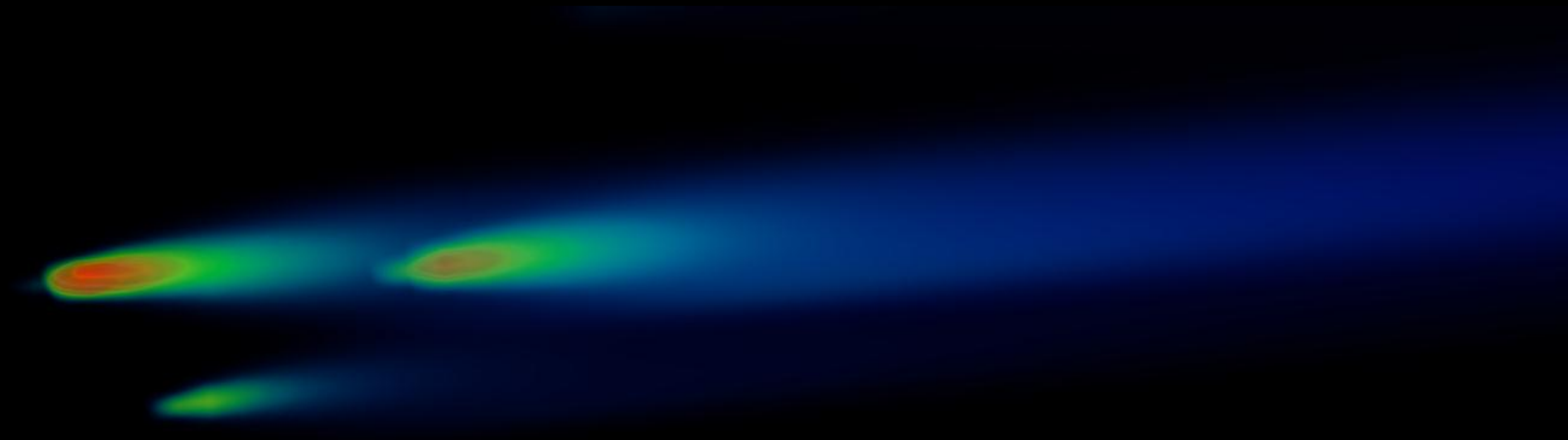


Multi-Source Atmospheric Dispersion Event Reconstruction Using Bayesian Inference and Composite Model Ranking

Derek Wade, Inanc Senocak

Boise State University,
Department of Mechanical and
Biomedical Engineering

DTRA Defense Threat Reduction Agency
Boise State University College of Engineering
NSF

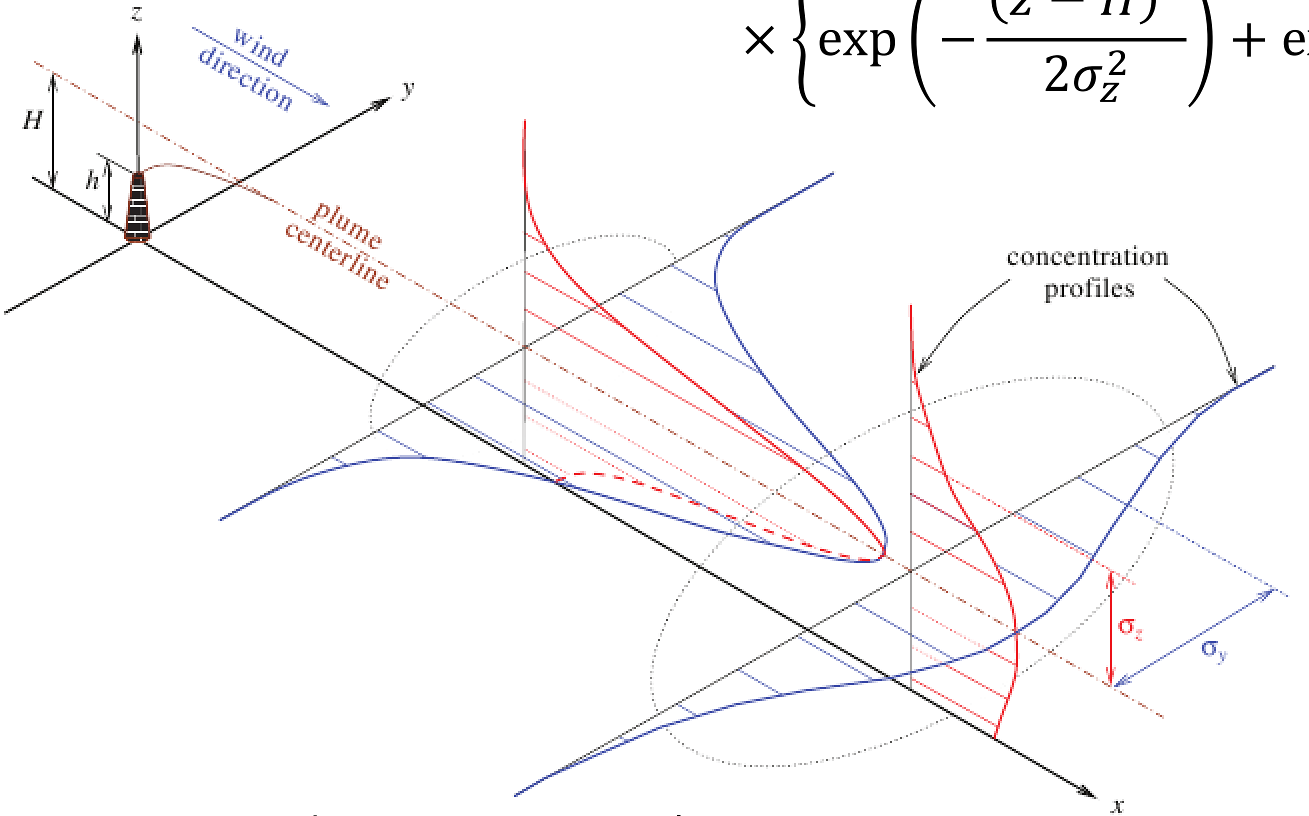
Motivation: Source term estimation of multi-source releases

(source: Red River Radio (NPR))

- Fukushima Nuclear Reactor Meltdown
- Reactor unreachable
- Radioactive material released from **multiple locations**

Gaussian Plume Forward Model

$$C_m(x,y,z) = \frac{Q}{2\pi U\sigma_y\sigma_z}\exp\left(-\frac{y^2}{2\sigma_y^2}\right)$$
$$\times\left\{\exp\left(-\frac{(z-H)^2}{2\sigma_z^2}\right)+\exp\left(-\frac{(z+H)^2}{2\sigma_z^2}\right)\right\}$$



Dispersion of passive contaminant over flat terrain with steady wind conditions

(Picture Source: Stockie, SIAM Review, vol 53, No. 2, pp. 349-372. 2010)

Inverse problem

$$\boldsymbol{m} \approx F^{-1}(\boldsymbol{d})$$

where,

- F is the forward model, the Gaussian plume model in this case (can be other)
- $\boldsymbol{d}$ is a vector of observed concentration values (i.e. sensor readings)
- $\boldsymbol{m}$ is a set of forward model parameters (i.e. source location, emission rate, etc.)
- Probabilistic approach

Bayesian Formulation

$$P(\boldsymbol{m}|\boldsymbol{d}) \propto L(\boldsymbol{d}|\boldsymbol{m})P(\boldsymbol{m})$$

where,

- $P(\boldsymbol{m}|\boldsymbol{d})$ is the probability of the forward model parameters, given the observed concentrations
- $L(\boldsymbol{d}|\boldsymbol{m})$ is the Likelihood of the observations given the forward model parameters
- $P(\boldsymbol{m})$ is a set of prior probabilities for each forward model parameter

Conditional Likelihood Formulation

$$L(d_i|\boldsymbol{m}) = \begin{cases} \exp\left(-\alpha \cdot \widehat{C}_i\right), & d_i = 0 \\ \dfrac{1-\exp\left(-\alpha \cdot \widehat{C}_i\right)}{\sqrt{2\pi}\sigma d_i} \exp\left(\dfrac{1}{2\sigma^2}\left(\ln(d_i) - \ln\left(\widehat{C}_i\right)\right)^2\right), & d_i > 0 \end{cases}$$

(Senocak et al., 2008)

- $\alpha = \dfrac{1}{C_{th}} \ln(2)$, due to ideal sensor detection with probability of ½.
- Given model parameters, $\boldsymbol{m}$, an ideal sensor, ξ_i, has a lognormal distribution with density:

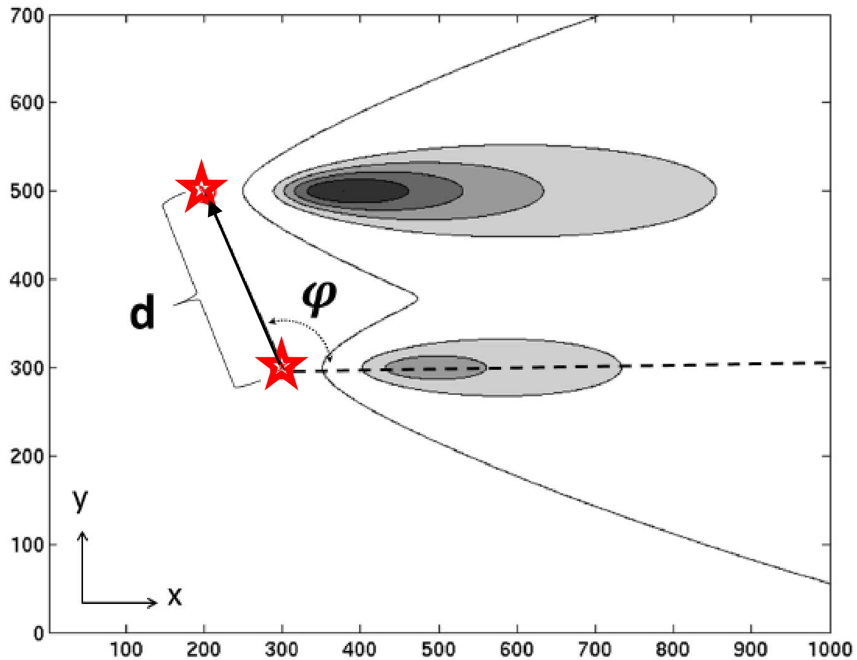
$$p(\xi_i|\boldsymbol{m}) = \frac{1}{\sqrt{2\pi}\sigma\xi_i} \exp\left(\frac{1}{2\sigma^2}\left(\ln(\xi_i) - \ln\left(\widehat{C}_i\right)\right)^2\right)$$

(Senocak et al., 2008)

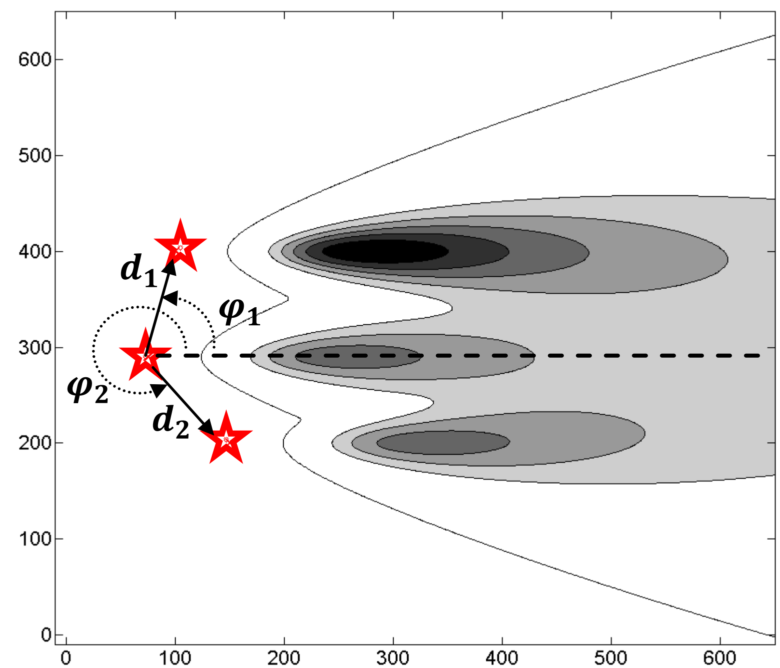
Extending it to Multi-source

$$\boldsymbol{m} = \left[x_{s1}, y_{s1}, \left(\frac{Q}{U}\right)_1, \theta, \zeta_1, \zeta_2, \sigma^2, d, \varphi, \left(\frac{Q}{U}\right)_2\right]$$

2 source



3 source



MCMC Sampling

- Markov Chain Monte Carlo (MCMC) via the Metropolis Algorithm (Metropolis et al. 1953) simulates samples from the posterior distributions
- Candidate state $\boldsymbol{m}^*$ is sampled from Gaussian distribution centered on previous state, $\boldsymbol{m}$, and accepted with probability

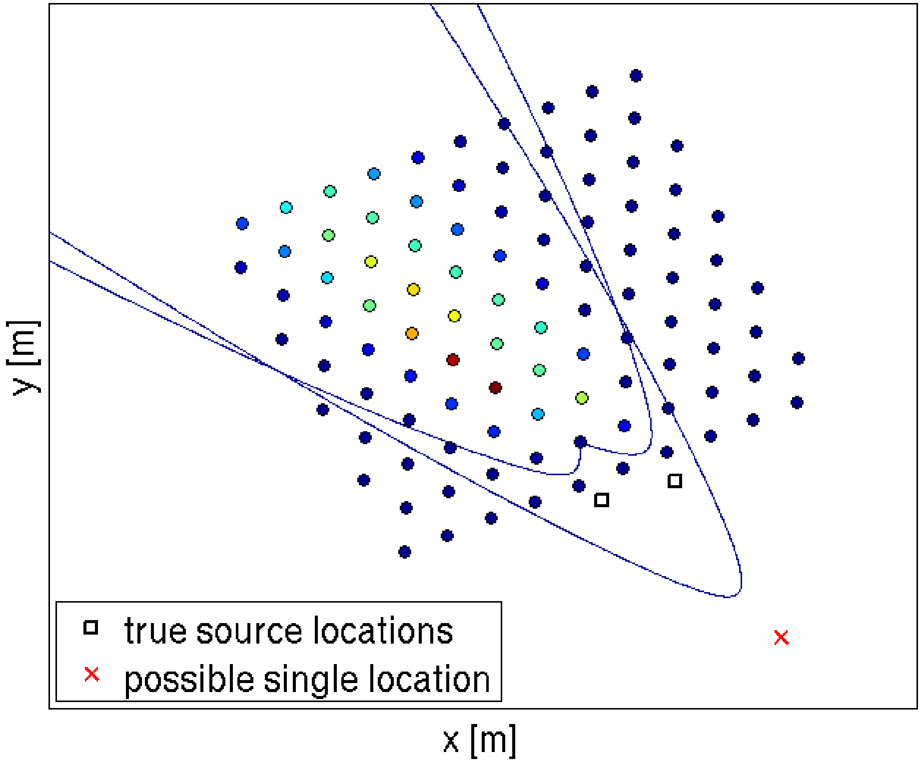
$$p(\boldsymbol{m}, \boldsymbol{m}^*) = \min\left(\frac{\pi(\boldsymbol{m}^*)}{\pi(\boldsymbol{m})}, 1\right)$$

Where target distribution is

$$\pi(\boldsymbol{m}) = L(\boldsymbol{d}|\boldsymbol{m}) \cdot p(\boldsymbol{m})$$

How many sources are there?

- Possible to get a reasonable result with incorrect number of sources
- Proposed a **composite model ranking** system to quantitatively choose the correct number of sources



Ranking Metrics

Ranking Formula

- Overall Ranking Range: 0 (worst) – 3 (best)

$$RANK = \left(1 - \|\boldsymbol{x}\|_{2,scaled}\right) + \left(1 - \frac{|FB|}{2}\right) + R^2$$

- Fractional Bias
  - Quantifies Bias (Over/Underestimation)

$$FB = 2\left(\frac{\bar{C} - \bar{\hat{C}}}{\bar{C} + \bar{\hat{C}}}\right)$$

- l_2 norm of relative error
  - Quantifies Error

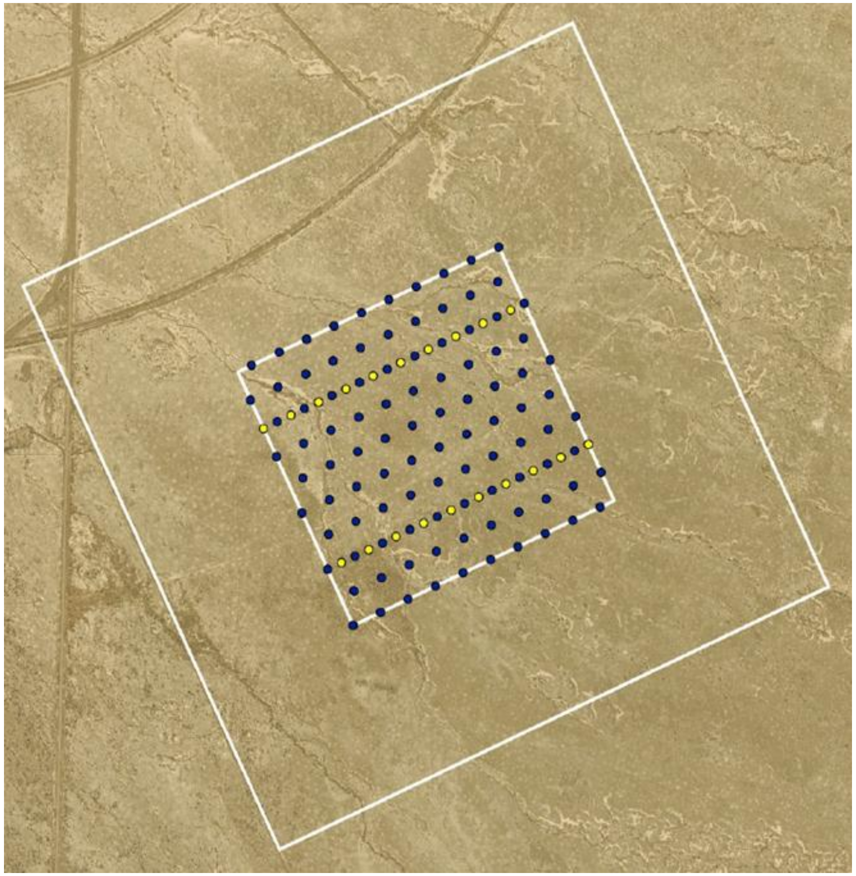
$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n}\left(\frac{|C_i - \hat{C}_i|}{C_i}\right)^2}$$

- Pearson's correlation coefficient
  - Quantifies Correlation

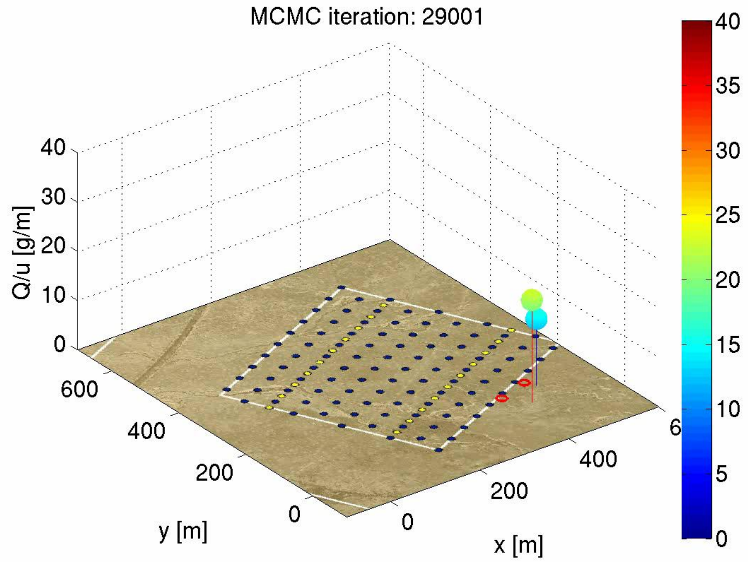
$$R = \frac{\sum_i \left[(C_i - \bar{C})\left(\hat{C}_i - \bar{\hat{C}}\right)\right]}{\left[\sqrt{\sum_i (C_i - \bar{C})^2}\right]\left[\sqrt{\sum_i \left(\hat{C}_i - \bar{\hat{C}}\right)^2}\right]}$$

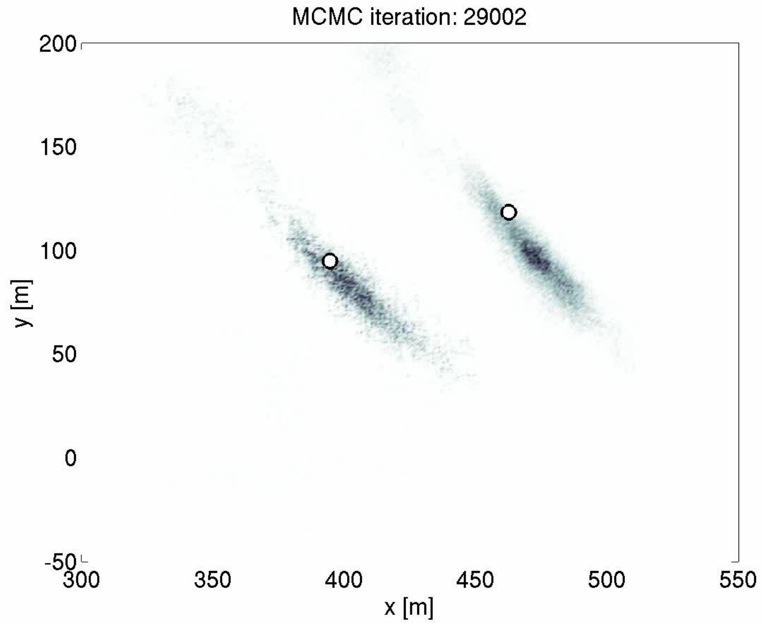
Application to FUSION Field Trials 2007 (FFT-07) and Synthetic Trials

- Real trials conducted by U.S. Army at Dugway Proving Grounds, Utah
- 100 sensors placed 50m apart in test space 1km x 1km
- Single and multiple source releases of Propylene Gas in continuous and puff trials as part of FFT-07 data set.
- **Goal**: based on sensor data, reconstruct the source and all associated parameters and correctly determine number of sources involved in release



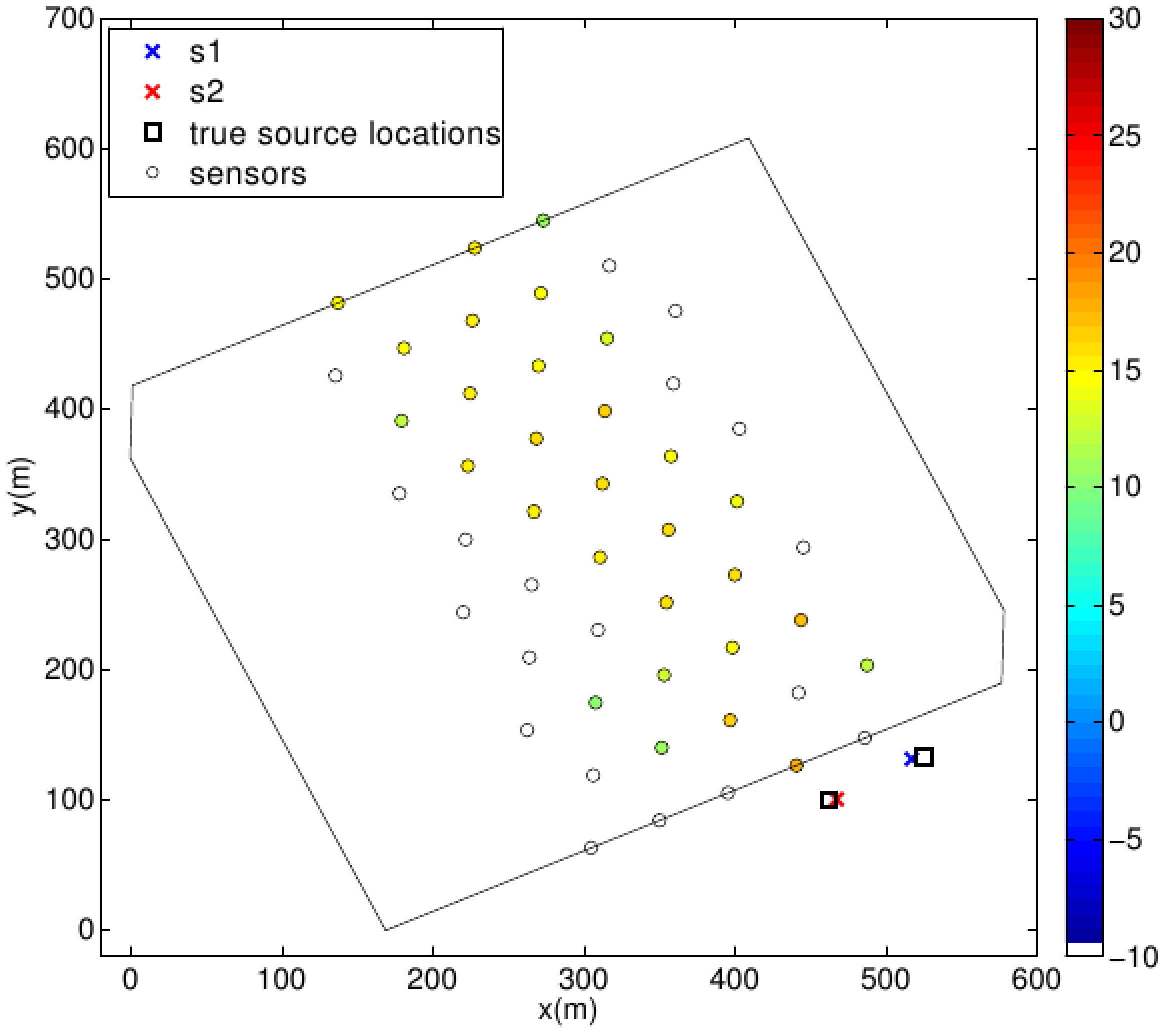
Synthetic Trial MCMC movie





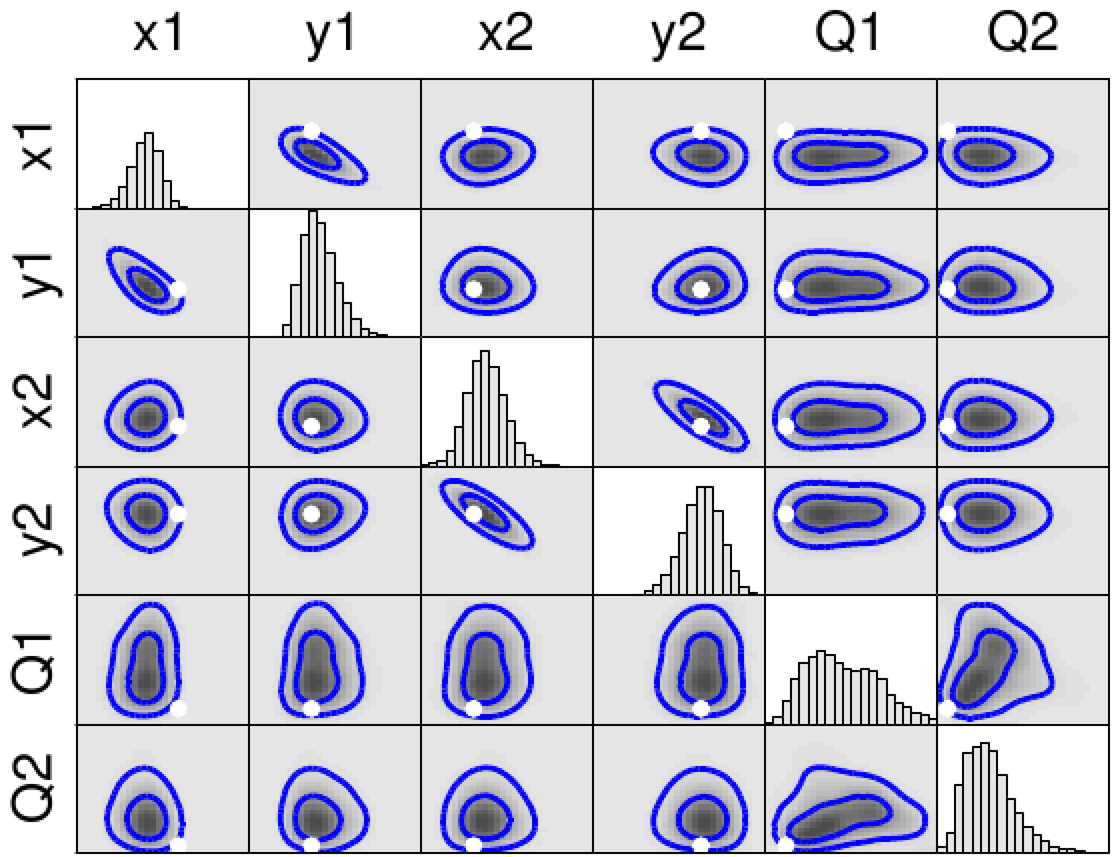
Trial 40 from FFT-07 Data Set

- 48 sensors used
- Error of 8m and 6m for source 1 and source 2, respectively



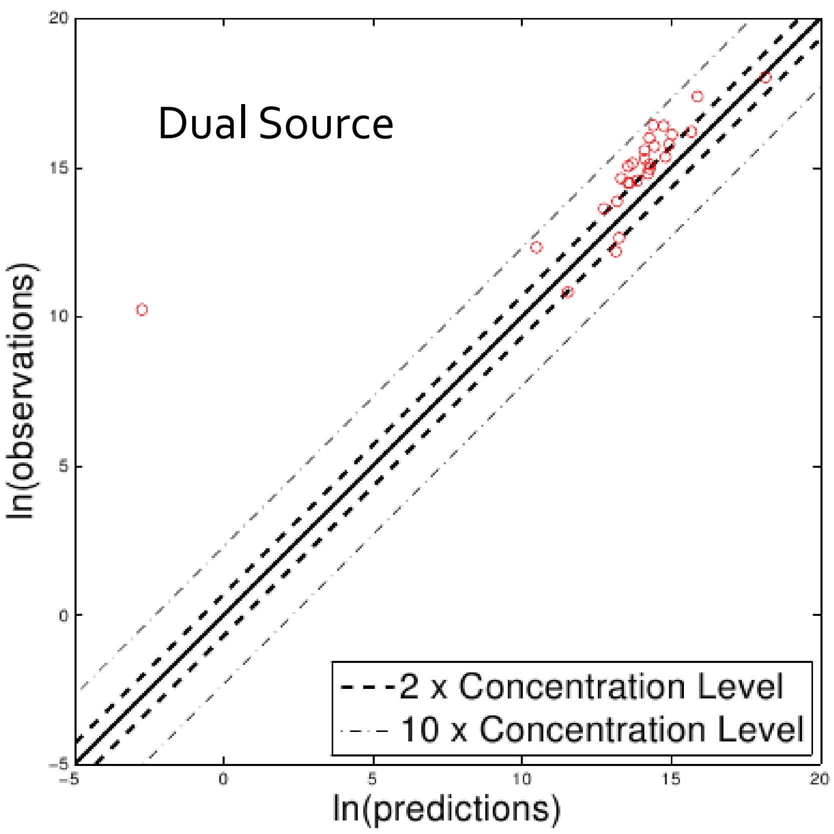
Trial 40 from FFT-07 Data Set

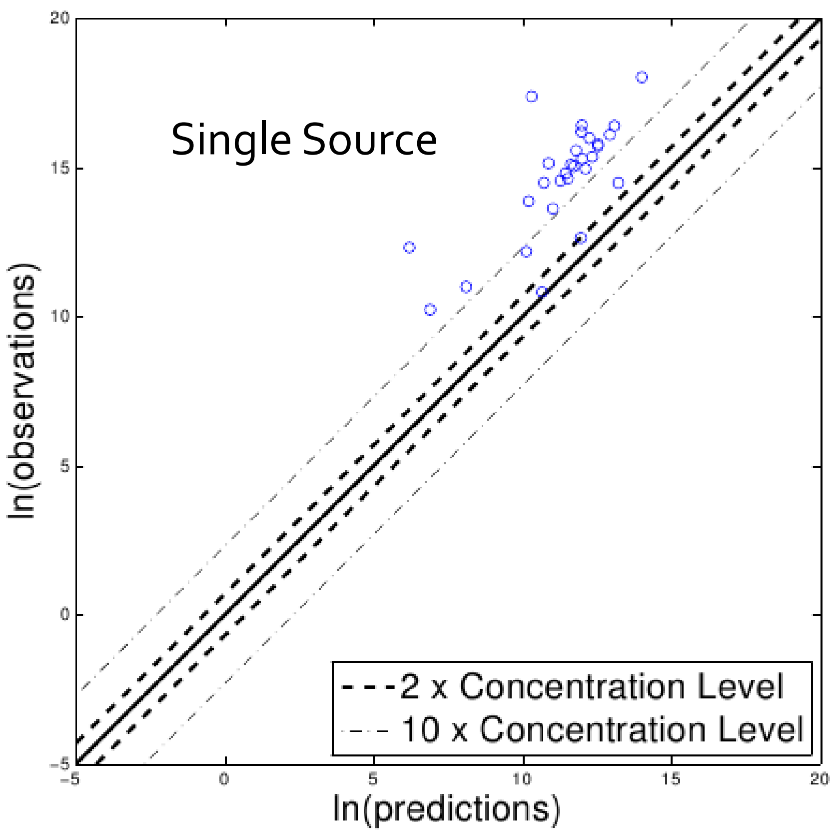
- Bivariate and marginal posterior probabilities for location and release rate, Q.
- Marginal distributions approximately Gaussian



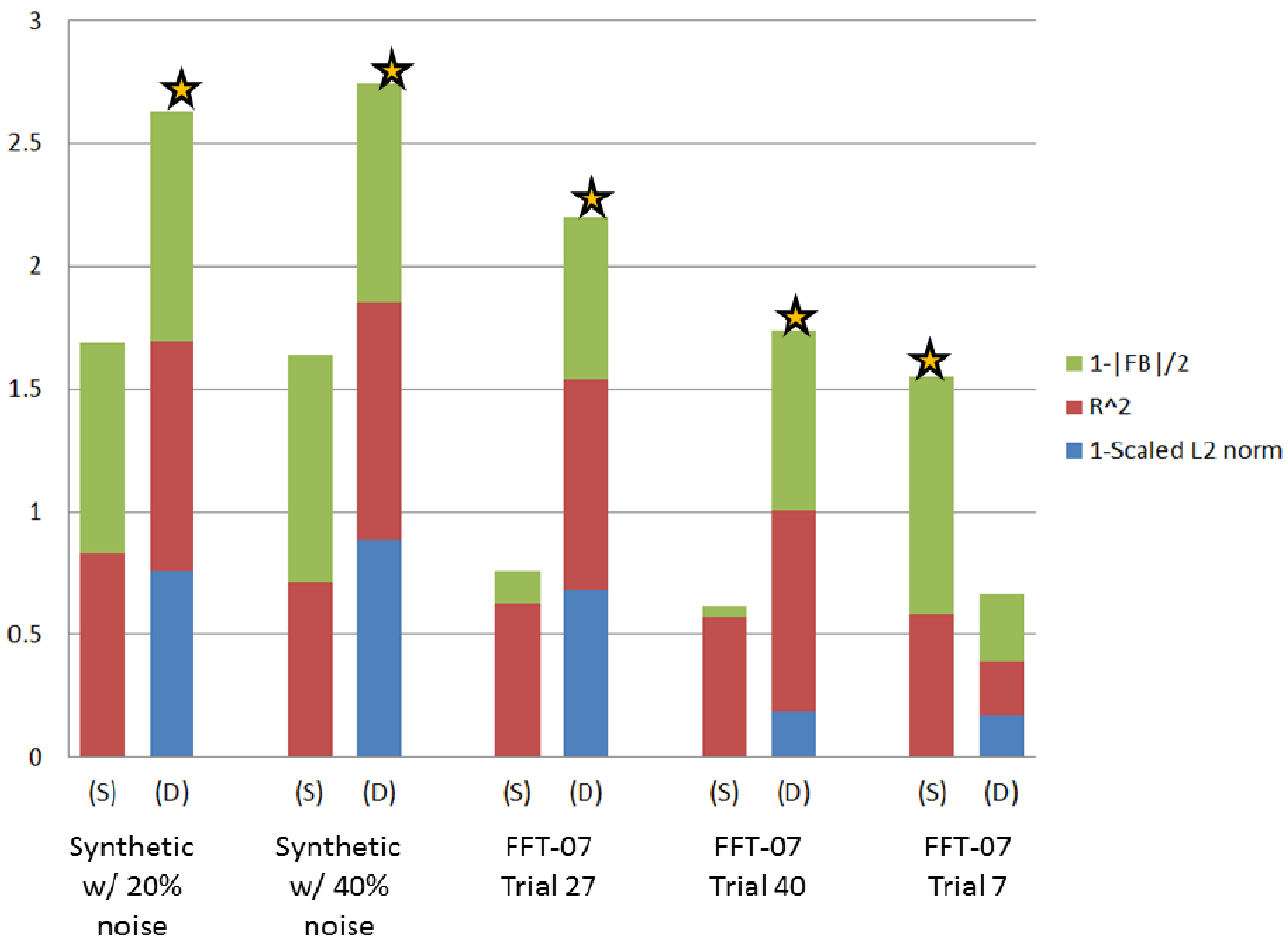
Trial 40 from FFT-07 Data Set

- Observed data vs. calculated (predicted) data





Composite Model Ranking Results



References* and Contact Info

Derek Wade
derekwade@u.boisestate.edu
Department of Mechanical and Biomedical Engineering:
High Performance Simulation Laboratory for Thermo-Fluids

Acknowledgements:
NSF-ATD Award No. 1043107

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, E., 1953. Equations of state calculations by fast computing machines. Journal of Chemical Physics 224, 560-568.

Senocak, I., Hengartner, N., Short, M., Daniel, W., 2008. Stochastic event reconstruction of atmospheric contaminant dispersion using bayesian inference. Atmospheric Environment 42, 7718-7727.

Stockie, J., 2011. The mathematics of atmospheric dispersion modeling. SIAM Review.

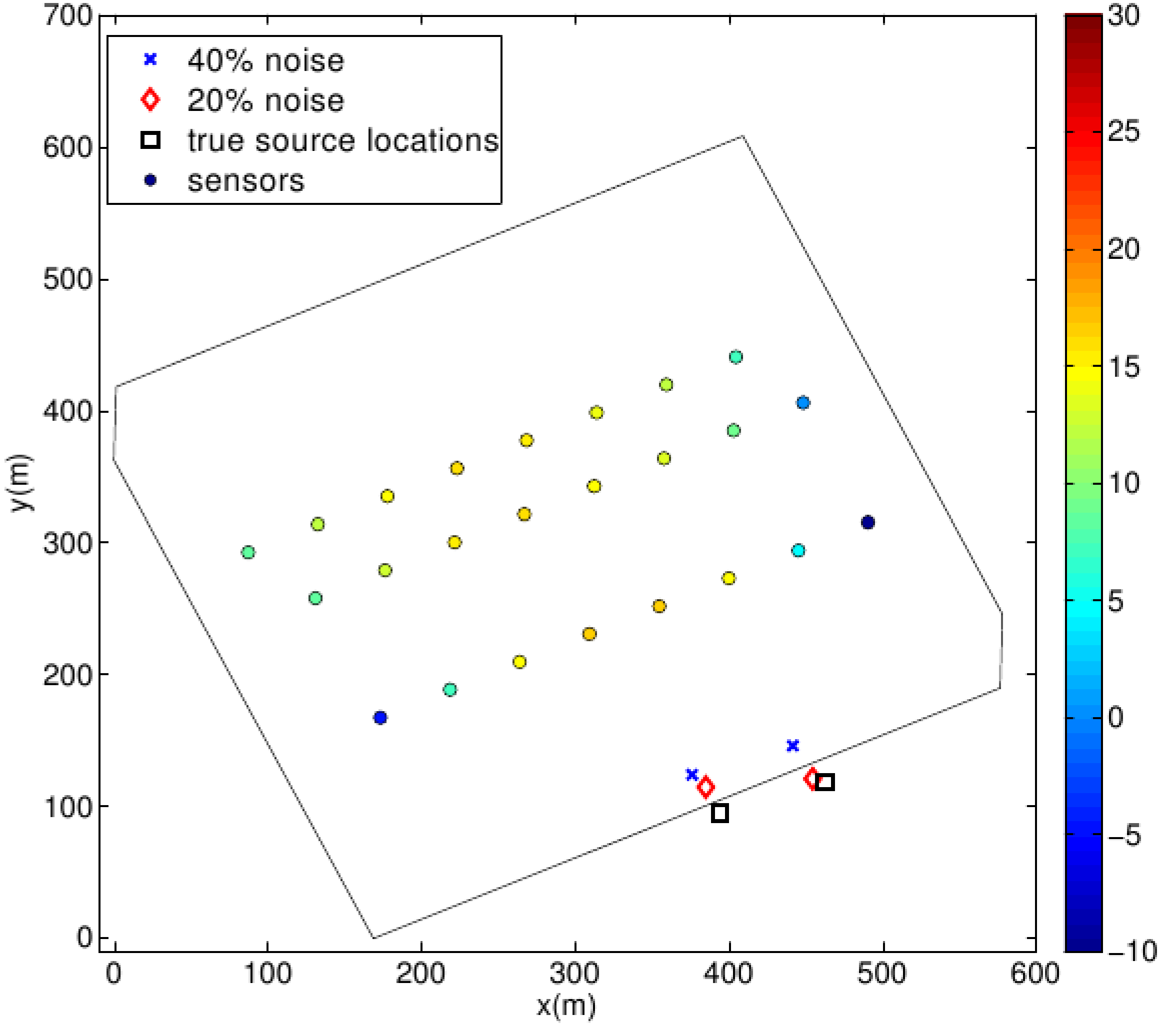
Storwold Jr., D., 2007. Detailed Test Plan for the Fusing Sensor Information from Observing Netwoks (FUSION) Field Trial 2007 (FFT-07). Technical Report WDTC- TP-07-078. U.S. Army.

*Extended Reference List Available

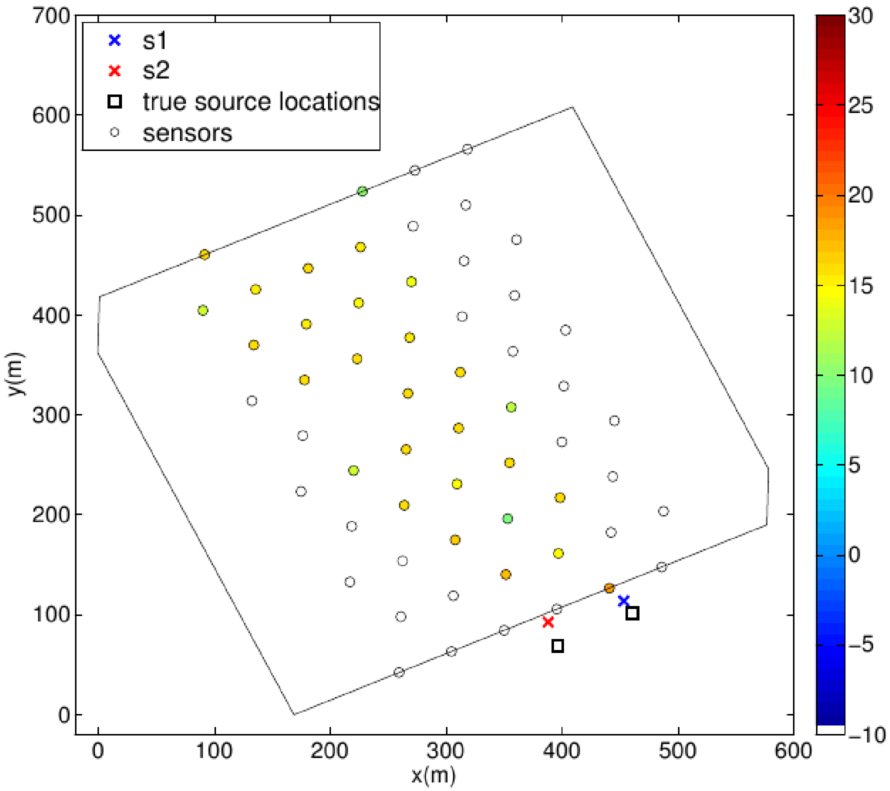
Additional Trial Data

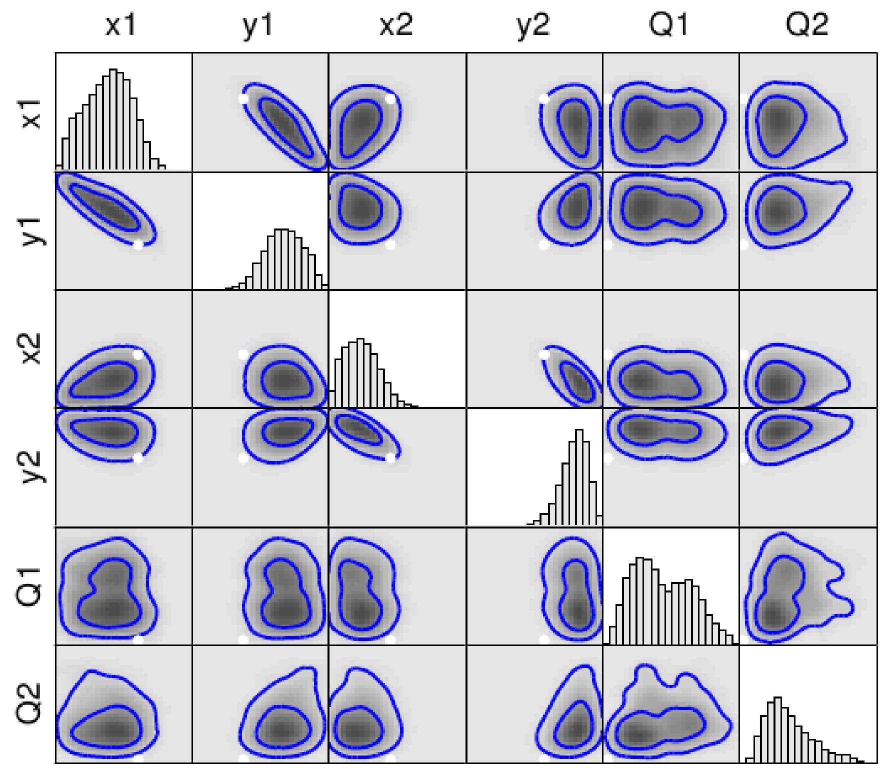
Synthetic Trial Results

- 24 sensors used
- Tested with 20% and 40% Gaussian noise
- Average error of 15.5m and 35.6m, respectively



Trial 27 from FFT-07 Data Set





- Source location Errors of 15m and 25m