



# Institute for Defense Analyses

4850 Mark Center Drive • Alexandria, Virginia 22311-1882

---

**9<sup>th</sup> International Conference on Harmonisation Within  
Atmospheric Dispersion Modelling for Regulatory Purposes,  
Garmisch-Partenkirchen / Germany  
June 1 - 4, 2004**

***Comparisons of Transport and  
Dispersion Model Predictions of the  
Urban 2000 Field Experiment***

**Steve Warner  
Nathan Platt  
James F. Heagy**



# Introduction and Outline

---

- **Introduction**

- Institute for Defense Analyses (IDA) is a non-profit research and development center.
- For this task, the U.S. Defense Threat Reduction Agency sponsored us to provide independent technical analyses to support model improvement and validation efforts.
- Within the U. S. Department of Defense, recent interest in characterizing and understanding transport and dispersion in an urban environment stems from the need to reliably estimate the population effects resulting from releases of chemical or biological agents. Such estimates require knowledge of the concentrations of dispersed material as a function of time and location.

- **Outline of Presentation**

- Urban 2000 field experiment
- Urban Hazardous Prediction and Assessment Capability (Urban HPAC) Model
- Protocol and Methodologies for Comparisons
- Example Results and Insights



## Brief Description of *Urban 2000*

---

- **Series of sulfur hexafluoride ( $\text{SF}_6$ ) releases; downtown Salt Lake City, Utah; October 2000**
  - Objective: collect tracer concentrations and meteorological observations throughout an urban area
  - Relatively small releases of short duration (1 hr) from a point or short line (30 m) source (nighttime releases and winds were generally light)
- **Six intensive operating periods (IOPs), that included  $\text{SF}_6$  releases, were associated with *Urban 2000***
  - Meteorological and tracer measurements were made throughout the urban region with an outermost arc of  $\text{SF}_6$  samplers located 6 km downwind of the release.
  - 66 ground locations ( $\text{SF}_6$  sampler sites) were examined.
  - 18 independent releases (3 per IOP)



# Brief Description of Urban HPAC: Four Modes of Operation

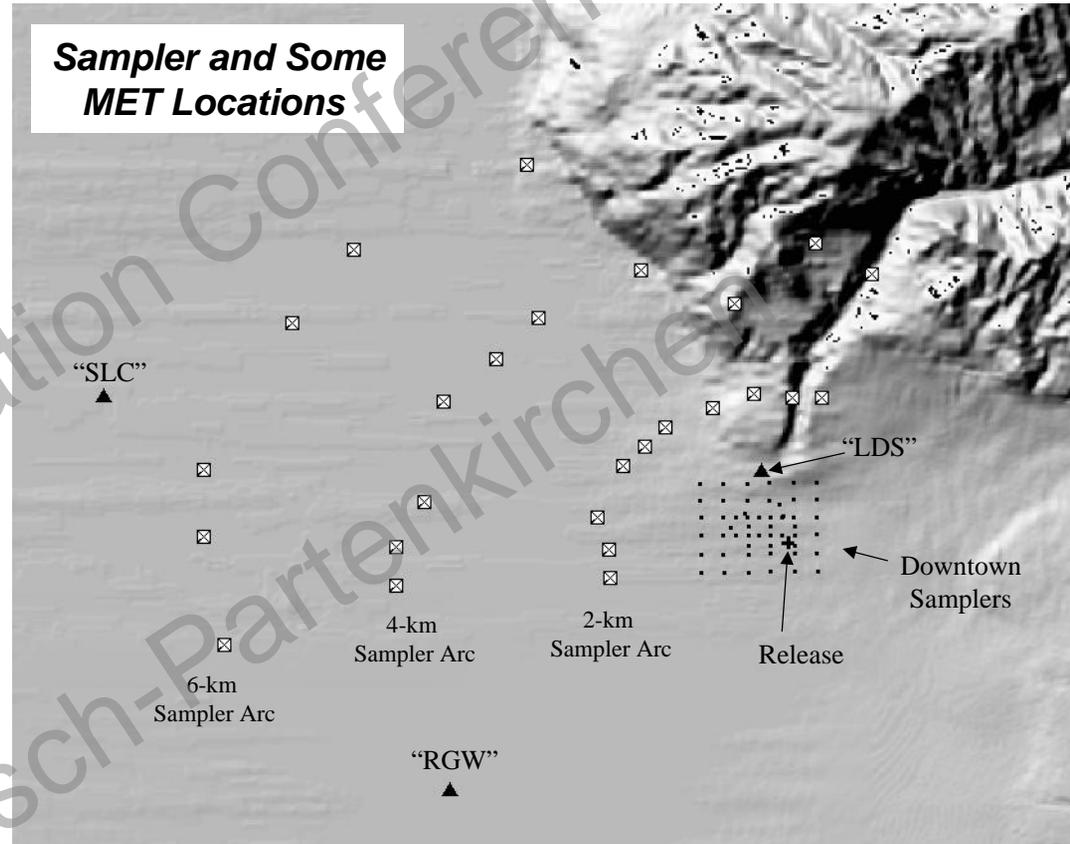
---

- HPAC uses the Second-Order Closure Integrated Puff (SCIPUFF) model and an associated mean wind field model
- **Baseline: Urban Canopy – “UC”**
  - For urban applications of HPAC, the vertical wind and turbulence profiles can be modified to account for urban effects – “urban canopy.”
- **Urban Dispersion Model – “DM”**
  - United Kingdom Defense Science and Technology Laboratory’s (DSTL) Urban Dispersion Model (UDM)
  - Empirical model based on wind tunnel studies and requires building database to support
- **Urban Windfield Module – “WM”**
  - Urban Windfield Module (UWM) predicts steady-state winds (speed and direction) inside the urban boundary layer using a canopy parameterization.
  - UWM is a computational fluid dynamics (CFD) code that is designed to provide a computationally fast solution, within a CFD framework, by considering spatially averaged obstacle effects. Therefore, the predicted winds of UWM represent spatio-temporal averages.
- **Both Urban Dispersion Model and Urban Windfield Module – “DW”**
  - Invoke both UDM and UWM – the expected optimum



# Meteorological Input Options Considered: Five Options Examined

- **Salt Lake City Airport – “SLC”**
  - Hourly surface and twice daily upper air, 10 km N-NW
- **Raging Waters Site - “RGW”**
  - Surface, sodar (20 – 300 m AGL), and radar profiler (172 – 3700 m AGL, 5 km SW)
- **Latter Day Saints’ Administration Building Top – “LDS”**
  - Wind speed and direction on top of building (124 m) within urban regime
- **All Meteorological (MET) Observations – “ALL”**
  - Including surface, building top, and sodars within and outside the urban area



- **OMEGA Forecast – “OMG”**

- 36 hour forecast (16 – 23 hours “old”), 9000 grid cells (horizontal resolution = 100 to 2-3 km and vertical resolution = 15 m (near ground) to 1 km (at top): Zafer Boybeyi – George Mason U.



# 20 Sets of Urban HPAC Predictions of the 18 Urban 2000 SF<sub>6</sub> Releases Were Created

<b>Meteorological Input Options</b>	<b>Urban HPAC Model Configuration</b>			
	<b>UC</b>	<b>DM</b>	<b>WM</b>	<b>DW</b>
<b>SLC</b>	UC_SLC	DM_SLC	WM_SLC	DW_SLC
<b>RGW</b>	UC_RGW	DM_RGW	WM_RGW	DW_RGW
<b>LDS</b>	UC_LDS	DM_LDS	WM_LDS	DW_LDS
<b>ALL</b>	UC_ALL	DM_ALL	WM_ALL	DW_ALL
<b>OMG</b>	UC_OMG	DM_OMG	WM_OMG	DW_OMG



# “Point-to-Point” Comparisons Defined

---

- **Predictions and Observations Paired in Space and Time**
  - For example,  $\approx 66$  locations  $\times$  18 releases
  - For each release, we examined 30-minute average concentrations (ppt)
    - 4 per release (4752 possible comparisons) and 2-hour dosages (ppt min).
  - Unlike analyses based on derived quantities (e.g., “plume” width, arcmax, crosswind integrated), this examination explicitly considers the size (amount), shape, and specific location (not just the downwind distance) of the cloud.

***Therefore, it is expected that the ability of the model (to include the input wind field information) to match the wind speeds and directions will be of vital importance for a point-to-point analysis.***



# Procedures and Protocol for Point-to-Point Comparisons

---

1. Add 3 ppt (background) to all predictions.
2. Observations less than 3 ppt were set to 3 ppt.
3. Observations denoted “-999” and corresponding predictions were removed.
4. 2-Hour Dosages were considered valid for comparison only when all four 30-minute average concentrations were valid (no interpolation only summation).

**94% of Concentration Pairings and  
81% of Dosage Pairings Used**

**Second protocol was also examined and documented:  
observation or prediction + 3 ppt < 14 ppt (MLOD) set to  
3 ppt. Results were found to be robust.**



# Metrics Examined for Concentration and Dosage

- **13 “Standard” Statistics Computed and Examined**
  - **Measures of Bias:** Bias, Fractional Bias (FB), and Geometric Mean (MG)
  - **Measures of Scatter:** Normalized Absolute Difference (NAD), Root Mean Square Error (RMSE), Normalized Mean Square Error (NMSE), Bounded Normalized Mean Square Error (BNMSE), and Geometric Variance (VG)
  - **Measures of Correlation:** linear Pearson correlation coefficient (R), correlation coefficient based on logarithms ( $R_{LN}$ ), and fraction of predictions within a factor of 2, 5, and 10 (FAC2, FAC5, and FAC10)
- **User-Oriented Measures of Effectiveness (MOE)**
  - Previously described at 2001 “Harmonisation Conference”
  - Threshold-based (concentration and dosage) MOEs (4 thresholds examined)
  - Summed concentration and dosage-based MOEs
- **Above statistics and MOEs were computed “globally” and as averages (based on 18 independent releases and 6 independent IOPs).**

**From the above, 114 metrics were computed for each model configuration.**



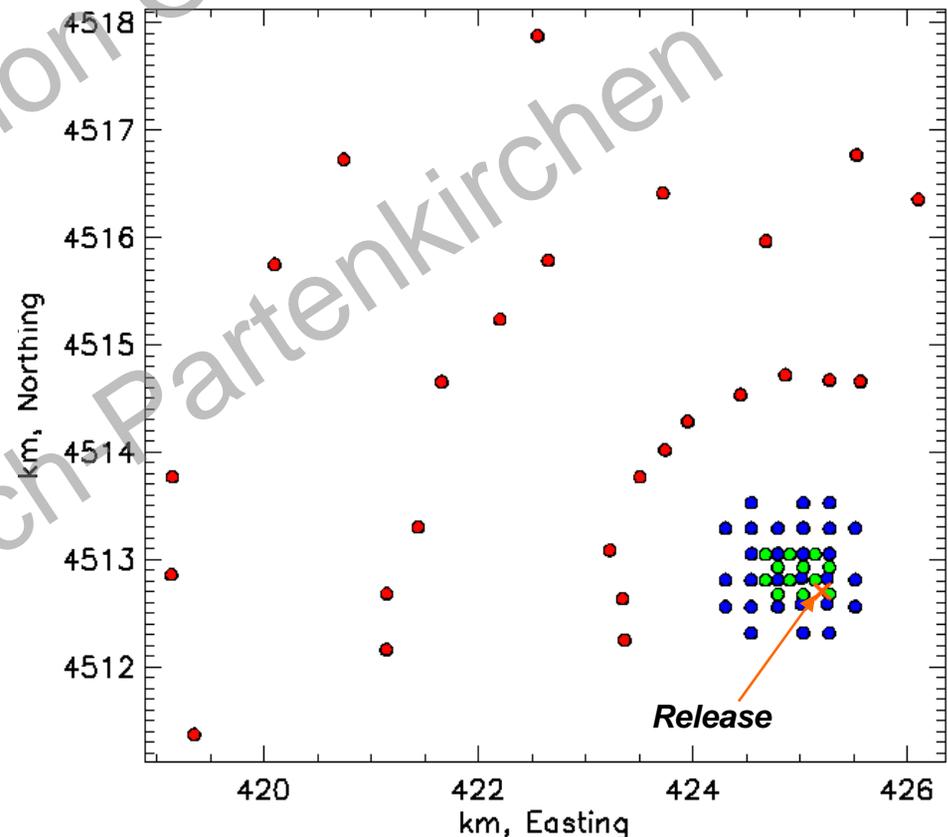
# Scope of Comparisons: Five Distance Regimes

**Urban 2000 Samplers and Downwind Distance Regimes: 12 Green Circles Correspond to the Downtown Mid-Block Sampler Locations, 28 Blue Circles Correspond to the Downtown Intersection Sampler Locations, 26 Red Circles Correspond to the Arc Sampler Locations, and the Orange "X" Marks the Release Point**

- Each metric was computed for each of the 20 model configurations.
- Also, each metric was computed for 5 downwind distance regimes.
- $114 \times 20 \times 5$

**Total metrics = 11 400**

- Approximate 0.95 confidence intervals and regions for two-dimensional metrics were appropriately computed.



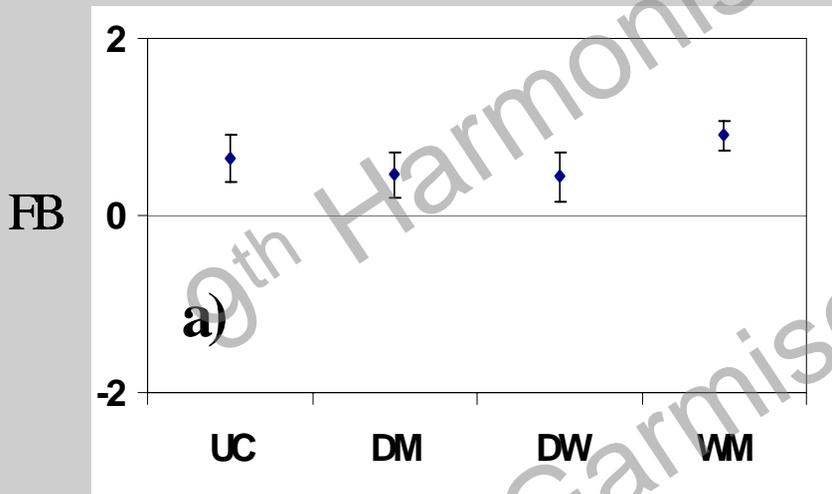


# Calculation of p-values / Hypothesis Testing

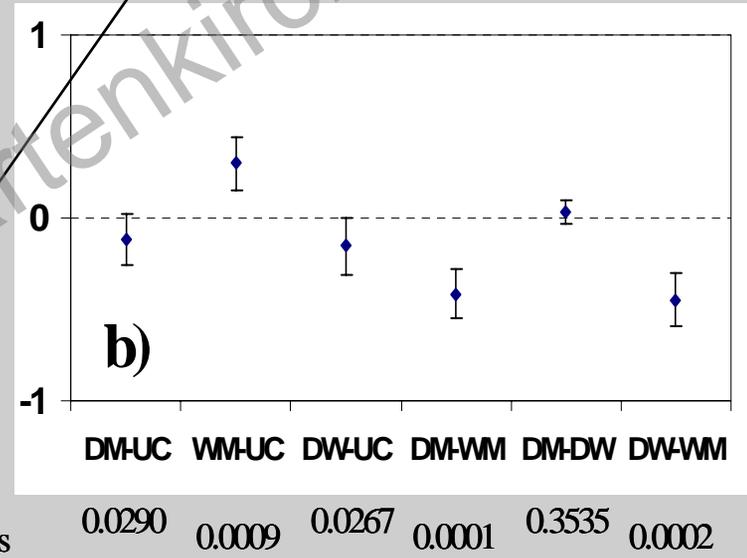
- Computation of p-values (hypothesis testing) was needed to aid examinations of this very large set (>25 000) of comparisons.
- For one-dimensional metrics, permutation test with general scores was chosen; for two-dimensional metrics, “2-D Sign test” procedures were followed. Both methods are non-parametric.

**ALL MET input option shown**

*based on 18 releases*



Mean FB Difference

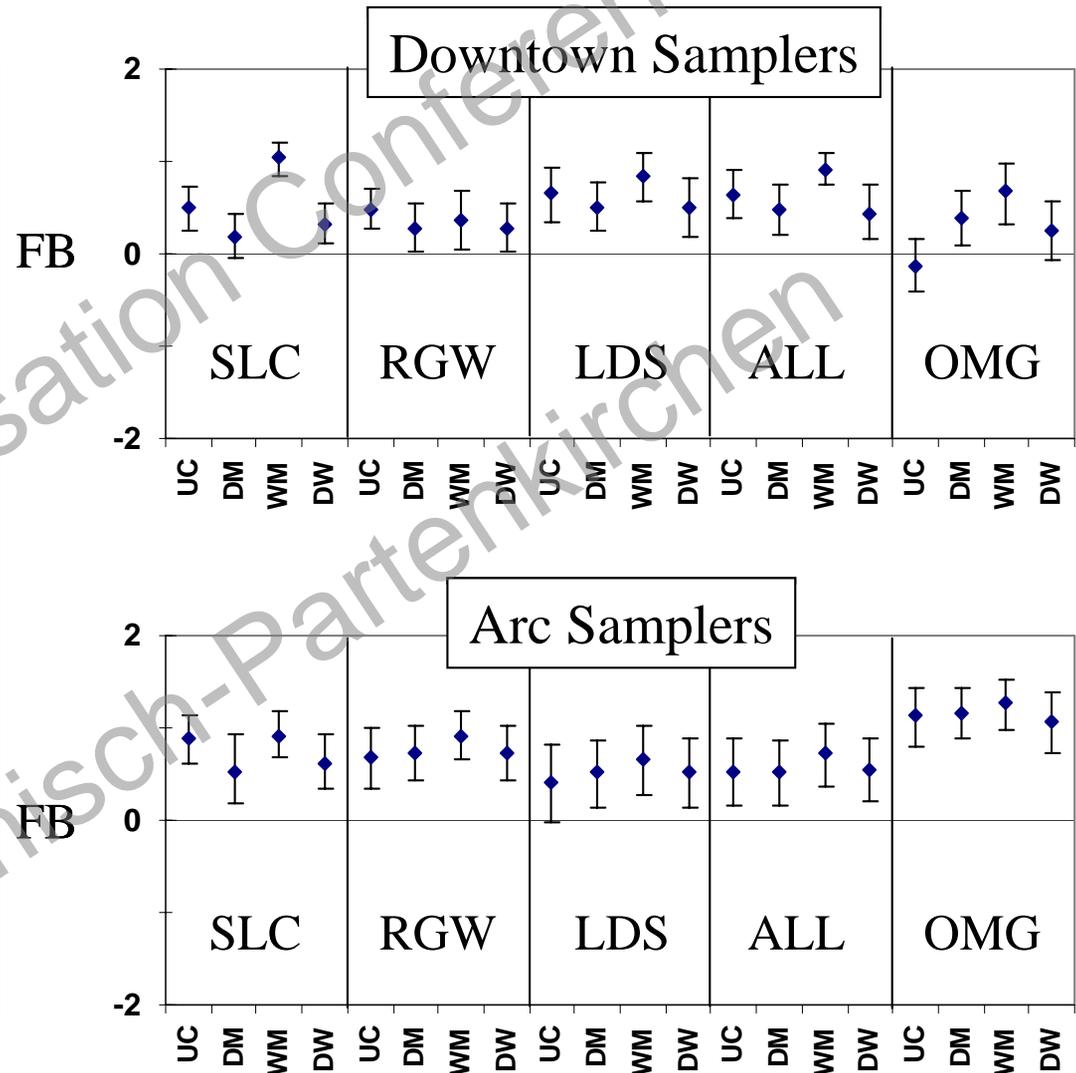


**FB, NAD, and MOE used here to support major findings.**



# In General, Urban HPAC Over-Predicted the Ground Level Concentrations and Dosages

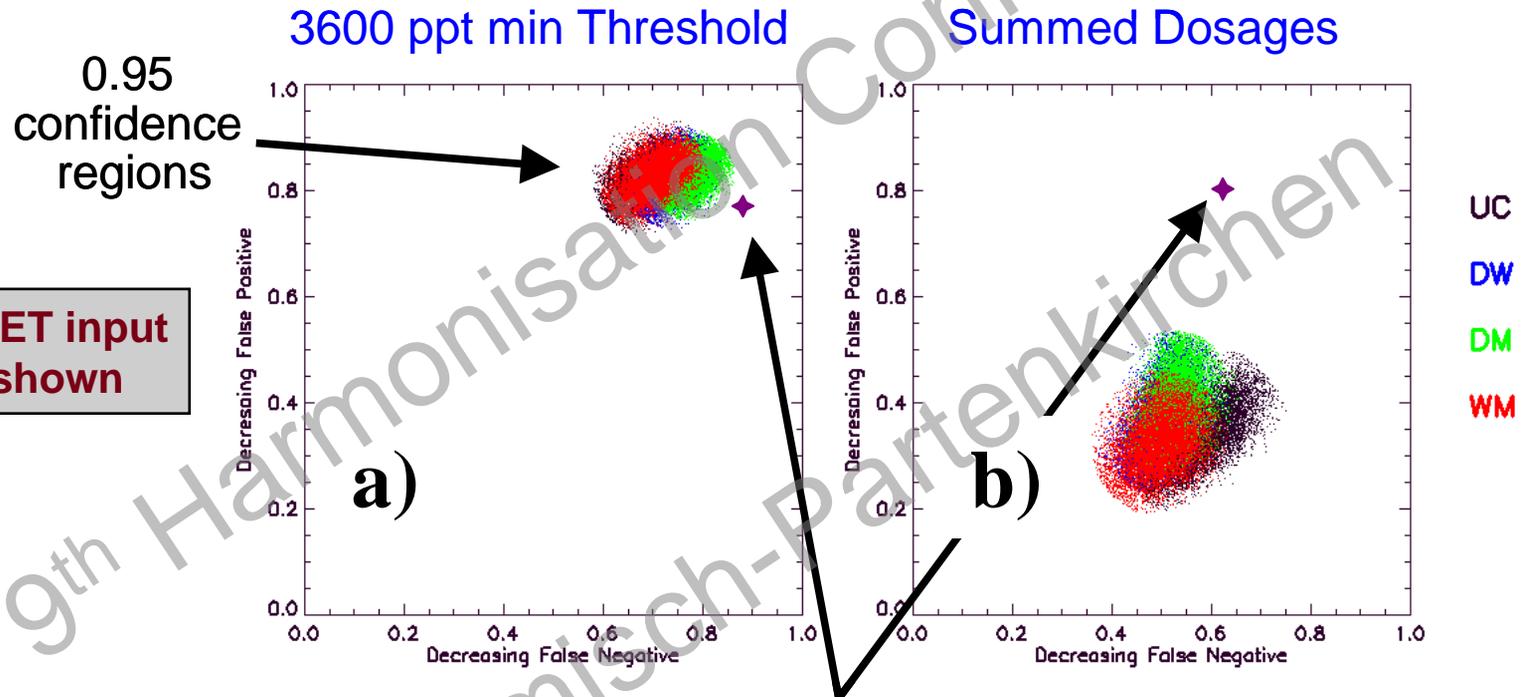
- **19 of 20 sets of Urban HPAC predictions led to an over-prediction.**
  - Exception was UC-OMG for downtown samplers
  - Average for downtown is 643 – 6670 ppt; for arcs, 39 – 264 ppt
- **Suggest fundamental underlying cause**
  - Hypotheses: too little initial lofting or too little vertical dispersion in the model predictions





# Predictions of Exceeding a Relatively Low Threshold Were Much More Accurate Than Predictions of Specific Amounts of Material

MOE = (1,1)  $\Rightarrow$  perfect.



Comparable MOE values for HPAC predictions of *Prairie Grass*

The implication is that for simply predicting “low-threshold” hazard regions, these models may be about as good in urban as in non-urban regions.



# Relative Rankings of Urban HPAC Configurations

**Example Table  
for RGW MET  
Input Option**

<i>p-values</i>	MOE	MOE	NAD	NAD
Comparison Being Tested	30-minute concentration	2-hour dosage	30-minute concentration	2-hour dosage
DM-DW	0.0216	0.0007	0.1789	0.2099
DM-WM	≤ 0.0001	0.0001	0.0053	0.0217
DM-UC	0.0050	0.0050	0.0191	0.0416
DW-WM	0.0771	0.2770	0.0104	0.0747
DW-UC	0.0216	0.2770	0.0808	0.1910
UC-WM	0.2270	0.5292	0.7604	0.5945

**Rankings as a  
Function of MET  
Input Option**

<u>SLC</u>	<u>RGW</u>	<u>LDS</u>	<u>ALL</u>	<u>OMG</u>
DM DW WM UC	DM DW WM UC	DM DW WM UC	DW DM WM UC	UC DM DW WM

- 1) In general, inclusion of UDM (“DM”) led to improvements.
- 2) Addition of UWM (“WM”) did not lead to significant improvement.
- 3) UC\_OMG is an exception (compensating errors, counteracting effects?)



# Relative Rankings of MET Input Options

## Example Table for UC

- 1) MET input options that included observations within the urban canopy (LDS and ALL) led to poorer predictive performance.
- 2) RGW led to the best performance.
- 3) How should one include (if at all), “in-canopy” observations?

<i>p-values</i>	MOE	MOE	NAD	NAD
Comparison Being Tested	30-minute concentration	2-hour dosage	30-minute concentration	2-hour dosage
OMG-RGW	0.0771	0.0050	0.0296	0.0550
OMG-SLC	<0.0001	0.0001	0.0003	0.0004
OMG-LDS	<0.0001	0.0050	< 0.0001	0.0003
OMG-ALL	0.0007	0.0216	0.0005	0.0013
RGW-SLC	0.0007	0.0001	0.0151	0.0150
RGW-LDS	0.0007	0.0771	0.0096	0.0257
RGW-ALL	0.0216	0.0216	0.0369	0.1181
SLC-LDS	0.5292	0.2270	0.6054	0.2399
ALL-SLC	0.2270	0.8972	0.7981	0.3924
ALL-LDS	0.0771	0.5292	0.4843	0.7814

## Rankings as a Function of Urban HPAC Configuration

UC	DM	WM	DW
OMG	SLC RGW OMG	RGW	SLC
RGW	LDS ALL	LDS SLC	OMG
SLC LDS ALL			LDS



# Conclusions

---

- An inclusive, paired in space and time – point-to-point – comparative protocol and quantitative hypothesis testing were used to assess the predictions of the *Urban 2000* field experiment by 20 Urban HPAC model configurations.
- **General Findings**
  - Over-predictions (with one exception)
  - Performance of predictions of whether or not a relatively low threshold was exceeded (e.g., hazard regions) was substantially improved relative to predictions associated with absolute amounts of material.
- **Relative Findings**
  - For these releases and predictions, addition of UDM improved performance (addition of UWM did not).
  - The two weather inputs that included meteorological information near the urban source resulted in relatively poor performance that was probably caused by including these more variable in-canopy observations as model input.

**In press, *J. Appl. Met.***

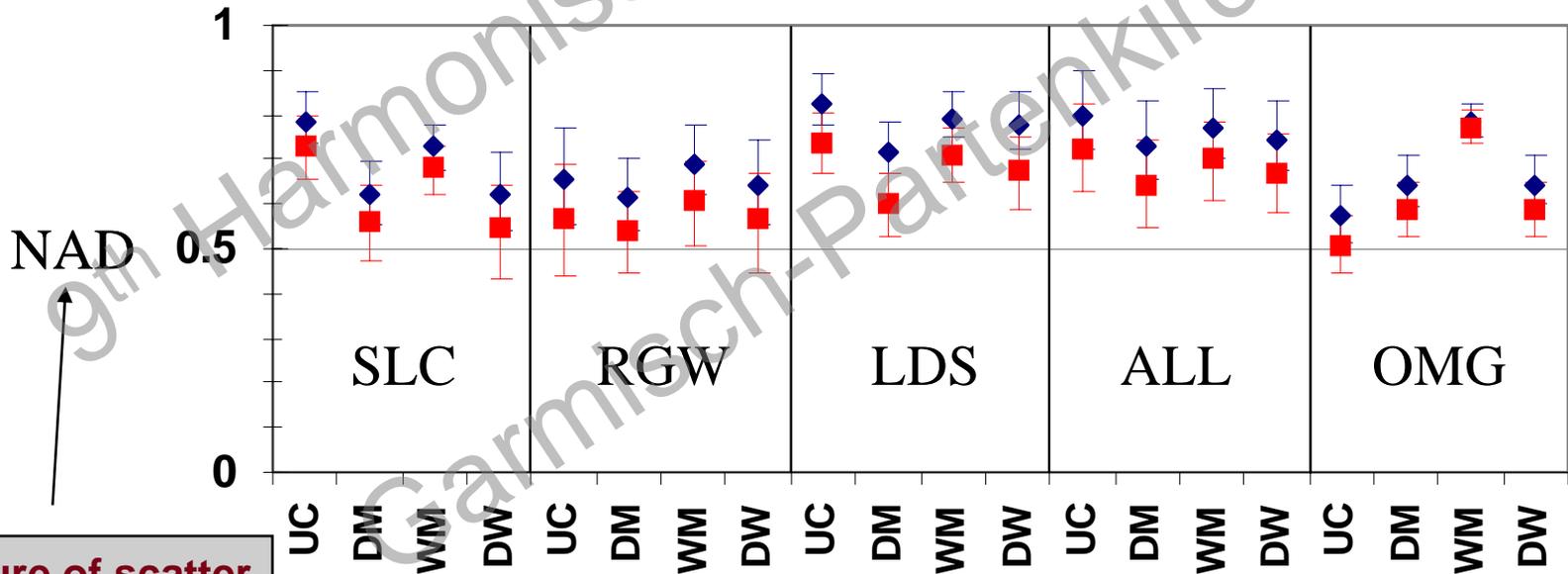


# Two Hour Dosages Were Better Predicted Than 30-Minute Average Concentrations

## Normalized Absolute Difference (NAD)

The removal (“smoothing”) of temporal component of point-to-point comparison leads to expected predictive improvements.

$$NAD = \frac{\sum_{i=1}^n |C_o^{(i)} - C_p^{(i)}|}{\sum_{i=1}^n (C_o^{(i)} + C_p^{(i)})}$$



Measure of scatter  
NAD = 0 is perfect

◆ 30-Minute Average Concentration ■ 2-Hour Dosage