Institute for Defense Analyses
4850 Mark Center Drive • Alexandria, Virginia 22311-1882

13th International Conference on Harmonisation within Atmospheric Modelling for Regulatory Purposes
Paris, France
1-4 June 2010

Comparative Investigation of Source Term Estimation Algorithms using FUSION Field Trial 2007 Data – Linear Regression Analysis

Nathan Platt
Dennis DeRiggi

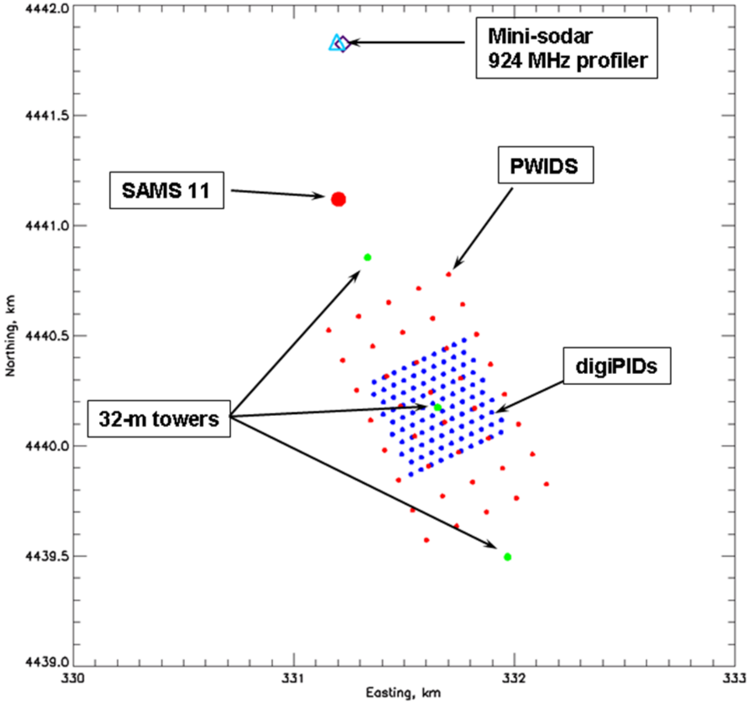
June 4, 2010

Outline

- **FUSION Field Trial 2007**
- **Phase I of Source Term Estimation Algorithm Comparison Exercise**
  - Phase I Data Statistics
  - Demonstration of individual case
  - Sets of Predictions Received
- **Inter-comparisons of algorithms**
  - Metrics used in the analysis
  - Using regression analysis to ascertain trends among algorithms
- **Summary**
- **Motivation for Phase II**

FUSION Field Trial 2007 (FFT 07)

- **FUsing Sensor Information from Observing Networks (FUSION)**
- **Conducted at U.S. Army Dugway Proving Ground in September 2007**
- **Objective: to provide a comprehensive tracer dispersion and meteorological dataset suitable for testing current and future chemical and biological (CB) sensor data fusion (SDF) algorithms**
- **Concept: to collect data from an abundance of research-grade tracer, sensor, and meteorological instruments, rather than employing an "optimal" placement strategy**
- **International Participation - Defence Research and Development Canada (DRDC), the UK Defence Science and Technology Laboratory (Dstl), and the Australian Defence Science and Technology Organisation (DSTO)**



Phase I of Source Term Estimation Algorithm Comparison Exercise

Why do we need exercise for STE algorithms:

- **To best allow for scientific insights from comparative analyses**
- **To provide for credible and fair comparisons among algorithms (in a *reasonably* realistic setting)**
  - To avoid perceived intentional, or more likely unintentional, model parameter tweaking to fit the unique data and observations of FFT 07
  - To give the most credible assessment of the state-of-the-art
- **To best allow information to be re-used for independent validation in the future (with newer algorithms)**
- **To clarify maturity of emerging STE algorithms for possible inclusion into JEM**

Eight STE Algorithm Developers Decided to Participate in This Exercise and Provided 14 sets of Predictions

Phase I Data Release Composition

Phase I Release Case Composition					
Condition	**All Trials**	**Single**	**Double**	**Triple**	**Quad**
none	104	40	40	16	8
Puff	52	20	20	8	4
Cont	52	20	20	8	4
Daytime	52	20	20	8	4
Nighttime	52	20	20	8	4
Daytime/Puff	26	10	10	4	2
Daytime/Cont	26	10	10	4	2
Nighttime/Puff	26	10	10	4	2
Nighttime/Cont	26	10	10	4	2

Phase I Dataset Consisting of 104 Cases was Released to Exercise Participants in September 2008

STE Prediction Sets

Composition of the Prediction Sets Recieved									
Organization	Total	Cont	Puff	Daytime	Nighttime	Single	Double	Triple	Quad
Aerodyne	104	52	52	52	52	40	40	16	8
Boise-State	33	14	19	21	12	13	13	4	3
Buffalo / GA	104	52	52	52	52	40	40	16	8
Buffalo / SA	70	34	36	34	36	26	26	12	6
DSTL	35	5	30	20	15	12	14	7	2
ENSCO / Set 1	102	51	51	50	52	39	39	16	8
ENSCO / Set 2	104	52	52	52	52	40	40	16	8
ENSCO / Set 3	42	24	18	19	23	13	15	10	4
NCAR / Variational	38	3	35	20	18	16	14	4	4
NCAR / Phase I	38	3	35	20	18	16	14	4	4
Sage-Mgt	104	52	52	52	52	40	40	16	8
PSU / Gaussian	50	26	24	25	25	18	20	8	4
PSU / SCIPUFF	50	26	24	25	25	18	20	8	4
PSU / MEFA	35	19	16	17	18	13	16	5	1

Algorithm Capabilities		
Organization	Number of Sources	Type
Aerodyne	Multi	Cont/Puff
Boise-State	Single	Cont/Puff
Buffalo / GA	Multi	Cont/Puff
Buffalo / SA	Mostly Single	Cont/Puff
DSTL	Single	Puff
ENSCO / Set 1	Multi	Cont/Puff
ENSCO / Set 2	Single	Cont
ENSCO / Set 3	Single	Cont
NCAR / Variational	Single	Puff
NCAR / Phase I	Single	Puff
Sage-Mgt	Single	Cont/Puff
PSU / Gaussian	Single	Cont/Puff
PSU / SCIPUFF	Single	Cont/Puff
PSU / MEFA	Multi	Cont/Puff

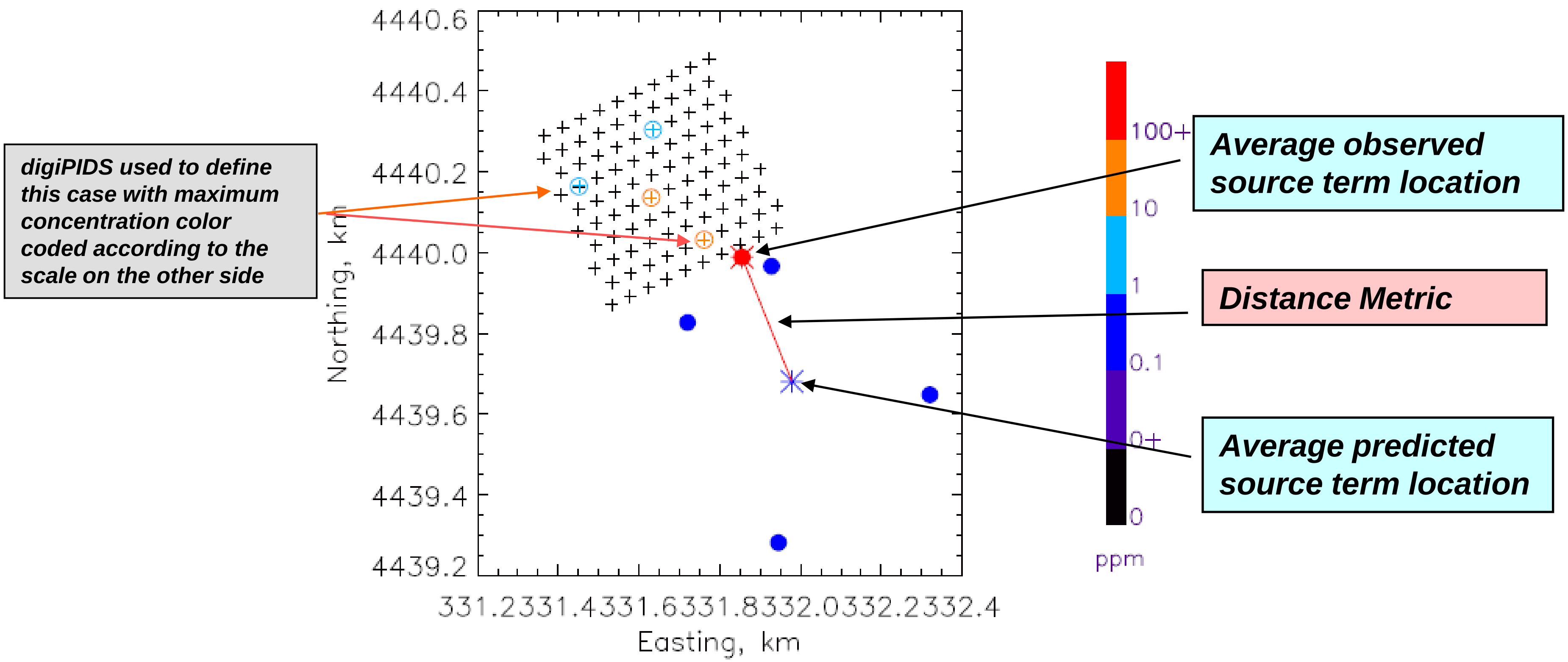
Notes

- ***Only cases when location is predicted are used in this table***
- ***Boise-State provided 53 predictions for cases 1-53 with 33 cases converging to a location estimate***
- ***PSU provided predictions for sixteen sensors cases only***

Blue ≥ 50% of cases predicted
Red – all cases predicted

Metrics Used in the Analysis

Sample Plot in *Location_Plots_Buffalo_GA.pdf*



Case Num: 4
Physical trial: xx
Trial type: PUFF
Puff Num: All Puffs
Prediction Identifier: Buffalo GA
Actual Sources: 1; Total Mass: 1.33200 kg
Predicted Sources: 4; Total Mass: 16.3704 kg

Total Mass Metric

Average Distance in km: 0.33145780

STE Algorithm Inter-Comparison

- Regression Analysis to Ascertain Trends Among Different Sets of Predictions is presented here
- Gross Algorithm Performance Trends using “Mean Missed Distance” and “Total Predicted/Actual Mass” Ratio Metrics are not presented here

Brief Description of Regression Analysis Performed

- **Two techniques are presented:**
  - Stepwise Regression
  - Backwards Regression

- **Stepwise**
  - Stepwise regression searches among the independent variables to determine which is most correlated with the dependent variable. That variable becomes the 1^{st} to enter the regression.
  - The next entry is the variable whose partial correlation (that is, after controlling for the effect of the 1^{st} independent variable) is the highest.
  - An F-test is now performed to determine what the effect would be of adding the 1^{st} independent variable to the regression if the 2^{nd} independent variable had entered first. If significant, the 1^{st} variable is retained. Otherwise it is removed.
  - The process now continues by examining the partial correlations of the remaining variables.

- **Backward**
  - Backward regression (backward elimination) enters all independent variables into the regression.
  - An F-test is performed for each variable as though it were the last to enter the regression; if not significant at some prescribed level, that variable is removed. Otherwise it is retained.

Independent Regression Variables

Case	Diurnal	MET Num	Sources	Sensors	Puff/Real
1	Night	Close-In	1	4	-1
2	Night	Close-In	2	4	1
3	Night	Close-In	1	4	-1
4	Night	Close-In	1	4	1
5	Night	Close-In	1	16	1
6	Night	Close-In	4	4	-1
7	Night	Close-In	2	4	1
8	Night	Close-In	4	16	-1
9	Night	Close-In	1	16	1
10	Night	Close-In	2	16	-1
11	Night	Operational	2	16	0
12	Night	Close-In	3	16	0
13	Night	Close-In	3	16	0
14	Night	Close-In	1	4	-1
15	Night	Operational	2	16	-1
16	Day	Operational	1	16	0
17	Night	Close-In	2	16	0
18	Night	Close-In	2	4	-1
19	Day	Close-In	2	16	0
20	Day	Close-In	3	4	1

$$\text{Puff Real} = \begin{cases} -1 & \text{if Continuous Release} \\ 0 & \text{if single realization of a Puff release} \\ 1 & \text{if multiple realizations of a Puff release} \end{cases}$$

Sample Dependent Regression Variables

Case	Mean (Dist)	Mass Ratio
1	0.18098393	1.276159841
2	0.51655648	10.4932407
3	0.17311404	0.206608389
4	0.13475478	4.307807958
5	0.025230382	1.108092215
6	0.10410637	0.235141559
7	0.095627225	11.41600705
8	0.10891281	0.170577897
9	0.061687421	0.710246583
10	0.044667524	0.883691805
11	0.057406344	0.215429999
12	0.036641343	0.461666624
13	0.11905685	2.403726708
14	0.063702853	0.264423135
15	0.034814414	0.200444062
16	0.060312748	1.16762176
17	0.06263416	1.096964541
18	0.13387494	4.959205386
19	0.02892583	1.409658618
20	0.055047161	4.350513428

Sample Summary Table of Regression Analysis

"Significant Variables" Table for Backward Regression

model	dependent	R2	significant factor	significant factor	significant factor
ENSCO 3	Mass Ratio	0.379	Puff Real (0.51, 2.49. 0)	Sources (-0.447, -1.9, 0.001)	
Buffalo SA	Mass Ratio	0.273	Sources (-0.348, -0.723, 0.002)	Met Num (0.235, 0.632, 0.031)	Diurnal (0.231, 0.508, 0.029)
DSTL	Mass Ratio	0.254	Puff Real (-0.567, -287.1, 0.001)	Sources (-0.376, -75.9, 0.026)	
ENSCO 2	Mass Ratio	0.221	Puff Real (0.37, 1.3, 0.0)	Sources (-0.32, -0.93,0)	Sensors (0.17, 0.074, 0.06)
PSA Gaussian	Mass Ratio	0.209	Puff Real (0.46, 0.059, 0.01)	SourceS (-0.407, -0.037, 0.02)	
PSU SCIPUFF	Mass Ratio	0.203	Sources (-0.5, -0.011, 0.035)		
Buffalo GA	Mass Ratio	0.172	Sources (-0.365, -2.376, 0)	Puff Real (0.183, 1.417, 0.044)	Diurnal (0.177, 1.224, 0.051)
ENSCO 1	Mass Ratio	0.15	Puff Real (0.398, 14.64, 0)		
Aerodyne	Mass Ratio	0.096	Puff Real (0.262, 0.852, 0.006)	Sensors (-0.212, -0.089, 0.026)	
NCAR Phase I	Mass Ratio	0	constant		
NCAR Variation	Mass Ratio	0			
SAGE Mgt August	Mass Ratio	0			
Boise State	Mass Ratio	-1.00E-06	NO DATA		
PSU MEFA	Mass Ratio	-1.00E-06	NO DATA		
model	dependent	R2	significant factor	significant factor	significant factor
DSTL	Mean	0.67	Puff Real (-0.725, -1.105, 0)	Sources (0.212,0.129, 0.056)	
NCAR Phase I	Mean	0.266	Sources (0.534, 0.09, 0.001)		
NCAR Variation	Mean	0.204	Sources (0.475, 0.09, 0.003)		
ENSCO 3	Mean	0.148	Sources (-0.366, -0.031, 0.015)	Sensors (0.258, 0.003, 0.08)	
PSA Gaussian	Mean	0.102	Sources(0.306, 0.055, 0.029)	Puff Real (-0.254, -0.057, 0.069)	
SAGE Mgt August	Mean	0.083	Sources (0.303, 0.204, 0.002)		
ENSCO 1	Mean	0.043	Met Num (0.228, 0.009, 0.021)		
ENSCO 2	Mean	0.04	Sensors (-0.173, -0.002, 0.076)	Met Num (0.169, 0.017, 0.083)	
Aerodyne	Mean	0.033	Sensors (-0.206, -0.003, 0.036)		
Boise State	Mean	0	constant		
Buffalo GA	Mean	0	constant		
Buffalo SA	Mean	0			
PSU MEFA	Mean	0	constant		
PSU SCIPUFF	Mean	0	constant		

Regression Analysis Results
Average Miss Distance

- **With respect to predicting average miss distance, regression analysis indicates**
  - "Day/Night" is not a significant variable for both backward and stepwise regressions
    - » Some confirmation of this for MET option could be seen in Excel chart distributed in the "Developer Feedback Package"
  - "Close-In/Operational MET" is not a significant variable for both backward and stepwise regressions for almost all algorithms
    - » Exception is ENSCO 1 and 2
    - » Some confirmation of this for MET option could be seen in Excel chart distributed in the "Developer Feedback Package"
  - "Number of sources" is a significant predictor of algorithm performance for six algorithms
    - » Six algorithms called by stepwise regression and four algorithms are called by backward regression
      - Although only two have adjusted R^2 greater than 0.2
  - "4 vs.16 Sensors" is a significant predictor of algorithm performance for only three algorithms indicating that most algorithms are not benefiting from having larger number of sensors
    - » None have R^2 greater than 0.2
    - » Some confirmation of this is seen in the Excel charts provided in the "Developer Feedback Package"
  - "Puff Real" is a significant variable for two algorithms using backward regression and one algorithm using stepwise regression
    - » Although only one algorithm have R^2 greater than 0.2

Regression Analysis Results
Total Predicted Mass

- **With respect to mass ratio variable, regression analysis indicates**
  - "Day/Night ", "Close-In/Operational MET", "4 vs. 16 Sensors" are not significant variables for most algorithms for both backward and stepwise regressions
    - » "Buffalo SA" calls "Close-in/Operational MET" for both regressions
    - » ENSCO 2 and Aerodyne calls "4 vs. 16 Sensors" for backward regression and Aerodyne calls "4 vs. 16 Sensors" for stepwise regression
    - » "Buffalo SA" and "Buffalo GA" calls "Day/Night" for backwards regression and "Buffalo SA" calls "Day/Night" for stepwise regression
  - "Number of Sources" is a significant variable for seven algorithms
    - » Six algorithms are called by stepwise regression and seven algorithms are called by backward regression
      - Five algorithms have adjusted R^2 greater than 0.2
  - "Puff Real" is a significant variable for seven algorithms
    - » Five algorithms are called by stepwise regression and seven algorithms are called by backward regression
      - Four algorithms have adjusted R^2 greater than 0.2

Regression analysis results should serve as a guide on further investigation of which algorithm/variable combinations are important. For instance, the regression analysis does not tell if algorithm performed as expected with respect to a given variable (e.g. averaged missed distance decreased when 16 sensors are used instead of 4 sensors)

Summary

- **Phase I of STE algorithms exercise involving predictions from eight organizations and 14 sets of "final" predictions was closed on Aug 31, 2009**
- **Developer Feedback Package was distributed to exercise participants in early September, 2009**
  - We hope that individual developers will find information provided in this feedback package useful for them to
    - » Help analyze their algorithm performance and find areas for improvement
    - » Help publish their results
- **Independent variables that *are not* significant indicators of STE algorithm performance include**
  - Atmospheric stability
  - Quality of meteorological input
    - » High frequency MET in the middle of the grid versus relatively course MET some downwind distance
  - Number of simulated sensors available to STE algorithms (e.g. 4 vs. 16)
  - Most likely explanations are
    - » Relatively small spatial scale of digiPID grid (450 by 450 meters)
    - » Proximity of release locations to each other and leading edge of the sensor grid

Phase II is Planned for FY10

- **"Reasonably" paced second phase of the STE algorithm evaluations will facilitate further development of algorithms**
  - To potentially include adding new features, fixing bugs, continuing to learn details about expected data that will be available operationally
  - It will help algorithm developers to continue their focus on making improvements to these algorithms

- **Continues to help guide algorithm developers to consider relatively realistic situations**
  - e.g., artificial limits on search box
  - e.g., using large number of sensors on 450 meters by 450 meters grid

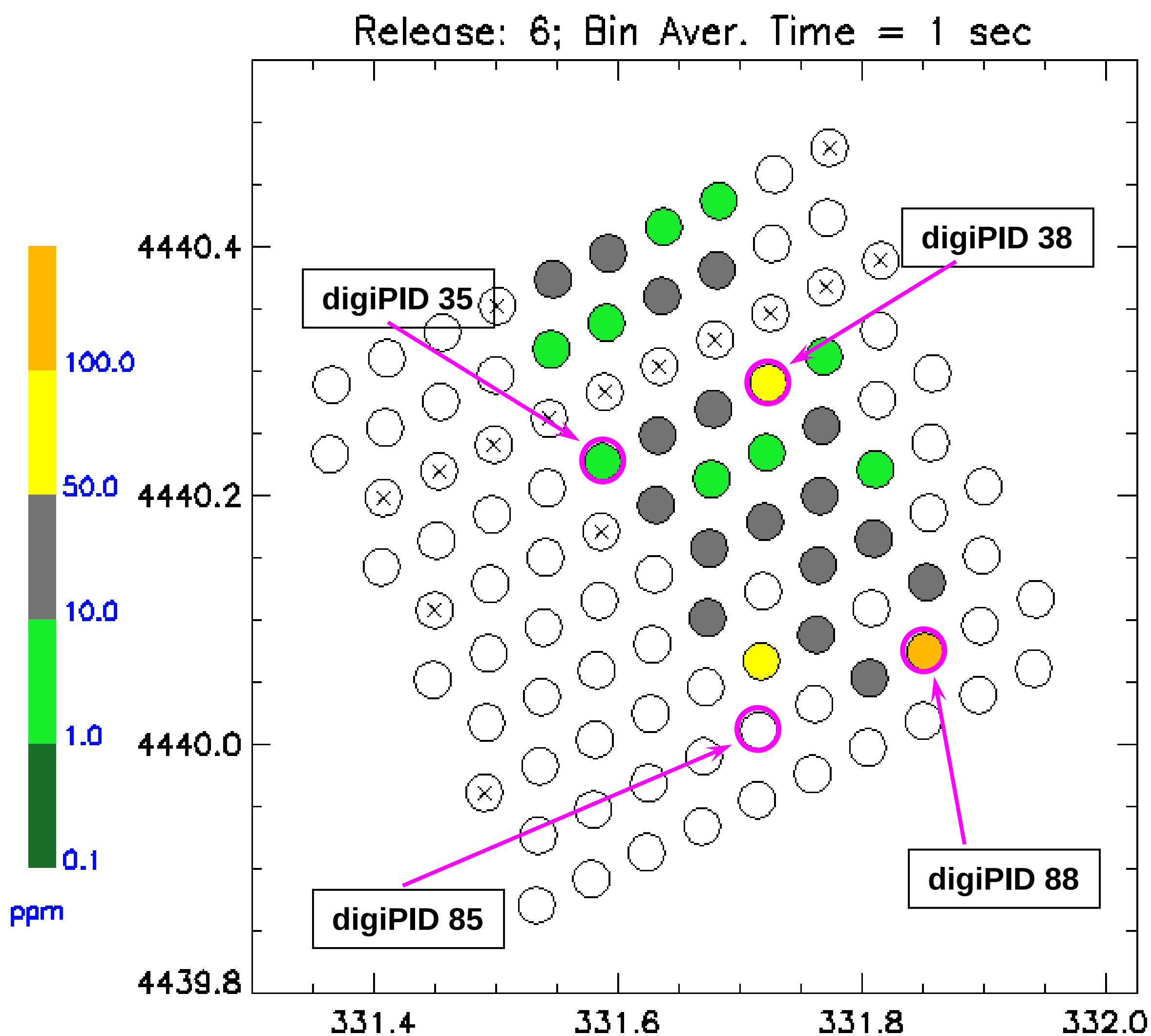
- **Broaden the scope of algorithm capabilities to better match data expected from actual chemical sensors**
  - Consider "Bar-Sensors"
  - Use VTHREAT simulation environment to
    - » Expand FFT 07 limited field trial data to "new" release locations, wind-directions, and eventually to larger "sensor placement area"
      - All FFT 07 trials were recently released

Research & Developments should play a role in informing future acquisition decisions. This work could have significant impacts in defining requirements as opposed to only satisfying requirements

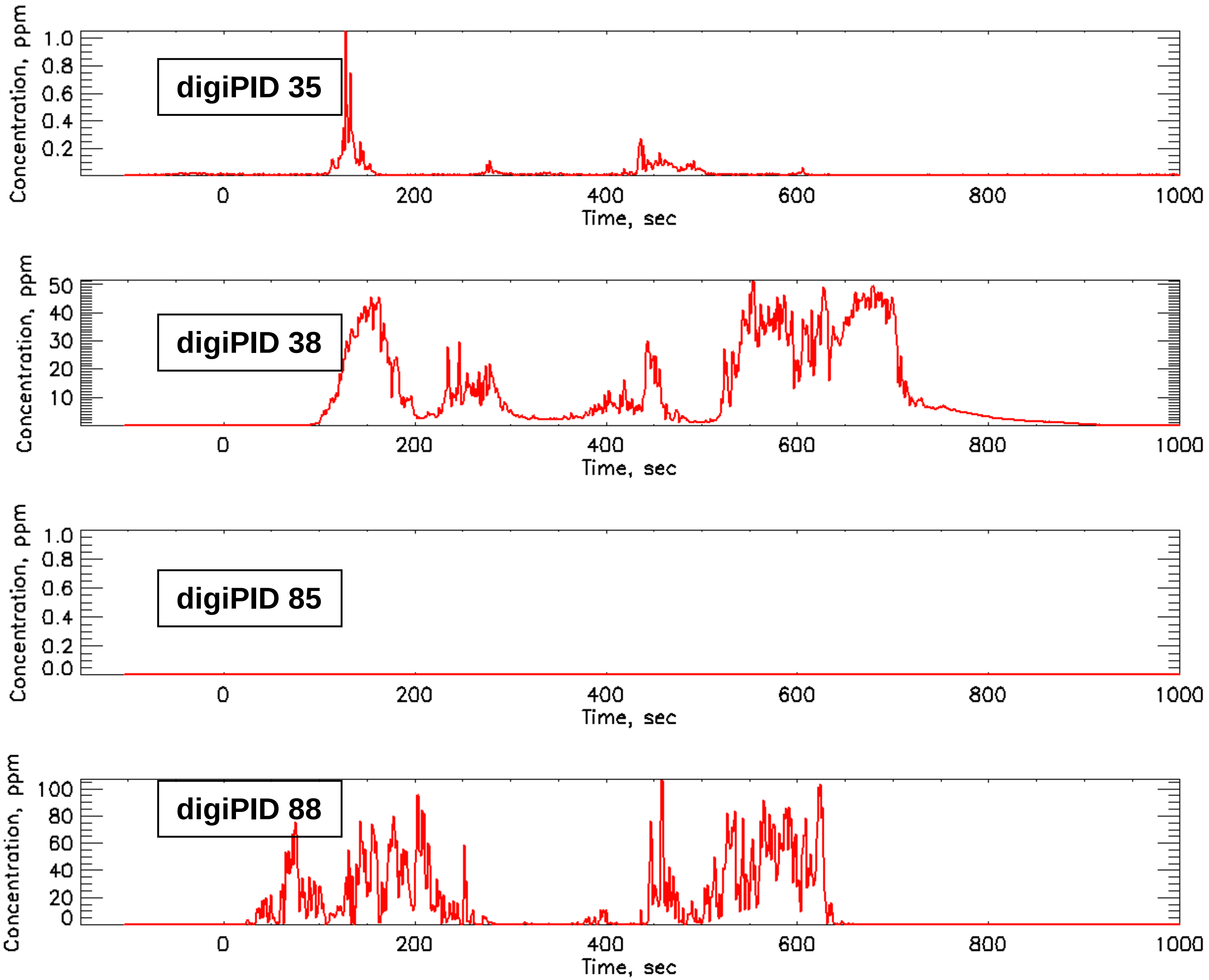
Backups

Creation of Phase I Cases

Selection of Sensors



Creation of Phase I Cases
Simulated Chemical Sensor Output

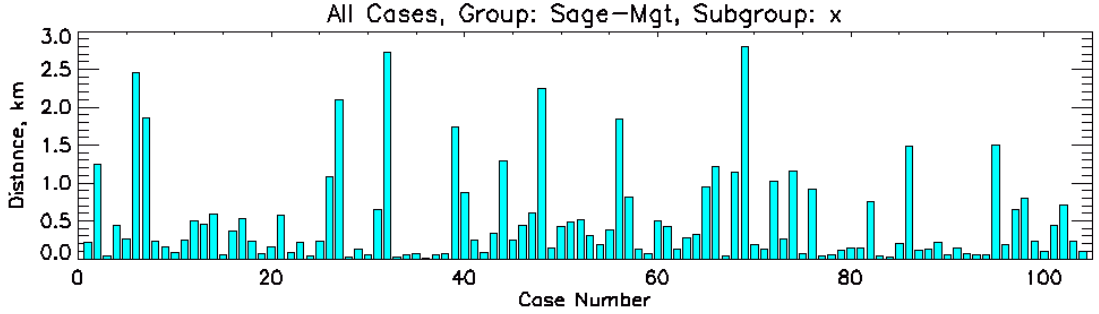


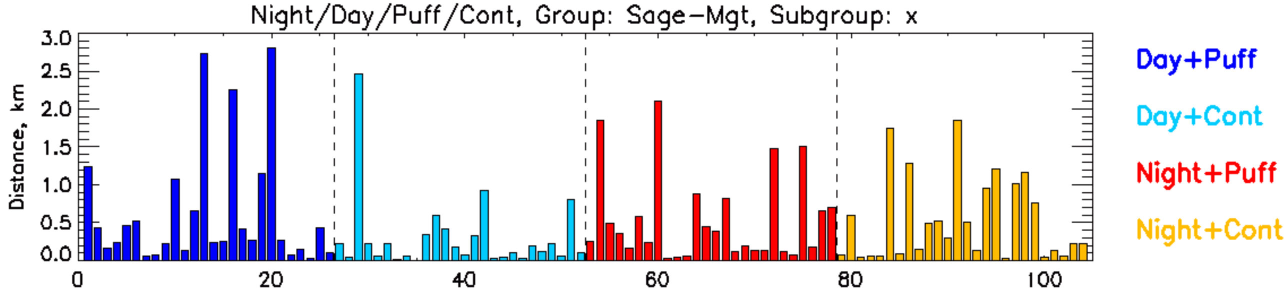
Summary Table of Regression Analysis

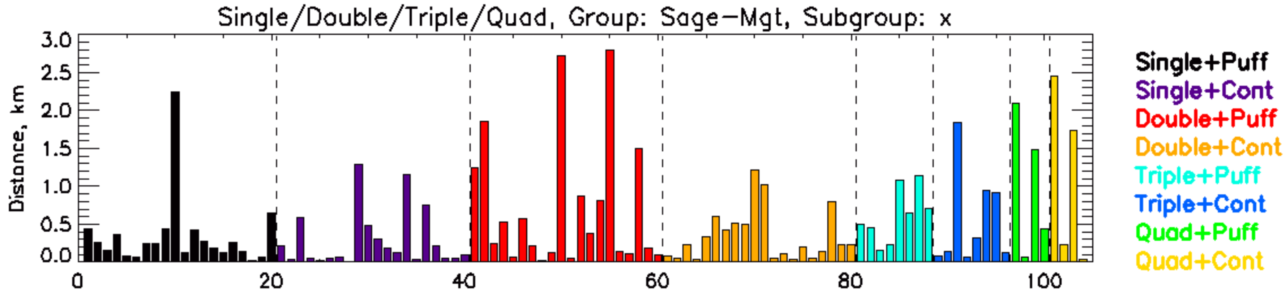
"Significant Variables" Table for Stepwise Regression

model	dependent	R2	significant factor	significant factor	significant factor
ENSCO 3	Mass Ratio	0.379	Puff Real (0.51, 2.49. 0)	Sources (-0.447, -1.9, 0.001)	
Buffalo SA	Mass Ratio	0.273	Sources (-0.348, -0.723, 0.002)	Met Num (0.235, 0.632, 0.031)	Diurnal (0.231, 0.508, 0.029)
DSTL	Mass Ratio	0.254	Puff Real (-0.567, -287.1, 0.001)	Sources (-0.376, -75.9, 0.026)	
PSU SCIPUFF	Mass Ratio	0.203	Sources (-0.5, -0.011, 0.035)		
ENSCO 2	Mass Ratio	0.201	Puff Real (0.37, 1.3, 0)	Sources (-0.32, -0.93, 0)	
ENSCO 1	Mass Ratio	0.15	Puff Real (0.398, 14.64, 0)		
Buffalo GA	Mass Ratio	0.125	Sources (-0.365, -2.376, 0)		
Aerodyne	Mass Ratio	0.096	Puff Real (0.262, 0.852, 0.006)	Sensors (-0.212, -0.089, 0.026)	
NCAR Phase I	Mass Ratio	0			
NCAR Variation	Mass Ratio	0			
PSU Gaussian	Mass Ratio	0			
SAGE Mgt August	Mass Ratio	0			
Boise State	Mass Ratio	-1	NO DATA		
PSU MEFA	Mass Ratio	-1	NO DATA		
model	dependent	R2	significant factor	significant factor	significant factor
DSTL	Mean	0.641	Puff Real (-0.807, -1.23, 0)		
NCAR Phase I	Mean	0.266	Sources (0.534, 0.09, 0.001)		
NCAR Variation	Mean	0.204	Sources (0.475, 0.09, 0.003)		
ENSCO 3	Mean	0.101	Sources (-0.35, -0.03, 0.023)		
SAGE Mgt August	Mean	0.083	Sources (0.303, 0.204, 0.002)		
ENSCO 1	Mean	0.043	Met Num (0.228, 0.009, 0.021)		
Aerodyne	Mean	0.033	Sensors (-0.206, -0.003, 0.036)		
Boise State	Mean	0			
Buffalo GA	Mean	0			
Buffalo SA	Mean	0			
ENSCO 2	Mean	0			
PSU Gaussian	Mean	0			
PSU MEFA	Mean	0			
PSU SCIPUFF	Mean	0			

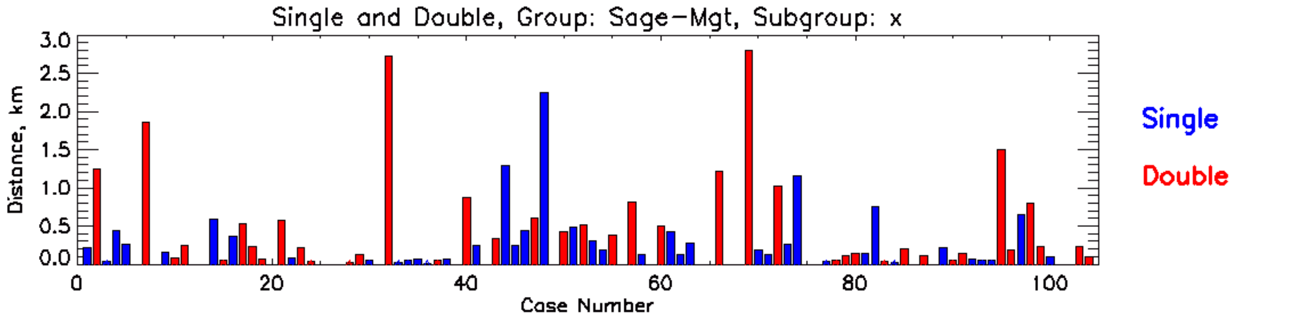
Typical “Distance Charts”
Sage-Mgt Predictions (Linear), All Cases

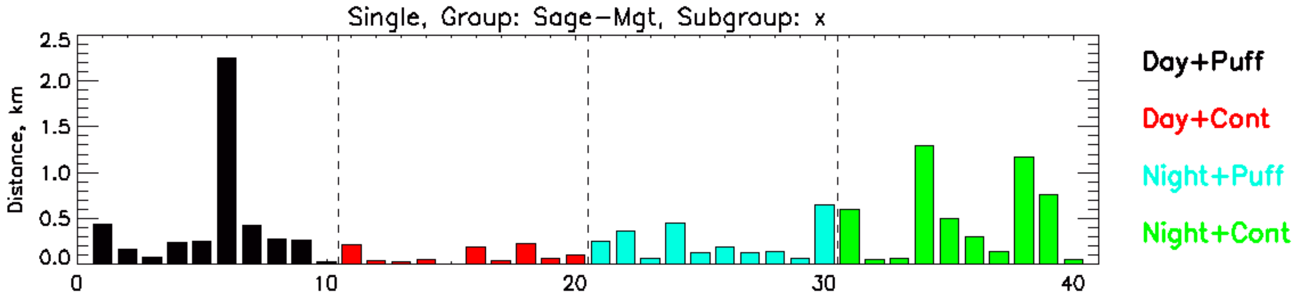


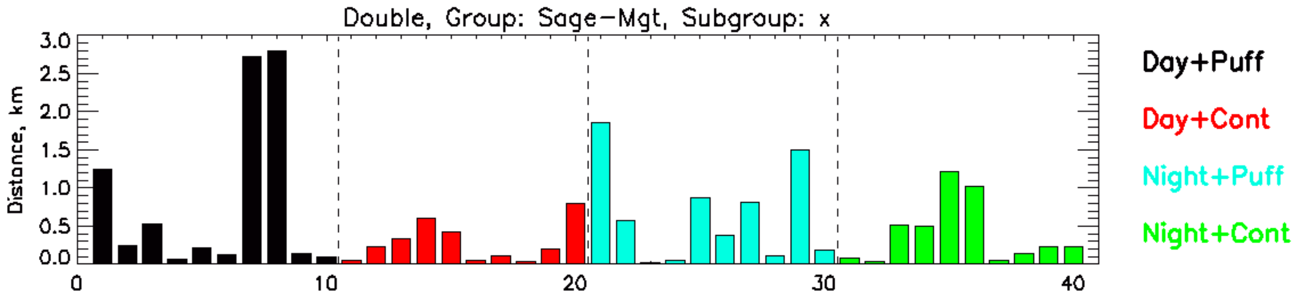




Typical "Distance Charts"
Sage-Mgt Predictions (Linear), Single and Double







Sample Aggregated Source Location Chart
PSU / Gaussian Predictions