CMAQ $PM_{2.5}$ Forecasts Adjusted to Errors in Model Wind Fields

Eun-Su Yang[1], Sundar A. Christopher[2]

[1]Earth System Science Center, UAHuntsville

[2]Department of Atmospheric Science, UAHuntsville

Shobha Kondragunta[3], and Xiaoyang Zhang[4]

[3]NOAA/NESDIS/STAR

[4]Earth Resources Technology Inc. at NOAA/NESDIS/STAR

Presentation to JCSDA 8th Workshop

May 4, 2010

Outline

- **Motivation**
- **CMAQ $PM_{2.5}$ forecasts during the Georgia fires in 2007**
- **Wind error propagation to prediction of smoke position**
    - analyze the MM5 model wind error along the trajectory path
    - introduce a first-order autoregressive process in wind error
    - derive accumulated transport errors
- **Suggestions to improve $PM_{2.5}$ forecasts**
- **Summary**

Motivation

The position of fire smoke depends on wind fields.

←→ There is no rigorous studies on error propagation of model wind fields.

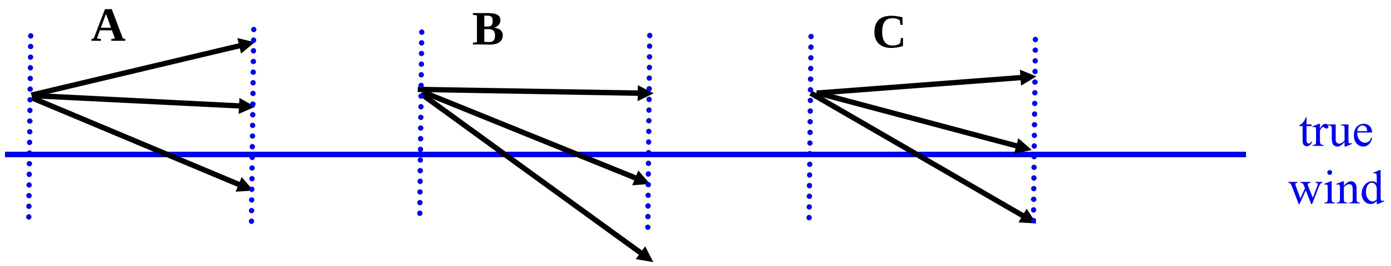
Q. If the model wind at 10 am is <u>larger</u> than the true wind,
what do you expect next hour?

The model wind at 11 am will be …

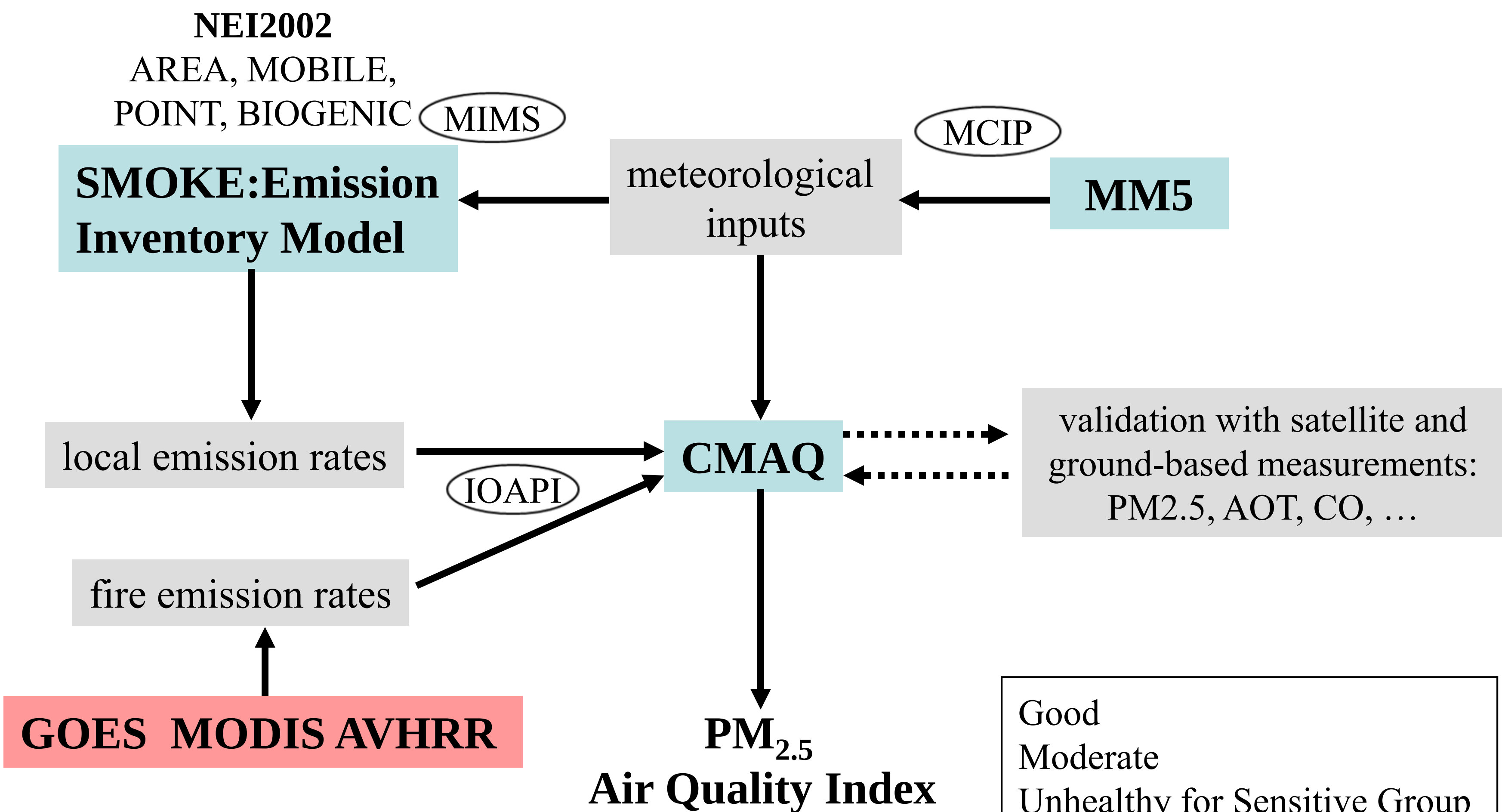
A. more likely to be larger than the true wind.

B. more likely to be smaller than the true wind.

C. equally larger or smaller than the true wind.



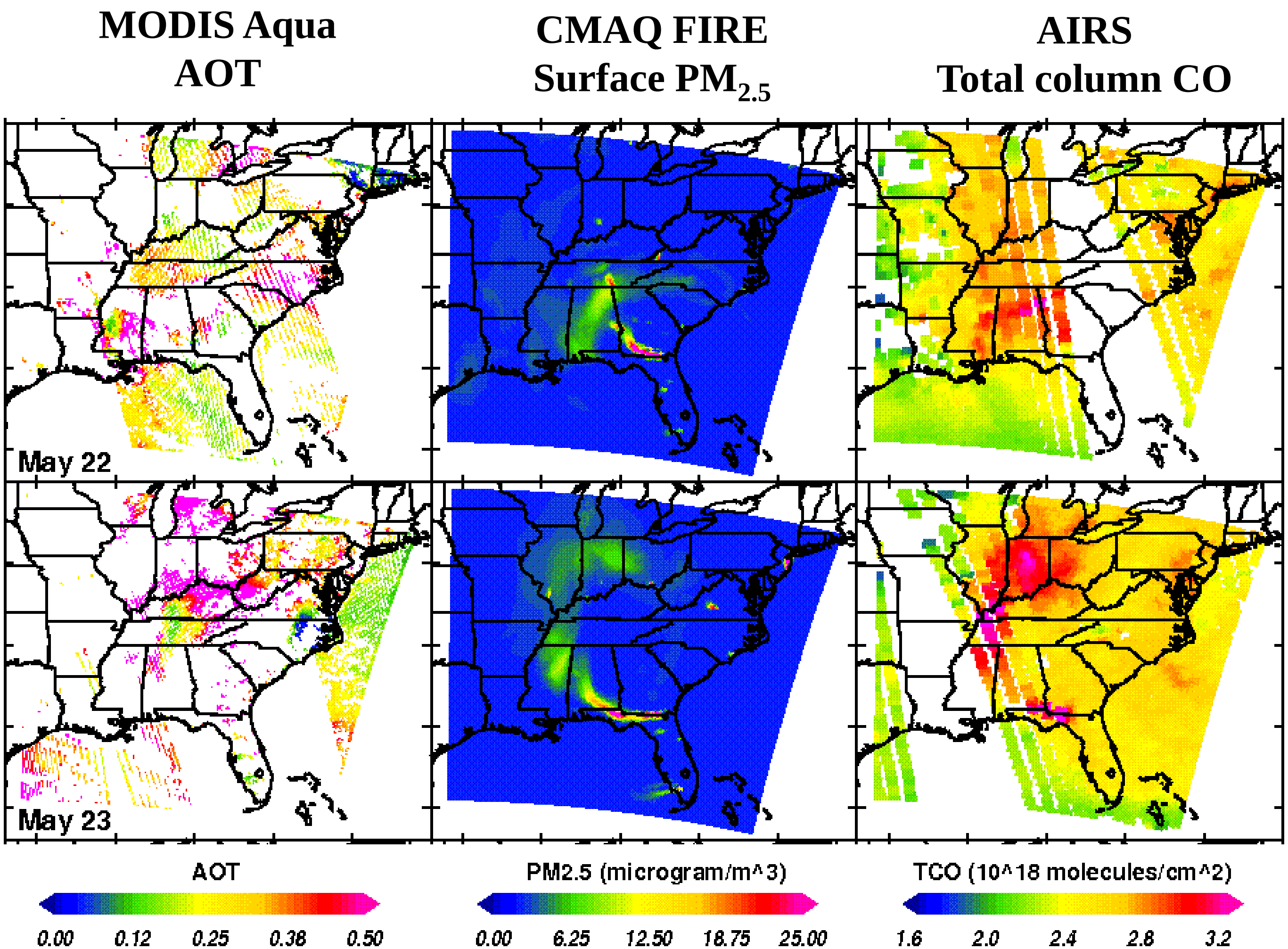
MM5/SMOKE/CMAQ for $PM_{2.5}$ forecasts



Good
Moderate
Unhealthy for Sensitive Group
Unhealthy
Very Unhealthy
Hazardous

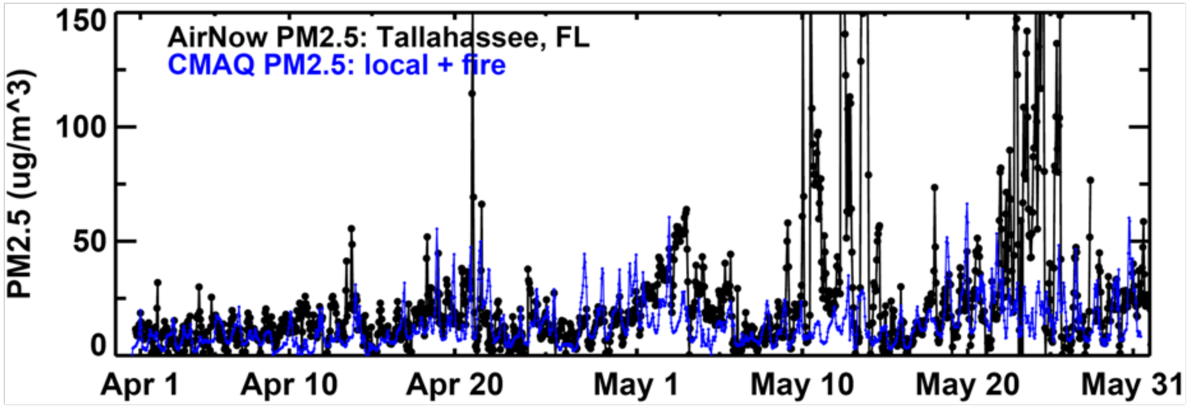
SMOKE: Sparse Matrix Operator Kernel Emissions
MCIP: Meteorology-Chemistry Interface Processor
MIMS: Multimedia Integrated Modeling System
I/O API: Input/Output Applications Programming Interface

Comparison with satellite observations

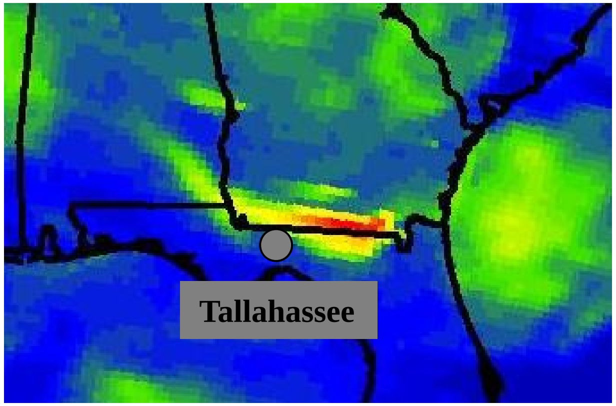


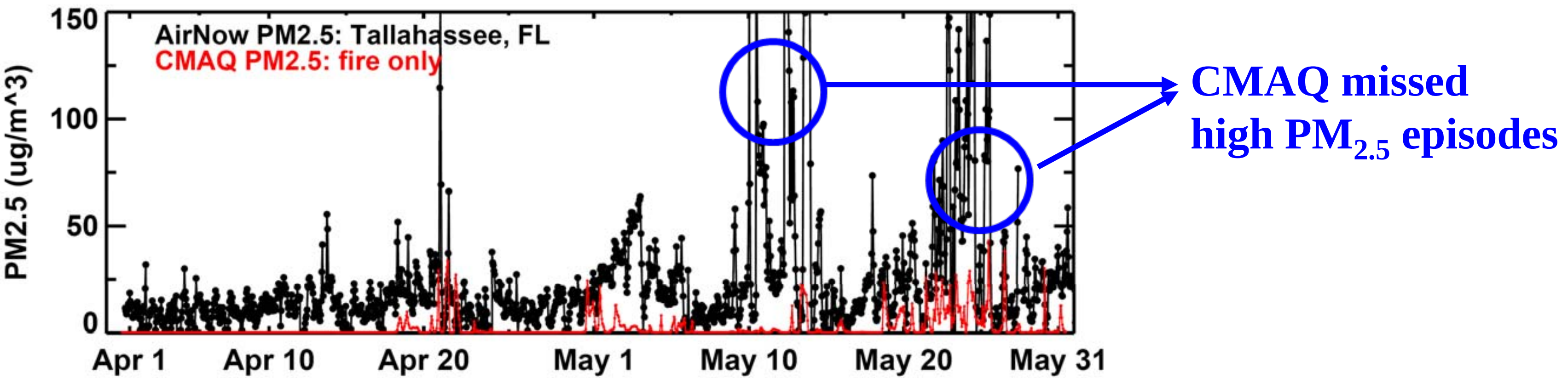
(Middle) Change in the surface-level PM2.5 simulations due to fires in μg m^{-3} at 19 UT, May 22-23, 2007. See MODIS AOT (left) and AIRS CO total column density (right) for comparison.

Comparison with ground-based $PM_{2.5}$ observations



May 22, 2007





CMAQ missed several high $PM_{2.5}$ episodes at Tallahassee.

Sensitivity of $PM_{2.5}$ to changes in initial conditions

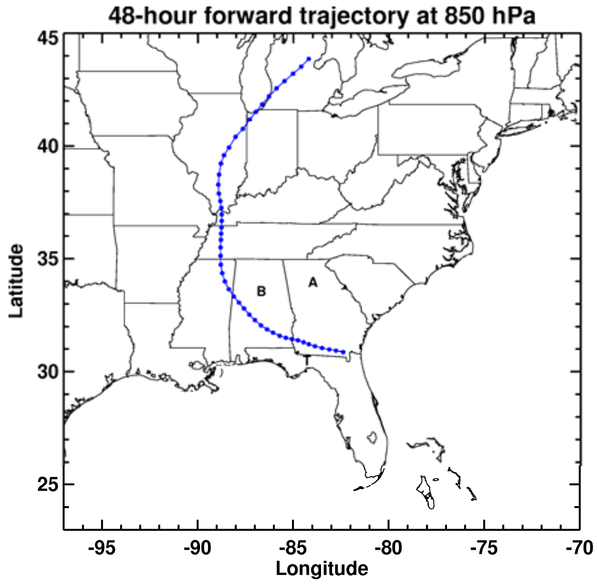
- Fire emissions are increased by 3 times.
- Fire emissions are injected
  - below the PBL height,
  - equally below and above the PBL height, or
  - into the lowest model layer.
- Fire emissions are put into 1, 9, or 16 grid cells.

→ **High $PM_{2.5}$ episodes were not well predicted at the ground stations.**

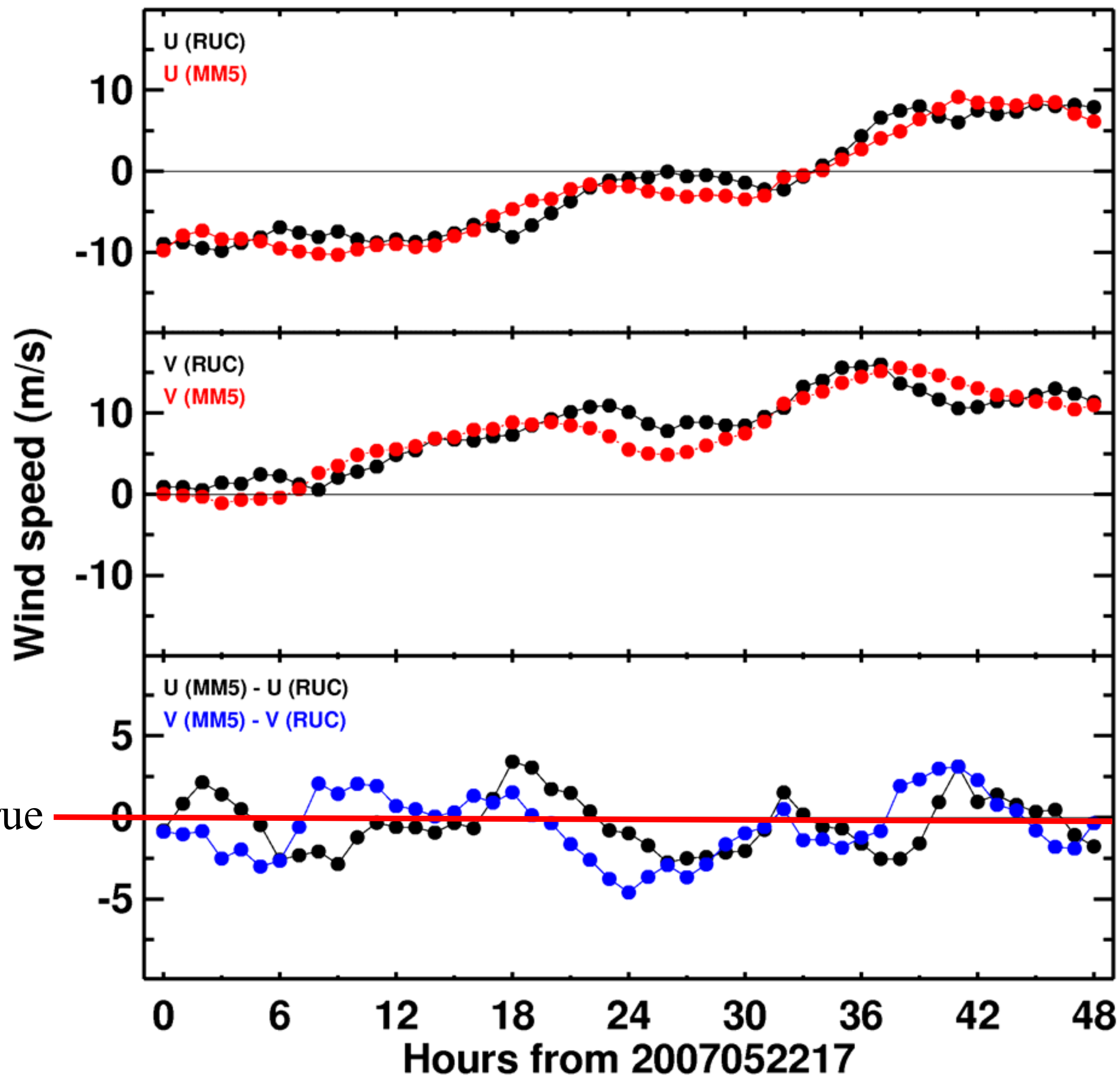
→ **CMAQ prediction of smoke position is accurate?**

→ **What is the errors in MM5-predicted winds (model winds)?**

MM5 (model) and RUC (analysis) winds along trajectory path

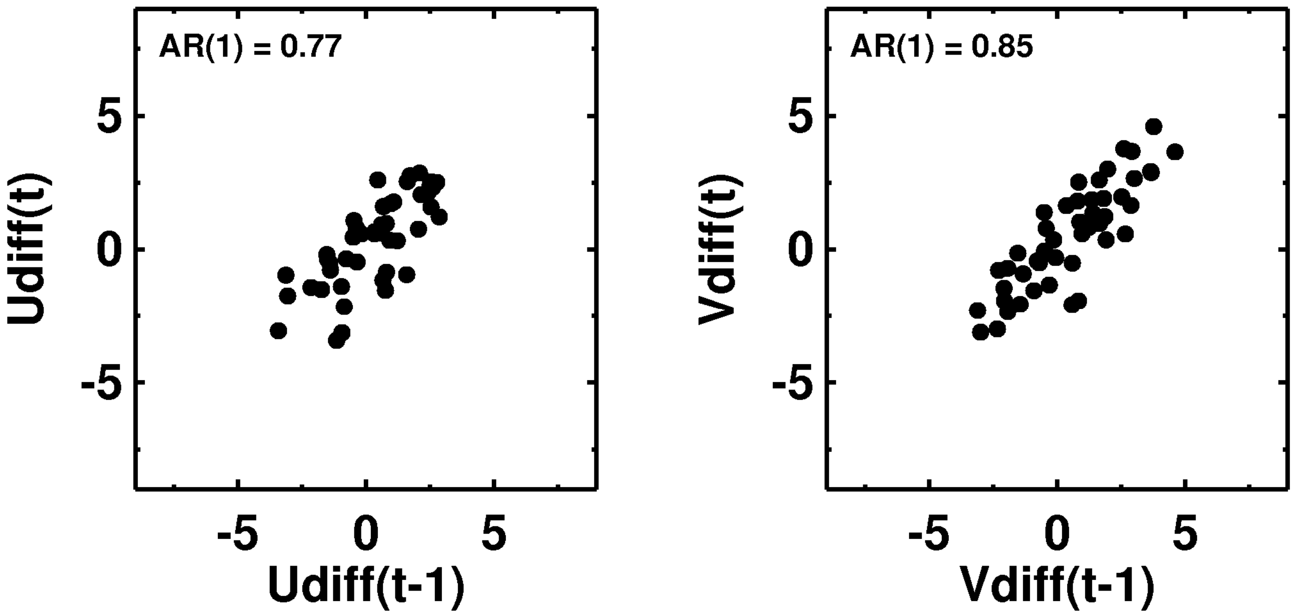


48-hour Trajectory from noon, May 22, 2007



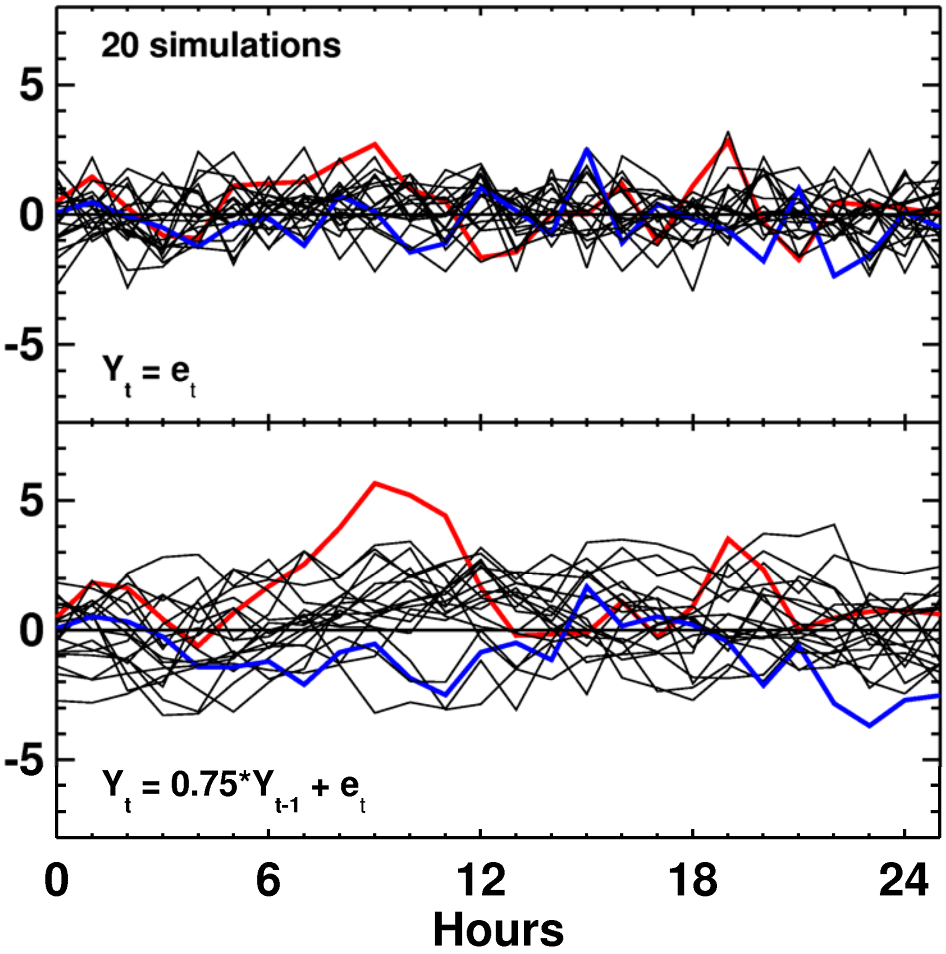
If the RUC wind is assumed true

$(U_{MM5} - U_{RUC})_{t-1}$ versus $(U_{MM5} - U_{RUC})_t$
$(V_{MM5} - V_{RUC})_{t-1}$ versus $(V_{MM5} - V_{RUC})_t$



0.75 < AR(1) < 0.95 for comparison of the MM5 wind with RUC analysis wind or ASOS surface wind observation

Error simulations without and with AR(1)



$e_t \sim N(0,1)$

- Variance increases with autocorrelation.
- Error looks drifting in case of a positive AR(1).

Derivation of variance with AR(1)

At $t = n_1 + 1$ (first data point after T_o), $E[u_t \cdot u_t]$

$$= E\left[\left(\varepsilon_{n1+1} + r\varepsilon_{n1} + r^2\varepsilon_{n1-1} + r^3\varepsilon_{n1-2} + \ldots\right) \cdot \left(\varepsilon_{n1+1} + r\varepsilon_{n1} + r^2\varepsilon_{n1-1} + r^3\varepsilon_{n1-2} + \ldots\right)\right]$$

$$= E\left[\left(\varepsilon^2_{n1+1} + r^2\varepsilon^2_{n1} + r^4\varepsilon^2_{n1-1} + r^6\varepsilon^2_{n1-2} + \ldots\right)\right]$$

$$= \sigma^2_\varepsilon + r^2\sigma^2_\varepsilon + r^4\sigma^2_\varepsilon + r^6\sigma^2_\varepsilon + \ldots$$

$$= 1 \cdot \sigma^2_\varepsilon/\left(1 - r^2\right)$$

At $t = n_1 + 2$ (second data point after T_o),

$$E[(u_t + u_{t-1}) \cdot (u_t + u_{t-1})]$$

$$= E\left[\left\{\left(\varepsilon_{n1+2} + r\varepsilon_{n1+1} + r^2\varepsilon_{n1} + r^3\varepsilon_{n1-1} + \ldots\right) + \left(\varepsilon_{n1+1} + r\varepsilon_{n1} + r^2\varepsilon_{n1-1} + r^3\varepsilon_{n1-2} + \ldots\right)\right\} \cdot \left\{\left(\varepsilon_{n1+2} + r\varepsilon_{n1+1} + r^2\varepsilon_{n1} + r^3\varepsilon_{n1-1} + \cdots\right) + \left(\varepsilon_{n1+1} + r\varepsilon_{n1} + r^2\varepsilon_{n1-1} + r^3\varepsilon_{n1-2} + \ldots\right)\right\}\right]$$

$$= E\left[\left(\varepsilon^2_{n1+2} + (1+r)^2\varepsilon^2_{n1+1} + r^2(1+r)^2\varepsilon^2_{n1} + r^4(1+r)^2\varepsilon^2_{n1-1} + \ldots\right)\right]$$

$$= \sigma^2_\varepsilon + (1+r)^2/\left(1 - r^2\right) \cdot \sigma^2_\varepsilon$$

$$= (2 + 2r) \cdot \sigma^2_\varepsilon/\left(1 - r^2\right)$$

At $t = n_1 + 3$ (third data point after T_o), $E[(u_t + u_{t-1} + u_{t-2}) \cdot (u_t + u_{t-1} + u_{t-2})]$

At $t = n_1 + n_2$ (last data point),

$$E[(u_t + u_{t-1} + u_{t-2} + u_{t-3} + \ldots) \cdot (u_t + u_{t-1} + u_{t-2} + u_{t-3} + \ldots)]$$

$$= E\left[\left\{\left(\varepsilon_{n1+n2} + r\varepsilon_{n1+n2-1} + r^2\varepsilon_{n1+n2-2} + r^3\varepsilon_{n1+n2-3} + \ldots\right) + \left(\varepsilon_{n1+n2-1} + r\varepsilon_{n1+n2-2} + r^2\varepsilon_{n1+n2-3} + r^3\varepsilon_{n1+n2-4} + \ldots\right) + \ldots\right\} \cdot \left\{\left(\varepsilon_{n1+n2} + r\varepsilon_{n1+n2-1} + r^2\varepsilon_{n1+n2-2} + r^3\varepsilon_{n1+n2-3} + \ldots\right). + \left(\varepsilon_{n1+n2-1} + r\varepsilon_{n1+n2-2} + r^2\varepsilon_{n1+n2-3} + r^3\varepsilon_{n1+n2-4} + \ldots\right) + \ldots\right\}\right]$$

$$= E\left[\left(\varepsilon^2_{n1+n2} + (1+r)^2\varepsilon^2_{n1+n2-1} + \left(1 + r + r^2\right)^2\varepsilon^2_{n1+n2-2} + \ldots.. + \left(1 + r + r^2 + \ldots + r^{n2}\right)^2\varepsilon^2_{n1+1} + r^2\left(1 + r + r^2 + \ldots + r^{n2}\right)^2\varepsilon^2_{n1} + r^4\left(1 + r + r^2 + \ldots + r^{n2}\right)^2\varepsilon^2_{n1-1} + \ldots\right)\right]$$

$$= \sigma^2_\varepsilon + (1+r)^2 \cdot \sigma^2_\varepsilon + \left(1 + r + r^2\right)^2 \cdot \sigma^2_\varepsilon + \left(1 + r + r^2 + r^3\right)^2 \cdot \sigma^2_\varepsilon + \ldots + \left(1 + r + r^2 + \ldots + r^{n2}\right)^2/\left(1 - r^2\right) \cdot \sigma^2_\varepsilon$$

$$= \left\{n_2 + 2(n_2 - 1)r + 2(n_2 - 2)r^2 + 2(n_2 - 3)r^3 + \ldots + 2(1)r^{n2-1}\right\} \cdot \sigma^2_\varepsilon/\left(1 - r^2\right) \approx \{n_2 + 2n_2 r/(1 - r)\} \cdot \sigma^2_\varepsilon/\left(1 - r^2\right)$$

$$= n_2 \cdot (1 + r)/(1 - r) \cdot \sigma^2_\varepsilon/\left(1 - r^2\right) = \sigma^2_\varepsilon/\left(1 - r^2\right) \cdot n_2 \cdot cf$$

Yang et al., 2006, JGR

Variance of accumulated deviation **without** AR(1)

$t = 1$, VAR $= 1 \cdot \sigma^2$
$t = 2$, VAR $= (2 \quad) \cdot \sigma^2$

$t = 3$, VAR $= (3 \quad) \cdot \sigma 2$

$t = 4$, VAR $= (4 \quad) \cdot \sigma 2$

$t = 5$, VAR $= (5 \quad) \cdot \sigma 2$

…

where φ is the AR(1) coefficient ($-1 < \varphi < 1$).

If $\varphi = 0$, VAR ~ t & standard deviation ~ $t^{1/2}$.

Variance of accumulated deviation **with** AR(1)

$t = 1, \ VAR = 1 \cdot \sigma^2$

$t = 2, \ VAR = (2 + 2\varphi) \cdot \sigma^2$

$t = 3, \ VAR = (3 + 4\varphi + 2\varphi^2) \cdot \sigma 2$

$t = 4, \ VAR = (4 + 6\varphi + 4\varphi^2 + 2\varphi^3) \cdot \sigma 2$

$t = 5, \ VAR = (5 + 8\varphi + 6\varphi^2 + 4\varphi^3 + 2\varphi^4) \cdot \sigma 2$

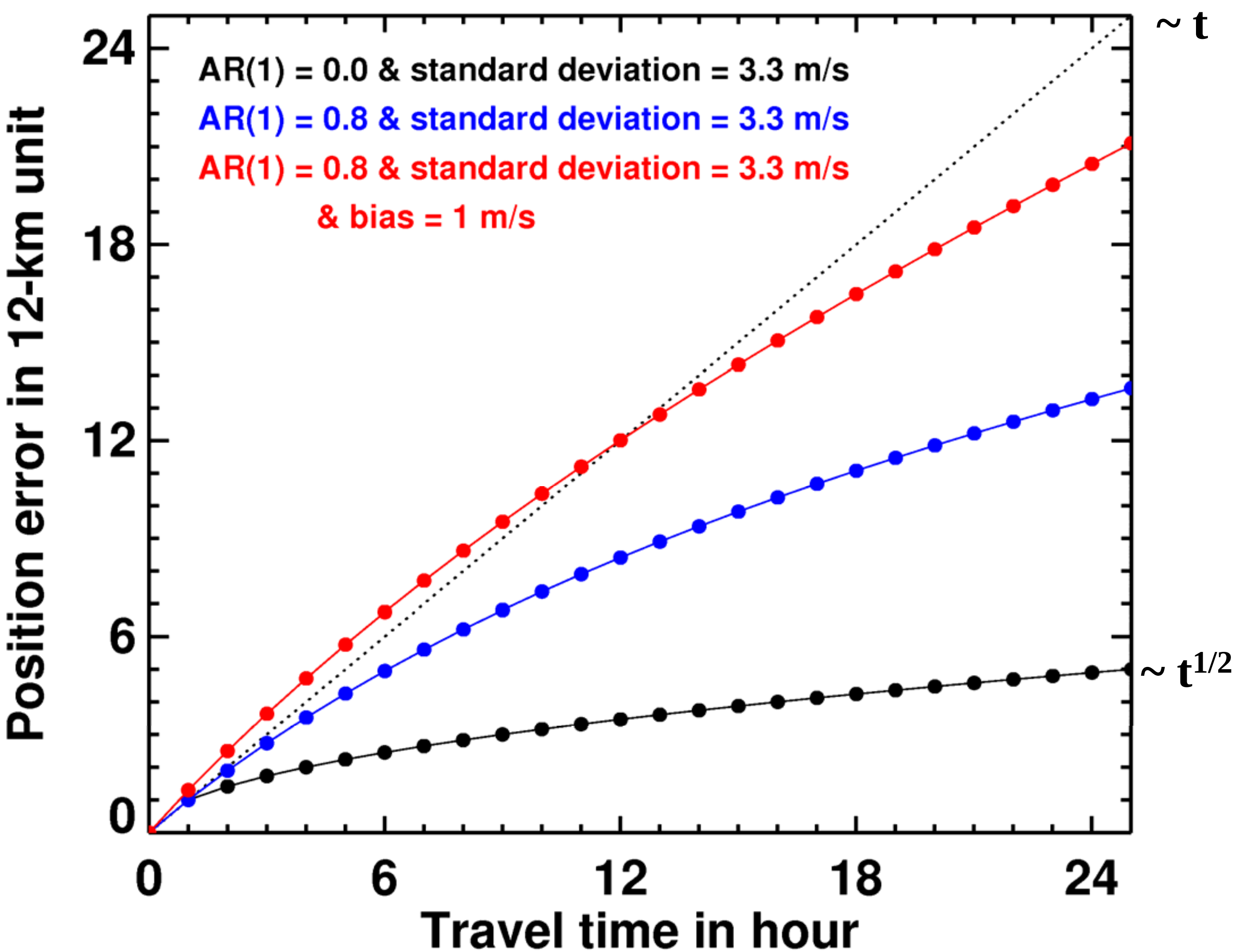
…

where φ is the AR(1) coefficient ($-1 < \varphi < 1$).

If $\varphi = 0$, VAR ~ t & standard deviation ~ $t^{1/2}$.

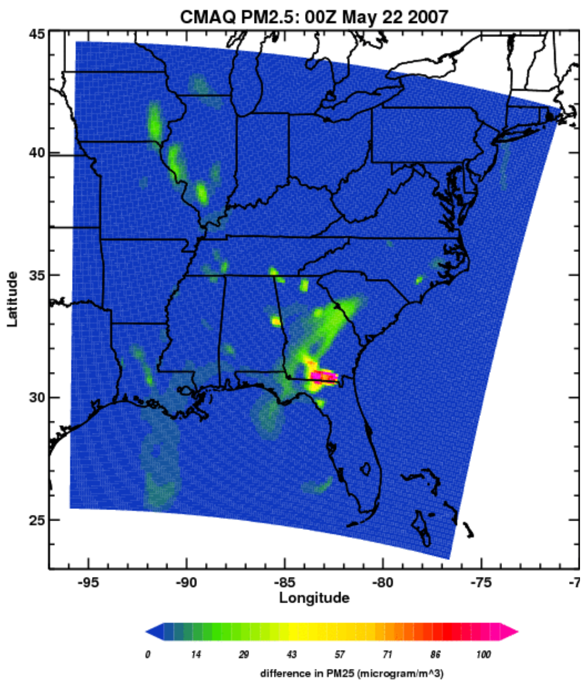
If $\varphi \rightarrow 1$, VAR $\rightarrow t^2$, standard deviation $\rightarrow$ t

Smoke position error along trajectory path

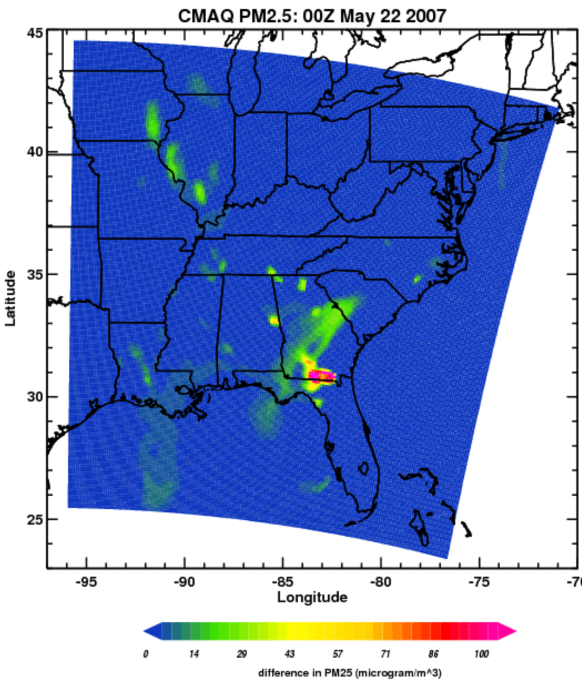


$PM_{2.5}$ forecasts adjusted to model wind errors

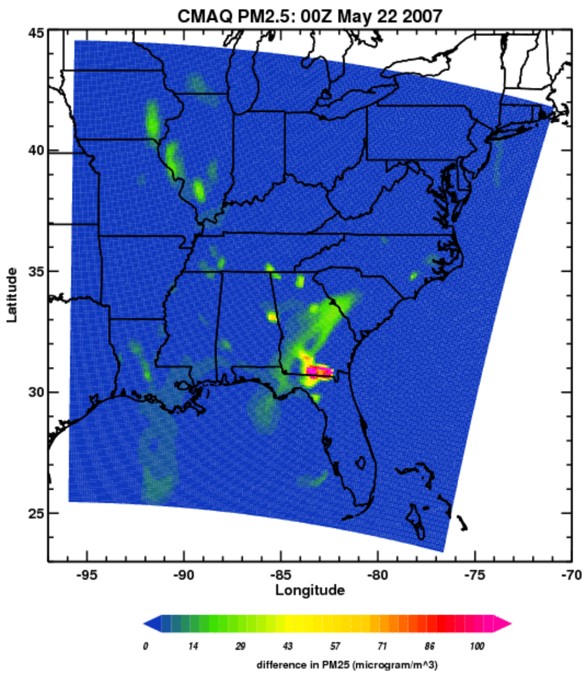
without error



with error, no AR(1)



with error, AR(1)



Summary

Q. If the model wind at 10 am is larger than the true wind, what do you expect next hour?

A. The model wind at 11 am will be more likely to be larger than the true wind.

- **Model winds are a key factor in predicting the position of fire smoke.**
- **Model wind errors are positively autocorrelated.**
- **Smoke transport error is better represented with AR(1).**
- **High $PM_{2.5}$ episodes can better be captured by wind error with AR(1).**