.105	(7.098, 7.510)
6.281	0.145	(5.981, 6.555)
6.175	0.140	(5.882, 6.442)
7.171	0.094	(6.982, 7.354)
6.545	0.078	(6.388, 6.698)
7.353	0.093	(7.174, 7.542)
8.416	0.083	(8.257, 8.581)
7.091	0.112	(6.876, 7.316)

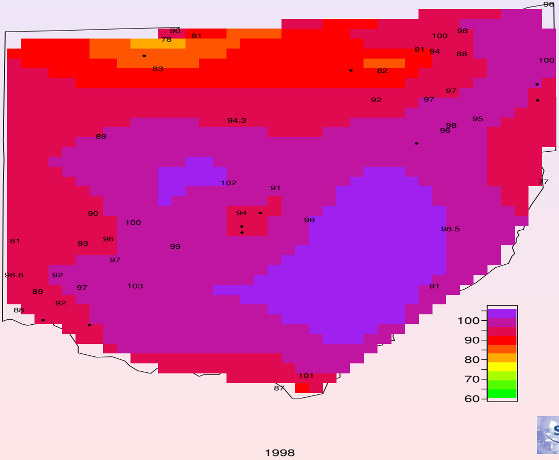
Table: $\sigma_1^2, \ldots, \sigma_8^2$.

mean	sd	95% interval
0.235	0.057	(0.146, 0.366)
0.425	0.098	(0.269, 0.649)
0.396	0.093	(0.248, 0.611)
0.196	0.047	(0.122, 0.305)
0.126	0.030	(0.078, 0.196)
0.197	0.046	(0.124, 0.302)
0.153	0.035	(0.097, 0.234)
0.278	0.063	(0.178, 0.424)

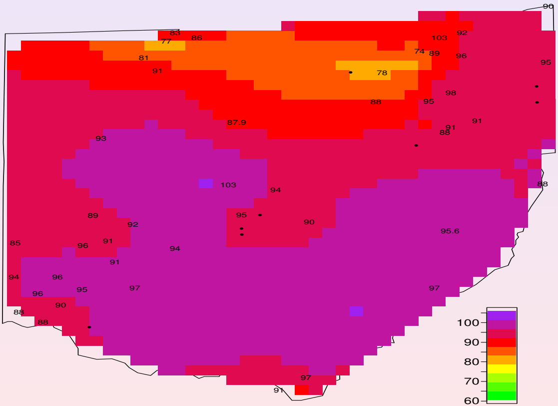
Model based interpolation of annual 4*th* highest maximum



Model based interpolation of annual 4*th* highest maximum

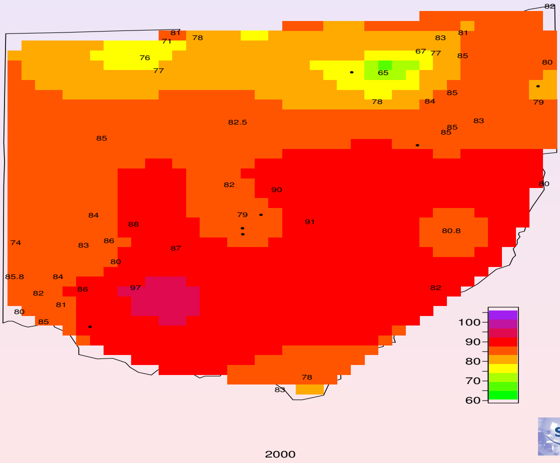


Model based interpolation of annual 4*th* highest maximum

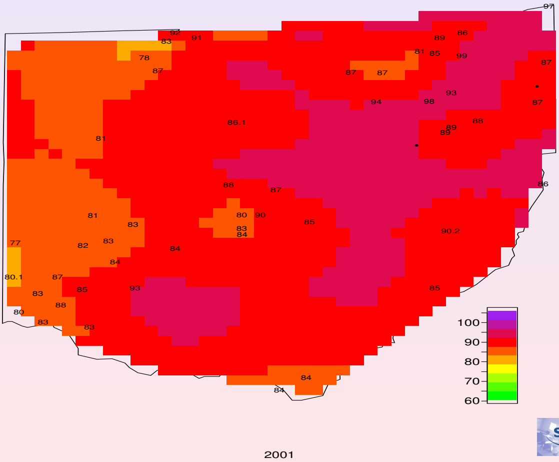


1999

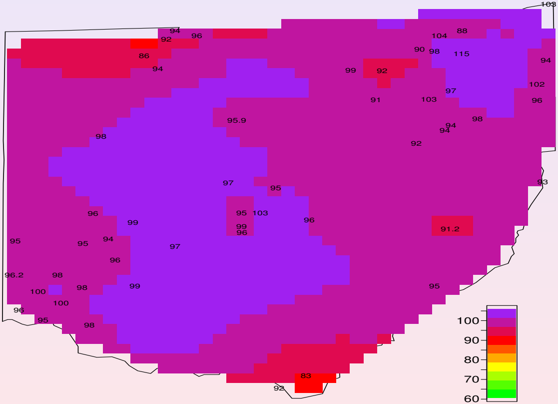
Model based interpolation of annual 4*th* highest maximum



Model based interpolation of annual 4*th* highest maximum

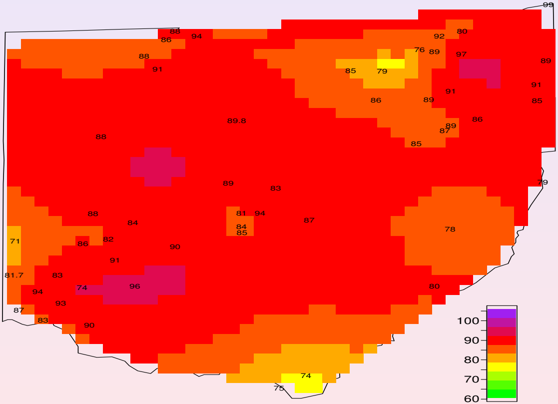


Model based interpolation of annual 4*th* highest maximum



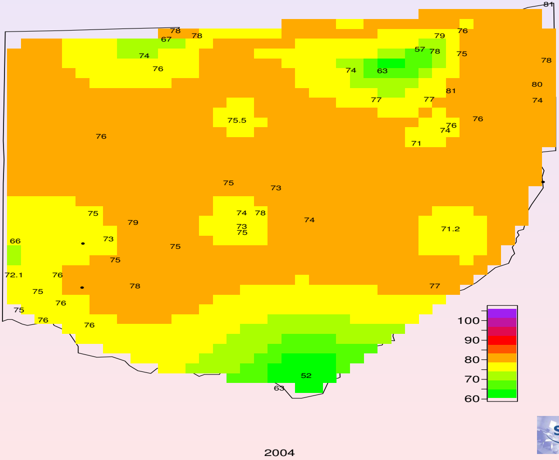
2002

Model based interpolation of annual 4*th* highest maximum

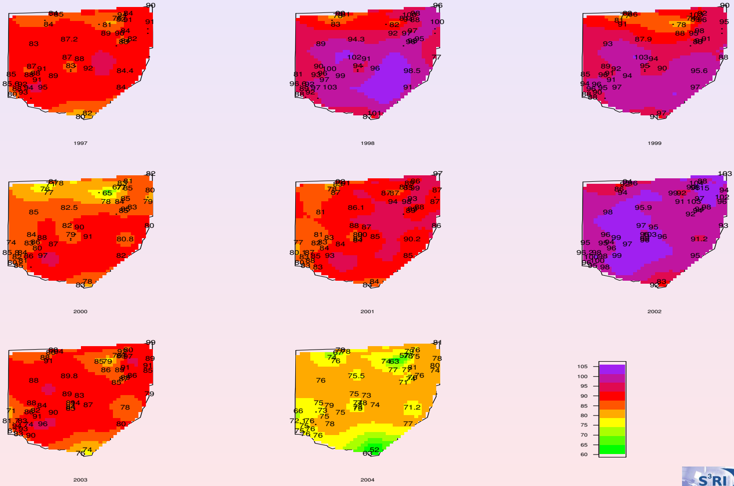


2003

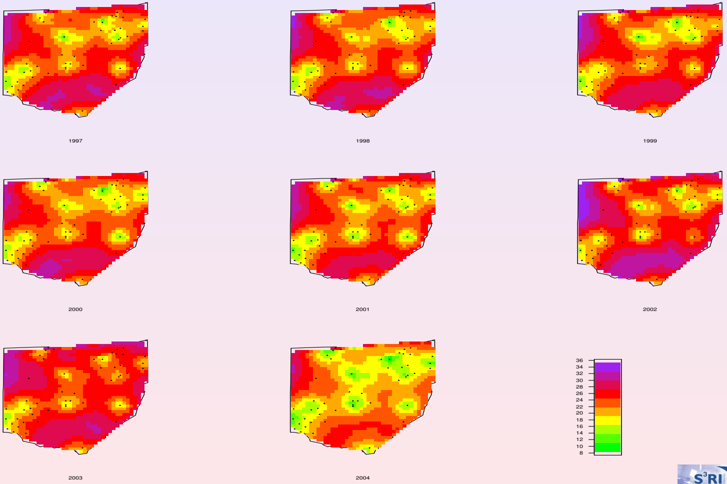
Model based interpolation of annual 4*th* highest maximum



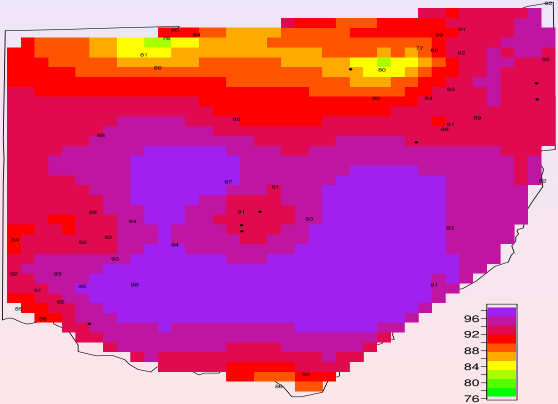
Model interpolation of the true annual 4*th* highest maximum ozone levels for 8 years.



Lengths of 95% intervals of the annual 4*th* highest maximum

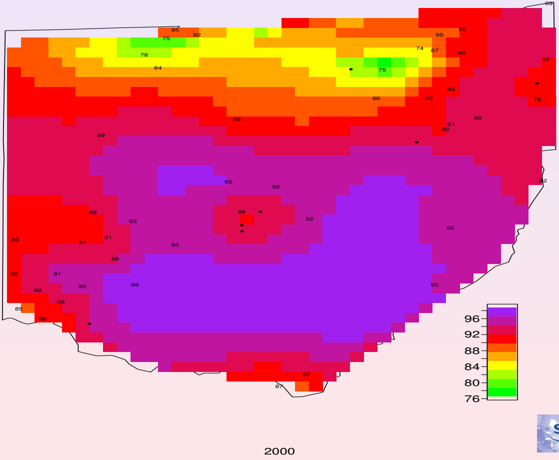


Model based interpolation of the 3-year rolling averages

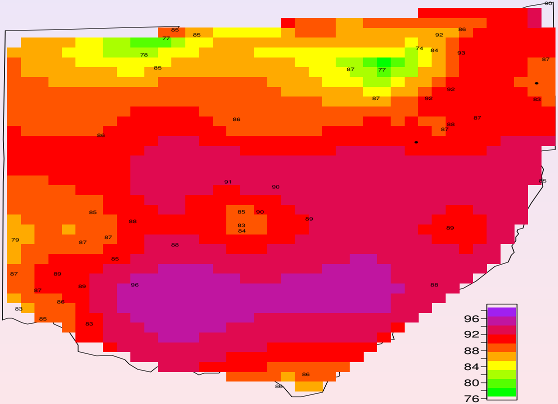


1999

Model based interpolation of the 3-year rolling averages

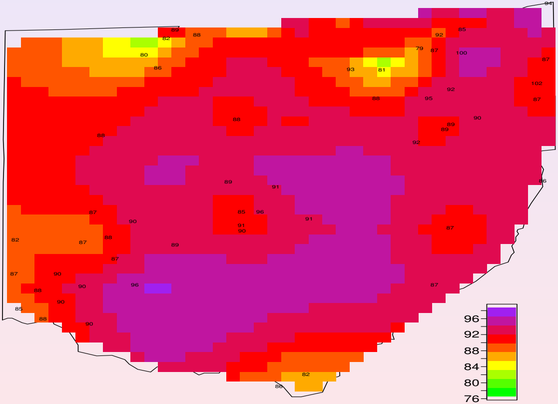


Model based interpolation of the 3-year rolling averages



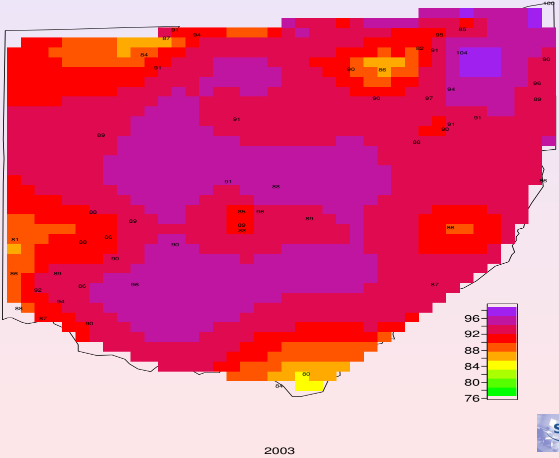
2001

Model based interpolation of the 3-year rolling averages

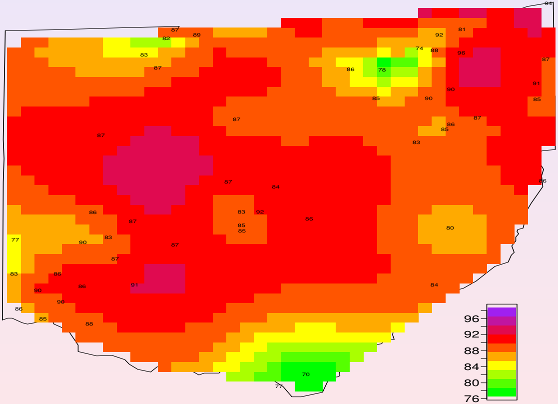


2002

Model based interpolation of the 3-year rolling averages

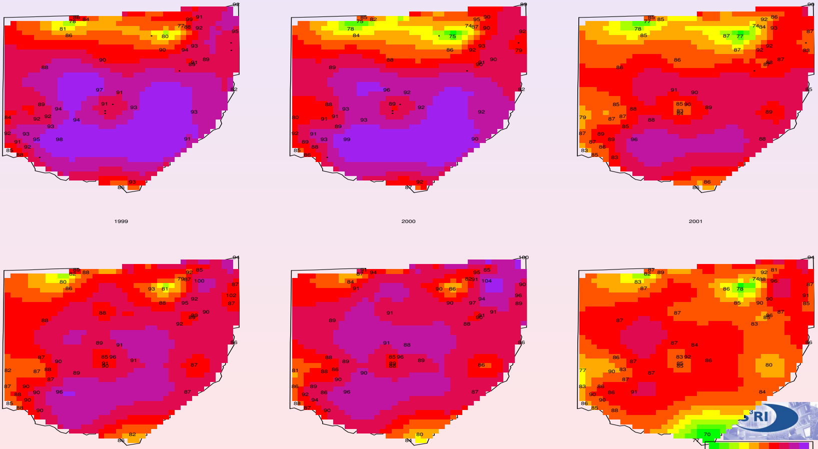


Model based interpolation of the 3-year rolling averages

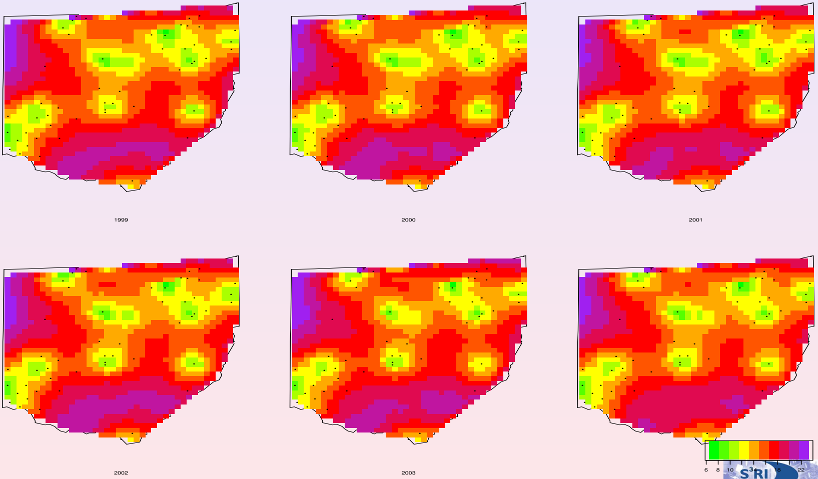


2004

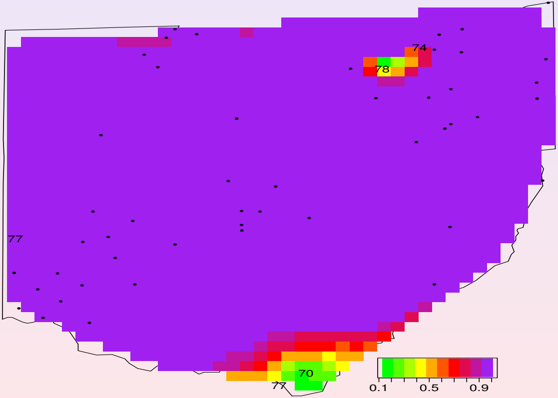
Model interpolation of the 3-year rolling averages of the true annual 4th highest maximum ozone levels.



Lengths of 95% intervals of the 3-year rolling averages



Probability that 3-year average exceeds 80 for 2004



year= 2004

Trends in Ozone

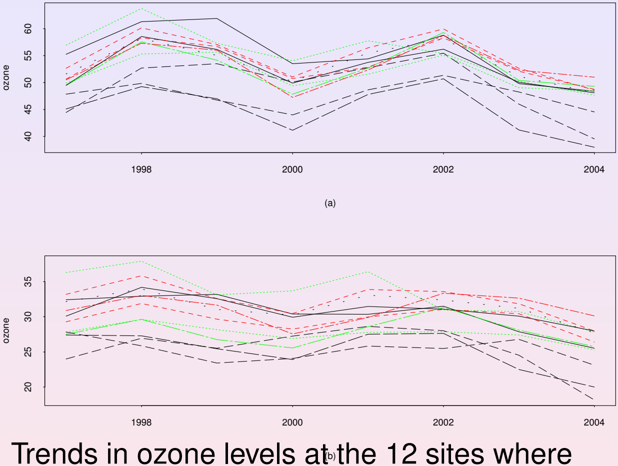


Figure: Trends in ozone levels at the 12 sites where met-variables have been observed, top panel (a) is for the un-adjusted trends; bottom panel (b) is for the met-adjusted trends.

Relative Percentage Trends

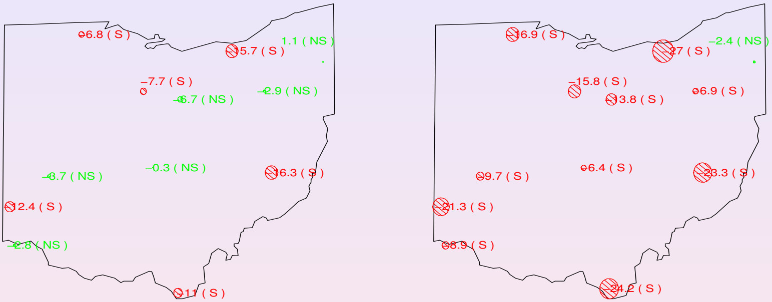


Figure: Relative percentage trends in the year 2004 (base=1997) at the 12 sites where met-variables have been observed, significant values (S) and non-significant values (NS). Radius of the circles are proportional to the values shown in the plots. Left panel (a) is for un-adjusted; right pabel (b) is for met-adjusted ones.

Discussion

- Very high resolution spatio-temporal model.
- Do not need to model absolute met variables!
- Summaries can be provided for any spatial or temporal aggregation.
- Evaluated trend without assuming any form for it.
- Methods/results are useful for many purposes, e.g. evaluating long-term emission reduction policies.
- Propagated uncertainty fully in the predictions through Bayesian modeling and computation.
- Need to do this for the eastern USA (1000 stations!)

This is joint work with Alan Gelfand (Duke University) and David Holland (USEPA), will appear in JASA and is available from www.maths.soton.ac.uk/staff/Sahu.