Efficient estimates of prior information and uncertainty with chi-square tests

Jodi L. Mead

Department of Mathematics

BOISE STATE UNIVERSITY

Collaborators
A. Bennett, M. Gribb, C. Hammerquist, J. McNamara, R. Renaut, I. Senocak

Large-scale Inverse Problems and Quantification of Uncertainty, June 10, 2011
Outline

Priors and Uncertainty in Ill-posed Inverse Problems

Lagrangian Ocean Data Assimilation

$\chi^2$ Tests for Parameter Estimation and Uncertainty Quantification

Soil Moisture Estimation

$\chi^2$ Tests for Nonlinear Problems

Event Reconstruction for a Contaminant Release

Summary
Ill-posed inverse problems

A problem is typically ill-posed because the system cannot be completely described by the data. It can be resolved by:

1. Adding more information.
   - Probabilistic view: Bayesian statistics.
   - Deterministic view: regularization.

2. Allowing flexibility in the solution.
   - Probabilistic view: uncertainty.
   - Deterministic view: weights.
Bayesian Inference

\[ p(x|d) \propto p(d|x)p(x) \]

- Characterize posterior distribution given data and prior information.
- An approach to regularization where modeling issues are taken into consideration.
- Monte Carlo methods for sampling the posterior
  - Sample based inference - MCMC.
  - Polynomial chaos expansions
    - Polynomials used to model stochastic process.
Point estimates

\[ d = F(x) + \epsilon \]

- MAP estimate - parameters that maximize probability data were observed:
  \[ \hat{x} = \arg \max_x p(x|d) \]
  - Known optimal solution for Gaussian errors -
    \[ \epsilon \sim N(0, \sigma^2_\epsilon), \quad x \sim N(x_0, \sigma^2_f) \]
  \[ p(x|d) \propto \exp\left\{ -\frac{1}{2}(d-F(x))^T C^{-1}_\epsilon (d-F(x)) - \frac{1}{2}(x-x_0)^T C^{-1}_f (x-x_0) \right\} \]
  - Optimal for non-Gaussian is challenging.
- Minimum mean-square error estimator:
  - Kalman filter - sequentially update the prior with posterior.
  - Variational - global minimum with specified priors.
Minimum mean-square error estimator

\[ \mathbf{d} = \mathbf{F}(\mathbf{x}) + \mathbf{\epsilon} \]
\[ \mathbf{x} = \mathbf{x}_0 + \mathbf{f} \]

\[ \hat{\mathbf{x}} = \arg\min_{\mathbf{x}} (\mathbf{d} - \mathbf{F}(\mathbf{x}))^T \mathbf{C}_\epsilon^{-1}(\mathbf{d} - \mathbf{F}(\mathbf{x})) + (\mathbf{x} - \mathbf{x}_0)^T \mathbf{C}_f^{-1}(\mathbf{x} - \mathbf{x}_0) \]

- Probabilistic view: sub-optimal MAP estimate when errors are non-Gaussian.
- Deterministic view: covariance matrices weight errors and regularize solution.
- Advantage: Known, unique solution.
- Disadvantage: Heavily relies on choice of \( \mathbf{C}_f \) and \( \mathbf{C}_\epsilon \) as illustrated in following application to Lagrangian dynamics.
- \( \chi^2 \) tests: Approach to finding good choices for \( \mathbf{C}_f \) or \( \mathbf{C}_\epsilon \).
Oceanographic float data in the North Atlantic

RAFOS sub-surface float
Oceanographic example with simulated data

Data: Lagrangian float data

\[ \mathbf{d} = [\lambda \ \theta \ h]^T + \mathbf{\epsilon} \]

Mathematical model: Lagrangian shallow water equations

\[
\begin{align*}
\frac{\partial^2 \lambda}{\partial t^2} &= 2 \tan \theta \frac{\partial \lambda}{\partial t} \frac{\partial \theta}{\partial t} + \frac{f}{\cos \theta} \frac{\partial \theta}{\partial t} - \frac{g}{r^2 \cos^2 \theta} J^{-1} \left( \frac{\partial h}{\partial \alpha} \frac{\partial \theta}{\partial \beta} - \frac{\partial \theta}{\partial \alpha} \frac{\partial h}{\partial \beta} \right) \\
&\quad + \text{friction}_\lambda + f \lambda \\
\frac{\partial^2 \theta}{\partial t^2} &= -\sin \theta \cos \theta \left( \frac{\partial \lambda}{\partial t} \right)^2 - f \cos \theta \frac{\partial \lambda}{\partial t} - \frac{g}{r^2} J^{-1} \left( \frac{\partial \lambda}{\partial \alpha} \frac{\partial h}{\partial \beta} - \frac{\partial h}{\partial \alpha} \frac{\partial \lambda}{\partial \beta} \right) \\
&\quad + \text{friction}_\theta + f \theta \\
\frac{\partial}{\partial t} (h \cos \theta J) &= f_h \\
J &\equiv \frac{\partial \lambda}{\partial \alpha} \frac{\partial \theta}{\partial \beta} - \frac{\partial \theta}{\partial \alpha} \frac{\partial \lambda}{\partial \beta} \\
\lambda(0) &= \alpha + i \lambda, \quad \theta(0) = \beta + i \theta, \quad h(0) = h_0 + i h
\end{align*}
\]
Minimum mean square estimator

**Variational Formulation**

*Find dynamics that fit data within specified errors*

\[
(\hat{\lambda}, \hat{\theta}, \hat{h}) = \arg\min_{(\lambda,\theta,h)} J
\]

\[
J (\lambda, \theta, h) = \int \int \int f_\lambda C_{f\lambda}^{-1} f_\lambda + \int \int \int f_\theta C_{f\theta}^{-1} f_\theta + \int \int \int f_h C_{f_h}^{-1} f_h
\]

\[
+ \int \int i_\lambda C_{i\lambda}^{-1} i_\lambda + \int \int i_\theta C_{i\theta}^{-1} i_\theta + \int \int i_h C_{i_h}^{-1} i_h
\]

\[
+ \epsilon^T C_\epsilon^{-1} \epsilon
\]

**Representer solution**

Choice of error variances

<table>
<thead>
<tr>
<th></th>
<th>Std. dev. as % of range of reasonable values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dynamics</td>
</tr>
<tr>
<td></td>
<td>acceleration</td>
</tr>
<tr>
<td>1</td>
<td>40 %</td>
</tr>
<tr>
<td>2</td>
<td>0.1 %</td>
</tr>
<tr>
<td>3</td>
<td>1 %</td>
</tr>
</tbody>
</table>

- Experiment 1 heavily weights the dynamics.
- Experiment 2 heavily weights data.
- Experiment 3 doesn’t heavily weight either.

*The solution will go wherever we place weight*
Assimilation Results from Experiment 2

- Data
- True ocean
- Initial guess
- 1st iterate
- 2nd iterate
- 3rd iterate

Longitude vs. Latitude plot showing the assimilation process.
Assimilation Results from Experiment 3

- **data**
- **true ocean**
- **initial guess**
- **1st iterate**
- **2nd iterate**
- **3rd iterate**
Approach to estimating error variances

**Standardize the problem as:**

\[
d = Ax + \epsilon \\
x = x_0 + f
\]

\[
J(x) = (d - Ax)^T C_\epsilon^{-1}(d - Ax) + (x - x_0)^T C_f^{-1}(x - x_0)
\]

**Statistical Hypothesis Testing:**

- Null hypothesis: \( \bar{\epsilon} = \bar{f} = 0, \ \epsilon\epsilon^T = C_\epsilon, \ f^f^T = C_f, \ \epsilon f = 0 \)
- Test Statistic: \( J(\hat{x}) \sim \chi^2_m \), where \( m \) is number of data and \( \hat{x} = \operatorname{argmin}_x J(x) \).
\[ \hat{x} = x_0 + C_f A^T (A C_f A^T + C_\epsilon)^{-1} r \]

where \( r = b - A x_0 \). Let \( P = A C_f A^T + C_\epsilon \) then

\[ J(\hat{x}) = r^T P^{-1} r \]

Covariance matrices are positive semi-definite so if none of \( \epsilon_i \) nor \( f_i \) has zero variance we define \( r = P^{1/2} k \) with

\[ J(\hat{x}) = k^T k = k_1^2 + \ldots + k_m^2 \]

If the \( k_i \) are independent, standard normal then

\[ J(\hat{x}) \sim \chi^2_m \]
χ² Test continued

- The χ² tests holds if $k_i$, and hence $r_i$ are independent and Gaussian.
- The χ² test still holds if $r_i$ are not Gaussian since

$$k_i = \sum_{j=1}^{m} P_{ij}^{-1/2} r_j \quad i = 1, \ldots, m$$

behave Gaussian when $m$ is large.
- The $k_i$ are necessarily standard normal since

$$E(kk^T) = P^{-1/2} E(rr^T)P^{-1/2} = I$$

and $\bar{k} = 0$ if $\bar{r} = 0$. If $\bar{r} \neq 0$ then $J(\hat{x})$ follows a non-central χ²_m distribution.
- If $J(\hat{x}) \sim \chi^2_m$ we fail to reject the null hypothesis with $C_\epsilon$ and $C_f$ as error covariance matrices.
\( \chi^2 \) Method

- Find \( C_\epsilon \) or \( C_f \) that ensures \( J(\hat{x}) \sim \chi^2_m \).
  - For \( C = C_f \) or \( C = C_\epsilon \) find roots of
    \[
    f(C) = J(\hat{x}) - m
    \]

**Relationship to Morozov’s discrepancy principle**

- Discrepancy: Find \( C_f = \delta I \) that ensures \( J_d(x) \sim \chi^2_m \)
  \[
  J_d(x) = (d - Ax)^T C^{-1}_\epsilon (d - Ax)
  \]
  - Implemented as finding roots of
    \[
    f(\delta) = J_d(\hat{x}) - m
    \]

which can give poor solutions since \( J_d(\hat{x}) \sim \chi^{2}_{m-n} \)

- Note that

\[
J(\hat{x}) = J_d(\hat{x}) + (\hat{x} - x_0)^T C_f (\hat{x} - x_0)^T \\
\sim \chi^2_{m-n} + \chi^2_n
\]
Approaches to calculating error variances

- Regularization
  - L-curve, GCV, UPRE, etc.
  - Assumes $C_f = \lambda^{-1}I$.

- Bayesian hierarchical models

\[
p(x, \gamma | d) \propto p(d|x)p(x|\gamma)p(\gamma)
\]

  - In addition to parameters, calculate hyperparameters $\gamma$.

- $\chi^2$ method
  - Use $\chi^2$ tests to estimate covariance matrices.
Scalar $\chi^2$ Method

- Finds $C_\epsilon$ or $C_f$ that pass the $\chi^2$ test.
  - For regularization find $C_f = \sigma_f^2 I$ with $C_f$ given via roots of
    \[
    f(\sigma_f) = r^T (P(\sigma_f, C_\epsilon))^{-1} r - m
    \]
  - Or find $C_\epsilon = \sigma_\epsilon^2 I$ with $C_f$ given via roots of
    \[
    f(\sigma_\epsilon) = r^T (P(C_f, \sigma_\epsilon))^{-1} r - m
    \]


Benchmark problem


Noise in data

Problem Phillips Right Hand Side

Problem Shaw Right Hand Side

- exact
- noise .1651

- exact
- noise .213
Convergence of Newton algorithm for $\chi^2$ Method

\[ n = 64 \text{ over 500 runs.} \]

<table>
<thead>
<tr>
<th></th>
<th>shaw</th>
<th></th>
<th>phillips</th>
<th></th>
<th>ilaplace</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>noise</td>
<td>$K$</td>
<td>noise</td>
<td>$K$</td>
<td>noise</td>
<td>$K$</td>
<td></td>
</tr>
<tr>
<td>0.008</td>
<td>9.0(2.0)</td>
<td>0.006</td>
<td>9.1(2.2)</td>
<td>0.003</td>
<td>7.2(1.1)</td>
<td></td>
</tr>
<tr>
<td>0.084</td>
<td>5.6(1.1)</td>
<td>0.064</td>
<td>7.8(2.3)</td>
<td>0.034</td>
<td>9.1(4.3)</td>
<td></td>
</tr>
<tr>
<td>0.166</td>
<td>5.2(1.1)</td>
<td>0.128</td>
<td>8.2(3.4)</td>
<td>0.069</td>
<td>7.9(4.1)</td>
<td></td>
</tr>
<tr>
<td>0.008</td>
<td>9.0(1.9)</td>
<td>0.006</td>
<td>9.1(2.2)</td>
<td>0.003</td>
<td>7.1(1.2)</td>
<td></td>
</tr>
<tr>
<td>0.084</td>
<td>5.6(1.1)</td>
<td>0.064</td>
<td>7.7(2.2)</td>
<td>0.034</td>
<td>9.4(4.4)</td>
<td></td>
</tr>
<tr>
<td>0.166</td>
<td>5.2(1.0)</td>
<td>0.128</td>
<td>8.1(3.4)</td>
<td>0.069</td>
<td>8.5(4.4)</td>
<td></td>
</tr>
</tbody>
</table>
Mean and Standard Deviation of Error

\[ n = 64 \text{ over 500 runs} \]

<table>
<thead>
<tr>
<th>Problem</th>
<th>noise</th>
<th>L-Curve</th>
<th>GCV</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>shaw</td>
<td>0.166</td>
<td>0.1211(0.0266)</td>
<td>0.4370(0.2934)</td>
<td>0.1019(0.0235)</td>
</tr>
<tr>
<td>phillips</td>
<td>0.128</td>
<td>0.1490(0.1191)</td>
<td>0.1686(0.2018)</td>
<td>0.1004(0.0010)</td>
</tr>
<tr>
<td>ilaplace</td>
<td>0.069</td>
<td>0.3791(0.2186)</td>
<td>0.2985(0.2464)</td>
<td>0.1473(0.1122)</td>
</tr>
</tbody>
</table>
Full $\chi^2$ Method

- Let $C = \text{diag}(\sigma_1^2, \ldots, \sigma_1^2, \ldots, \sigma_2^2, \ldots, \sigma_2^2, \ldots, \sigma_d^2, \ldots, \sigma_d^2)$
- Solve for $\sigma_1, \sigma_2, \ldots, \sigma_d$ with multiple $\chi^2$ principles:
  
  
  \[
  k_1^2 + \ldots + k_{m_1}^2 = m_1 \\
  k_{m_1+1}^2 + \ldots + k_{m_1+m_2}^2 = m_2 \\
  \vdots \\
  k_{m+1-m_d}^2 + \ldots + k_m^2 = m_d 
  \]

  \[
  \sum_{i=1}^{d} m_i = m.
  \]
- Solve $F(C) = 0$ with a multidimensional Newton Method, submitted to *Computational Statistics and Data Analysis*.
Discontinuous test problem with full $\chi^2$ method

- For regularization with $F(C_f) = 0$, $d = 3$, i.e. find $\sigma_1$, $\sigma_2$, $\sigma_3$

$$\hat{x} = \arg\min_x \left\{ (b - Ax)^T C_\epsilon^{-1} (b - Ax) \right\}$$

$$+ (x - x_0)^T \begin{bmatrix} \sigma_1 I & 0 & 0 \\ 0 & \sigma_2 I & 0 \\ 0 & 0 & \sigma_3 I \end{bmatrix} (x - x_0)$$

**Benchmark problem**

_P.C. Hansen, 2007_

Wing - 1D problem with discontinuous solution, $m = n = 96$.

**Numerics**

Matlab functions _fsolve_ and _sqrtm_.

---

**Priors and Uncertainty**

**Lagrangian Ocean**

**$\chi^2$ Tests**

**Soil Moisture**

**Nonlinear**

**Event Reconstruction**

**Summary**
Results from discontinuous test problem *wing*
Hydrological Application

Dry Creek Watershed, Boise, ID

- Continuous representation of soil moisture in space and time, \( \theta(x, t) \) can be obtained from Richards equation (1931):

\[
\frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla \psi) + \frac{\partial K}{\partial z}
\]

- Hydraulic conductivity \( K \) is not measurable and depends on fitting parameters \( \alpha, n \), and soil hydraulic properties \( \theta_s \) and \( \theta_r \) found via van Genuchten’s equation (1980):

\[
\theta(\psi) = \theta_r + \frac{\theta_s - \theta_r}{(1 + |\alpha \psi|^n)^m}
\]

- These parameters are a function of the percentage of sand, silt, clay, etc. in the soil.
van Genuchten parameters

Approaches to parameter estimation

- In-situ measurements of $\theta$, $\psi$ and invert for parameters. Very time consuming.

- Collect soil samples, measure % of sand, silt, clay etc. and input in neural network algorithm (Rosetta).
\[ \chi^2 \text{ Method applied to Soil Moisture in Dry Creek Watershed} \]

- van Genuchten equation and parameters - \( A \) and \( x \)
- In-situ measurements - \( b \)
  - Uncertainties due to measurement error and spatial variability, i.e. \( C_\epsilon \) are unknown.
  - Single \( \chi^2 \) Principle found \( C_\epsilon = \sigma I \)
  - Multiple \( \chi^2 \) Principles found
    \[ C_\epsilon = \text{diag}(\sigma_1^2, \ldots, \sigma_2^2, \ldots, \sigma_N^2, \ldots, \sigma_N^2), \quad N = 2, 3, 4, 5, 10. \]
- Laboratory parameter estimates - \( x_0 \)
  - Uncertainties \( C_f \) are given by neural network algorithm.
- Model spatial variability as uncertainty
  - Consolidate measurements over multiple soil pits.
    - The scale over which one parameter set can be used is determined by value of \( C_\epsilon \).
    - These parameters and corresponding uncertainties can be propagated in flow models.
Dry Creek Instrumentation

North facing slope

South facing slope
In-situ

Full $\chi^2$ Method
SU10_40

- one $\chi^2$ test
- two $\chi^2$ test
- three $\chi^2$ test
- four $\chi^2$ test
- five $\chi^2$ test
- ten $\chi^2$ test
### Scalar $\chi^2$ Method - Estimates of $\sigma_\epsilon$

#### Standard Deviation for Single Pits

<table>
<thead>
<tr>
<th></th>
<th>NU10_15</th>
<th>SU10_20</th>
<th>SU10_40</th>
<th>SD5_15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$ principle</td>
<td>0.005 (2.6%)</td>
<td>0.011 (6.4%)</td>
<td>0.044 (23.1%)</td>
<td>0.011 (6.4%)</td>
</tr>
</tbody>
</table>

#### Standard Deviation for South facing slope

<table>
<thead>
<tr>
<th></th>
<th>SU10_20 and SU10_40</th>
<th>SU10_20 and SD5_15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$ principle</td>
<td>0.048 (22.5%)</td>
<td>0.031 (14.5%)</td>
</tr>
<tr>
<td>in-situ</td>
<td>0.017 (8.0%)</td>
<td>0.011 (5.1%)</td>
</tr>
</tbody>
</table>

#### Standard Deviation for North and South facing slopes

<table>
<thead>
<tr>
<th></th>
<th>SU10_20 and NU10_15</th>
<th>SD5_15 and NU10_15</th>
<th>SU10_20 and SD5_15 and NU10_15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$ principle</td>
<td>0.014 (7.7%)</td>
<td>0.025 (11.9%)</td>
<td>0.030 (14.0%)</td>
</tr>
<tr>
<td>in-situ</td>
<td>0.010 (5.2%)</td>
<td>0.010 (4.7%)</td>
<td>0.010 (4.8%)</td>
</tr>
</tbody>
</table>
### Full $\chi^2$ Method - Estimates of $\sigma_\epsilon$

#### SU10_40

<table>
<thead>
<tr>
<th>( \sigma_i )</th>
<th>0.044 (23%)</th>
<th>0.054 (28%)</th>
<th>0.052 (27%)</th>
<th>0.052 (27%)</th>
<th>0.052 (27%)</th>
<th>0.054 (29%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td></td>
<td>0.029 (15%)</td>
<td>0.038 (20%)</td>
<td>0.050 (26%)</td>
<td>0.051 (27%)</td>
<td>0.050 (26%)</td>
</tr>
<tr>
<td>( \sigma_3 )</td>
<td></td>
<td></td>
<td>0.021 (11%)</td>
<td>0.026 (14%)</td>
<td>0.046 (24%)</td>
<td>0.052 (27%)</td>
</tr>
<tr>
<td>( \sigma_4 )</td>
<td></td>
<td></td>
<td></td>
<td>0.025 (13%)</td>
<td>0.013 (7%)</td>
<td>0.048 (25%)</td>
</tr>
<tr>
<td>( \sigma_5 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.029 (15%)</td>
<td>0.053 (28%)</td>
</tr>
<tr>
<td>( \sigma_6 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.040 (21%)</td>
</tr>
<tr>
<td>( \sigma_7 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.016 (9%)</td>
</tr>
<tr>
<td>( \sigma_8 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.007 (4%)</td>
</tr>
<tr>
<td>( \sigma_9 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.024 (13%)</td>
</tr>
<tr>
<td>( \sigma_{10} )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.036 (19%)</td>
</tr>
</tbody>
</table>
Hydrologic Conclusions

- The high uncertainty at depth 40cm dominates.
- Use same parameters on north and south slopes.
- May want to use different parameters upstream than down.
- Wet or dry soil has more uncertain measurements than in between.
- Single $\chi^2$ Principle reasonably captures uncertainty across all pressure heads.
Nonlinear inversion

\[ d = F(x) + \epsilon \]
\[ x = x_0 + f \]

\[ J(x) = (d - F(x))^T C_\epsilon^{-1} (d - F(x)) + (x - x_0)^T C_f^{-1} (x - x_0) \]

Newton’s method

\[ \nabla^2 J(x_k)(x_{k+1} - x_k) = -\nabla J(x_k) \]

gives

\[ x_{k+1} = x_k + (C_f^{-1} + J_k^T C_\epsilon J_k)^{-1} (J_k C_\epsilon^{-1} (d - F(x_k)) - C_f^{-1} (x_k - x_0)) \]

where \( J_k \) is Jacobian of \( F(x_k) \).
Nonlinear $\chi^2$ test

Newton method for minimum gives at each iterate:

$$ J(\hat{x}_{k+1}) \sim \text{non-central } \chi^2_m $$

with non-central parameter

$$ \lambda_k = \sum_{i=1}^{m} (P_k^{-1/2} J_k (x_k - x_0))^2_i $$

where $P_k = J_k C_f J_k^T + C_\epsilon$.

- Null hypothesis: $\bar{\epsilon} = \bar{f} = 0$, $\epsilon \epsilon^T = C_\epsilon$, $\bar{f} \bar{f}^T = C_f$, $\epsilon \bar{f} = 0$
- Test Statistic: $\bar{J}(\hat{x}_{k+1}) \approx m + \lambda_k$.
- Non-central $\chi^2$ test holds at each iterate if $d_i$ are identically distributed, not necessarily normal, and $m$ is large.
Atmospheric releases in an urban environment

- Chemical or biological agent released into atmosphere
  - Homeland security (bioterrorism)
  - Environmental monitoring (nuclear, pollution)
- Possibility of multiple sources
• Determine
  • source location
  • emission rate or strength
• Simulate spatial and temporal evolution of contaminant
Sensor Measurements

- Fixed network of concentration measurements
Atmospheric models for concentration of contaminants

- Advection and diffusion
- Lagrangian particle
- Steady state
  - Gaussian Plume Model

\[
C(x, y, z) = \frac{Q}{2\pi U \sigma_y \sigma_x} e^{-\frac{(y-y_s)^2}{2\sigma_y^2}} \left( e^{-\frac{(z-H)^2}{2\sigma_z^2}} + e^{-\frac{(z+H)^2}{2\sigma_z^2}} \right)
\]

\[
\sigma_y = \frac{\nu_1 x}{\sqrt{1 + 0.004(x - x_s)}} \quad \sigma_z = \nu_2(x - x_s)
\]

\(Q\) - emission strength, \(U\) - mean wind speed, \(H\) - height of release, \((x_s, y_s)\) - source location, \(x\) - direction parallel to wind, \(y\) - crosswind direction.
Sources of uncertainty

- **Sensors**
  - Measurement error, false readings
- **Models**
  - Incomplete physics, numerical approximations
- **Unknown parameters, initial or boundary conditions**

**Prediction**

![Prediction Image]

**Uncertainty**

![Uncertainty Image]
Tracer gas experiments

- Sulphurhexafluoride was released from a tower at a height of 115 m on October 19, 1978 in Copenhagen.
- Network of sensors measured 2-3 m above ground at positions 2-6 km from point of release
- Average of three consecutive 20 min tracer concentrations
Mean square error estimator with $\chi^2$ test.
Variational assimilation - comparison of variance estimation

Bivariate plot of parameter estimates and posterior covariance

Variance from Bayesian Inference  Variance from $\chi^2$ test
Comparison of Variational Assimilation and Bayesian Methods

Variational - $\chi^2$

Bayes - MCMC
Comparison of means and standard deviations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Value</th>
<th>Bayes MCMC</th>
<th>Variational assimilation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Uncertainty - given</td>
</tr>
<tr>
<td>$x_s$ (m)</td>
<td>0</td>
<td>$-23 \pm 305$</td>
<td>$-231 \pm 436$</td>
</tr>
<tr>
<td>$y_s$ (m)</td>
<td>0</td>
<td>$4 \pm 145$</td>
<td>$69 \pm 221$</td>
</tr>
<tr>
<td>$Q$ (g/s)</td>
<td>3.2</td>
<td>$13.7 \pm 20$</td>
<td>$3.75 \pm 3.15$</td>
</tr>
<tr>
<td>$H$ (m)</td>
<td>115</td>
<td>$146 \pm 117$</td>
<td>$106 \pm 55$</td>
</tr>
</tbody>
</table>

Variational assimilation gives:

- estimates of $(x_s, y_s)$ source with slightly higher uncertainty, but lie within true value.
- better estimates of emission strength.
- comparable estimates of the height of the release.
Event reconstruction conclusions

- Bayesian inference implies the solution is a probability distribution. Goal is to calculate the posterior distribution of the parameters, given the data.
  - Optimization can be computationally prohibitive for time dependent problems.
- Mean square estimator can be used to find the mean and covariance of the parameters.
  - For the steady-state model, showed that there is not a significant loss in accuracy with mean square estimate, as compared to Bayesian inference, when data are not Gaussian.
  - Success of mean square estimate depends on accurate estimate of error covariance matrices, and the \( \chi^2 \) method successfully found estimates.
Summary and main ideas

- Inverse methods can be used as a mechanism to propagate and determine uncertainty in data, parameters and models.
  - The “right” answer from an inversion is the one that combines data and models within their uncertainty bounds. The difficult part is determining these uncertainties \textit{a priori}.
- Least squares and Tikhonov regularization often smooth solutions, but if we choose to weight with covariance matrices rather than scalars we can obtain non-smooth solutions.
  - The $\chi^2$ method is an approach to estimating these covariance matrices.