Using Genetic Algorithms to Assimilate Sensor Data into Atmospheric Models

Sue Ellen Haupt
Head, Department of Atmospheric & Oceanic Physics, Applied Research Lab
Associate Professor of Meteorology

February 2, 2009
Happy Groundhog Day!!

The world's most famous groundhog saw his shadow Monday morning, predicting that this already long winter will last for six more weeks. - AP
• Why Computational Intelligence?
• Introduction to Genetic Algorithms
  – Guppy Sex – binary genetic algorithms
  – Real-valued genetic algorithms
• Application to Data Assimilation
  – Forward assimilation
  – Back-calculating the source and meteorology
• Summary
Observing Nature to Make Predictions
Red at Night –
Sailor’s Delight
Red in the Morning –
Sailors Take Warning
Observation and Classification

- **Norwegian Cyclone Model** (V. Bjerknes and J. Bjerknes)
- Explains passage of standard weather systems
- Helped in forecasting based on recent events
• F=ma

• Primitive Equations:

\[
\frac{D\vec{V}}{Dt} = -\frac{1}{\rho} \nabla P - 2\Omega \times \vec{V} + \vec{g} + \vec{F}
\]

\[
\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \vec{V})
\]

\[
P = \rho RT
\]

\[
Q = c_v \frac{dT}{dt} + P \frac{d\alpha}{dt}
\]

• If discretize, predict time rate of change

-> Numerical Forecasting
Imagine a large hall like a theatre, except that the circles and galleries go right round through the space usually occupied by the stage. …

A myriad computers are at work upon the weather of the part of the map where each sits, but each computer attends only to one equation or part of an equation. …

From the floor of the pit a tall pillar rises to half the height of the hall. It carries a large pulpit on its top. In this sits the man in charge of the whole theatre; he is surrounded by several assistants and messengers. …

In this respect he is like the conductor of an orchestra in which the instruments are slide-rules and calculating machines. But instead of waving a baton he turns a beam of rosy light upon any region that is running ahead …

L.F. Richardson, 1922
The ENIAC machine occupied a room thirty by fifty feet. 1946 Charney used filtered equations to produce first numerical forecast.

http://www.library.upenn.edu/special/gallery/mauchly/jwm0-1.html
• computing power increased

Lorenz and Recognition of Chaos

- Sensitivity to initial conditions
- Chaos – limits to predictability
- Think in terms of attractors & manifolds
- Requires
  - assimilation
  - initialization
  - statistical forecasting
  - ensemble forecasts
  - empirical models
Two distinct approaches to weather forecasting

1. Equation based – numerical integration and pre- and post-processing
2. Empirically based – begin with data and find patterns → Computational intelligence
Computational Intelligence

- New methods of using observed data to make sense of observations
- Combine with increases in computer power
- Artificial intelligence methods both leverage and offer an alternative to traditional methods
Part I – Intro to AI
- Environmental Science Models and Artificial Intelligence
- Basic Statistics and Basic AI: Neural Networks
- Performance Measures and Uncertainty
- Decision Trees
- Introduction to Genetic Algorithms
- Introduction to Fuzzy Logic
- Missing Data Imputation through Machine Learning Algorithms

Part II – Application of AI in Environmental Science
Genetic Algorithms (GA) – Search and Optimization

Combine:

- Genetics (Meiosis)
- Evolution (Survival of the Fittest)
Combines biological ideas of evolution and genetics
- Binary code representing fish
  - 11100001
  - 11000000
  - 01111000
  - 10010111

- Fish receive rating for interaction with predators, environment, and mates. In this example brightness for mates.
  - 3.5
  - 2.5
  - 3.5
  - 3.0

- Fish selected for mating
  - 01111000
  - 11100001

- Offspring replace the weaker in the population
  - 01110001
  - 11101000

- Introduce mutations
  - 11000000
  - 10000000
Guppy Evolution: Binary Encoding

Code parameters in *Genes*

<table>
<thead>
<tr>
<th>Attractiveness (other guppies)</th>
<th>Tail</th>
<th>Attractiveness (predator)</th>
<th>Dappling</th>
<th>Temperature Tolerance</th>
<th>Disease Tolerance</th>
<th>Food Requirements (amount)</th>
<th>Feeding Requirements (time between)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop-dead gorgeous</td>
<td>Short</td>
<td>tasty</td>
<td>none</td>
<td>little</td>
<td>hardy</td>
<td>Bottomless Pit</td>
<td>Frequent Feeder</td>
</tr>
<tr>
<td>10</td>
<td>01</td>
<td>01</td>
<td>00</td>
<td>01</td>
<td>00</td>
<td>11</td>
<td>00</td>
</tr>
</tbody>
</table>

Concatenate Genes into *Chromosomes*

1001010001001100
Guppy Evolution: The Population

Assemble chromosomes into matrix

\[
\begin{bmatrix}
1001010001001100 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
1001010001001100 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
00011001101001110110011001110010 \\
0110001001001110 \\
000011101010111001010011100101110011000111001010 \\
1100011010001101 \\
0
% guppy evolution
% Bonny Haupt
% habitat 1

wts(1,:)=[1.0, 0.8, 0.0, 0.7, 0.0, 0.0, 0.0, 0.0]; % probability of mating
wts(2,:)=[0.0, 0.0, 0.5, 0.8, 0.6, 0.2, 0.9, 0.4] % probability of getting eaten

attractive = x(:,1:2); % grabs the first two genes

% likeliness to mate
% attractiveness (brightness positive)
if attractive(ind,:) == [1 1]
    adapt(ind,1)=2.0; % drop dead gorgeous
elseif attractive(ind,:) == [1 0]
    adapt(ind,1)=1.5; % very handsome
elseif attractive(ind,:) == [0 1]
    adapt(ind,1)=1.0; % passable
else % [0 0]
    adapt(ind,1)=0.5; % if the female guppies are desperate
end

f = - ( (wts(1,:)*adapt')' + 3*(wts(2,:)*adapt')' );
The fitness or cost of each chromosome is determined.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001010001001100</td>
<td>-5.4</td>
</tr>
<tr>
<td>0110001001001110</td>
<td>-4.7</td>
</tr>
<tr>
<td>0000111010101110</td>
<td>-2.1</td>
</tr>
<tr>
<td>0101001110010111</td>
<td>-4.2</td>
</tr>
<tr>
<td>0011000111001010</td>
<td>-6.7</td>
</tr>
<tr>
<td>1100011010001101</td>
<td>-7.3</td>
</tr>
<tr>
<td>0001100110100111</td>
<td>-5.9</td>
</tr>
<tr>
<td>01100110011110010</td>
<td>-1.8</td>
</tr>
</tbody>
</table>

Cost Function

\[
\begin{bmatrix}
1001010001001100 \\
0110001001001110 \\
0000111010101110 \\
0101001110010111 \\
0011000111001010 \\
1100011010001101 \\
0001100110100111 \\
01100110011110010 \\
\end{bmatrix}
= 
\begin{bmatrix}
-5.4 \\
-4.7 \\
-2.1 \\
-4.2 \\
-6.7 \\
-7.3 \\
-5.9 \\
-1.8 \\
\end{bmatrix}
\]
Rank the Chromosomes

Cost

\[
\begin{bmatrix}
1100011010001101 \\
0011000111001010 \\
0001100110100111 \\
1001010001001100 \\
0110001001001110 \\
0101001110010111 \\
0000111010101110 \\
10110011001110010
\end{bmatrix}
= 
\begin{bmatrix}
-7.3 \\
-6.7 \\
-5.9 \\
-5.4 \\
-4.7 \\
-4.2 \\
-2.1 \\
-1.8
\end{bmatrix}
\]
Natural Selection

Poorly adapted guppies (high cost) get eaten
Mating

Selection

Crossover
Generating Offspring

\[
\begin{bmatrix}
1100011011001101 \\
0011000111001010
\end{bmatrix}
= \begin{bmatrix}
-7.3 \\
-6.7
\end{bmatrix}
\]

\[
\begin{bmatrix}
1100011011001010 \\
0011000110001101
\end{bmatrix}
= \begin{bmatrix}
-7.4 \\
-3.7
\end{bmatrix}
\]
Elitism – don’t change best chromosome

Randomly mutate selected bits
GA parameters: popsize=16; mutation rate = 0.2
<table>
<thead>
<tr>
<th>Advantages of the GA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doesn't require derivative information,</td>
</tr>
<tr>
<td>Simultaneously searches from a wide sampling of the cost surface,</td>
</tr>
<tr>
<td>Optimizes with continuous or discrete parameters,</td>
</tr>
<tr>
<td>Deals with a large number of parameters,</td>
</tr>
<tr>
<td>Is well suited for parallel computers,</td>
</tr>
<tr>
<td>Optimizes parameters with extremely complex cost surfaces,</td>
</tr>
<tr>
<td>Provides a list of semi-optimum parameters, not just a single solution,</td>
</tr>
<tr>
<td>May encode the parameters so that the optimization is done with the encoded parameters, and</td>
</tr>
<tr>
<td>Works with numerically generated data, experimental data, or analytical functions.</td>
</tr>
</tbody>
</table>
Continuous Genetic Algorithm

1. Initialize Population (continuous Numbers)
2. Evaluate Cost
3. Converge?
   - yes: Solution
   - no: Mating (blend information) → Mate Selection → Mutation → Evaluate Cost → Converge?

Initialize

Parameters \( N_{par} \)

\[
\text{chromosome} = \left[ p_1, p_2, p_3, \ldots p_{N_{par}} \right]
\]

Cost Function

\[
\text{cost} = F(\text{chromosome}) = F\left[ p_1, p_2, p_3, \ldots p_{N_{par}} \right]
\]
Select crossover point \( \alpha = \text{roundup} \{ \text{random} \times N_{\text{par}} \} \)

Select parents

\[
\text{parent}_1 = \begin{bmatrix} p_{m1} & p_{m2} & \cdots & p_{m\alpha} & \cdots & p_{mN_{\text{par}}} \end{bmatrix}
\]

\[
\text{parent}_2 = \begin{bmatrix} p_{d1} & p_{d2} & \cdots & p_{d\alpha} & \cdots & p_{dN_{\text{par}}} \end{bmatrix}
\]

Perform crossover

\[
\text{offspring}_1 = p_{m\alpha} - \beta \left[ p_{m\alpha} - p_{d\alpha} \right] p_{dN_{\text{par}}}
\]

\[
\text{offspring}_2 = p_{d\alpha} - \beta \left[ p_{m\alpha} - p_{d\alpha} \right] p_{mN_{\text{par}}}
\]
For a continuous GA, merely generate a new random number in place of mutated variable.

\[
\text{chromosome} = \begin{bmatrix} p_1, p_2, p_3, p_4, \ldots p_{N\text{\,var}} \end{bmatrix}
\]

\[
\text{mutated \, chromosome} = \begin{bmatrix} p_1, p_2, p_{3\text{\,new}}, p_4, \ldots p_{N\text{\,var}} \end{bmatrix}
\]

\(p_{3\text{\,new}}\) newly generated random number
Find the minimum of:

\[ f(x, y) = \sin(x)J_1(y) \]

On the interval:

\[ 0 \leq x \leq 10 \]
\[ 0 \leq y \leq 10 \]

Minimum: \(-0.58186\)

at: 4.71, 1.81
**Bes-sin: Continuous GA**

<table>
<thead>
<tr>
<th>initial population</th>
<th>initial costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.77109</td>
<td>0.04337</td>
</tr>
<tr>
<td>0.67016</td>
<td>0.07454</td>
</tr>
<tr>
<td>0.35830</td>
<td>0.89600</td>
</tr>
<tr>
<td>0.99596</td>
<td>0.06927</td>
</tr>
<tr>
<td>0.10709</td>
<td>0.36673</td>
</tr>
<tr>
<td>0.89394</td>
<td>0.67496</td>
</tr>
<tr>
<td>0.43533</td>
<td>0.08984</td>
</tr>
<tr>
<td>0.63978</td>
<td>0.03043</td>
</tr>
<tr>
<td>0.75431</td>
<td>0.57726</td>
</tr>
<tr>
<td>0.70620</td>
<td>0.06602</td>
</tr>
<tr>
<td>0.03515</td>
<td>0.63859</td>
</tr>
<tr>
<td>0.89856</td>
<td>0.80849</td>
</tr>
</tbody>
</table>

Population size: 12
Mutation rate: 0.2
Crossover rate: 0.5
### Real GA: Bes-sin

#### Mating

<table>
<thead>
<tr>
<th>probability of mating</th>
<th>cum probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.28571</td>
<td>0.28571</td>
</tr>
<tr>
<td>0.2381</td>
<td>0.52381</td>
</tr>
<tr>
<td>0.19048</td>
<td>0.71429</td>
</tr>
<tr>
<td>0.14286</td>
<td>0.85714</td>
</tr>
<tr>
<td>0.095238</td>
<td>0.95238</td>
</tr>
<tr>
<td>0.0476191</td>
<td>1.</td>
</tr>
</tbody>
</table>

#### Cost Weighted Roulette Wheel

Roulette Wheel selection
pick1 = 0.26006
pick2 = 0.89538
cost
ma = 1 : 0.43533 0.089845 -0.37954
pa = 5 : 0.035158 0.63859 -0.063866
crossover point
xp = 2
mixing parameter
r = 0.32178
cost
offspring: 0.43533 0.26642 -0.2689
0.035158 0.46202 -0.24959

#### Mutation

<table>
<thead>
<tr>
<th>Original x</th>
<th>Original y</th>
<th>Original cost</th>
<th>Mutated x</th>
<th>Mutated y</th>
<th>Mutated cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.43533</td>
<td>0.089845</td>
<td>-0.37954</td>
<td>0.43533</td>
<td>0.089845</td>
<td>-0.37954</td>
</tr>
<tr>
<td>0.75431</td>
<td>0.57726</td>
<td>-0.29981</td>
<td>0.75431</td>
<td>0.57726</td>
<td>-0.29981</td>
</tr>
<tr>
<td>0.99596</td>
<td>0.069279</td>
<td>-0.16617</td>
<td>0.99596</td>
<td>0.070595</td>
<td>-0.16893</td>
</tr>
<tr>
<td>0.3583</td>
<td>0.896</td>
<td>-0.10674</td>
<td>0.3583</td>
<td>0.896</td>
<td>-0.10674</td>
</tr>
<tr>
<td>0.035158</td>
<td>0.63859</td>
<td>-0.06387</td>
<td>0.035158</td>
<td>0.63859</td>
<td>-0.06387</td>
</tr>
<tr>
<td>0.89394</td>
<td>0.67496</td>
<td>-0.03753</td>
<td>0.89394</td>
<td>0.67496</td>
<td>-0.03753</td>
</tr>
<tr>
<td>0.43533</td>
<td>0.26642</td>
<td>-0.42353</td>
<td>0.72802</td>
<td>0.49394</td>
<td>-0.2689</td>
</tr>
<tr>
<td>0.035158</td>
<td>0.46202</td>
<td>-0.09</td>
<td>0.8155</td>
<td>0.46202</td>
<td>-0.24959</td>
</tr>
<tr>
<td>0.035158</td>
<td>0.57588</td>
<td>-0.10909</td>
<td>0.035158</td>
<td>0.57588</td>
<td>-0.10909</td>
</tr>
<tr>
<td>0.99596</td>
<td>0.13199</td>
<td>-0.26825</td>
<td>0.03702</td>
<td>0.13199</td>
<td>0.19042</td>
</tr>
<tr>
<td>0.43533</td>
<td>0.35033</td>
<td>-0.12733</td>
<td>0.43533</td>
<td>0.35033</td>
<td>-0.12733</td>
</tr>
<tr>
<td>0.3583</td>
<td>0.63552</td>
<td>0.082746</td>
<td>0.3583</td>
<td>0.63552</td>
<td>0.082746</td>
</tr>
</tbody>
</table>
Real GA: Bes-sin Convergence

26-Jan-2006 09:14:56
continuous genetic algorithm
optimized function is funbesin_cga
popsize = 12 mutrate = 0.2 # par = 2
#generations=10
best cost=-0.5813
best solution:
0.47113    0.18935
• Problem Dependent

• In general, smallest number of cost function evaluations for:
  - small population size (order 8-16)
  - moderately large mutation rate (.15-.25)
Includes MATLAB code
(some HPF Fortran)
Collaborators:

George Young – Professor of Meteorology
Randy Haupt – ARL, Electrical Engineering
Kerrie Long, Chris Allen, Anke Beyer-Lout, Luna Rodriguez, Andrew Annunzio, Jim Limbacher, McKenzie McNeal
In case of Chemical, Biological, Radiological, or Nuclear (CBRN) release, predict:

- Release Characteristics
- Subsequent Dispersion

SCIPUFF

Effects on Humans, Equipment, Troop Mobilization

Decontamination

Medical Response
Fundamental Question: How do we predict transport and dispersion of hazardous contaminant given:

- inexact source information
- inexact meteorological conditions
- monitored data with errors
- inherent uncertainties of turbulent diffusion
- realization vs. ensemble average of models
Atmospheric Turbulence

- There is an inherent discrepancy between observed realizations and traditional ensemble model predictions.
- This discrepancy could be alleviated through careful data assimilation to produce an assimilated ensemble prediction.

a. Single realization

b. Ensemble average prediction

c. Assimilated ensemble prediction
Assimilation Process

Observations
  • Chem/bio
  • Met data

Assimilation
  • Statistical Interpolation
  • Initialization

Analysis

Forecast Model

Model Predictions

Forecast
GA-Var Assimilation Procedure

Concentration Assimilation

1. Use “guessed” wind data to predict concentration.
2. Compute difference between concentration prediction and observation.
3. Using assimilation algorithm, use difference to update the wind prediction.

Repeat until converged

dynamically assimilate one time before going on to next time
Dynamical Prediction System: \[ \frac{\partial x}{\partial t} = Mx + \eta \]

Assimilation Process:
\[ \frac{\partial x}{\partial t} = Mx + \eta + G(x^0, x^f) \]

Objectives:
1. Determine realization characteristics
2. Assimilate data into forecast

Can separate into wind and concentration equations
\[ \frac{\partial \vec{v}}{\partial t} = M_v(\vec{v})\vec{v} + \eta_v + G_v(\vec{v}^0, \vec{v}^f, C^0, C^f) \]
\[ \frac{\partial C}{\partial t} = M_C(\vec{v})C + \eta_C + G_C(\vec{v}^0, \vec{v}^f, C^0, C^f) \]
We wish to assimilate a puff in a meandering wind field to reconstruct time dependent wind by assimilating observations of dispersed contaminant concentrations.
Identical Twin Experiment

- Sinusoidally varying wind field
- Puff dispersion
- Source characteristics are known
- Seek to compute wind direction given concentration observations

Results
Comparison

Truth

Back-calculation
Comparison

Exact Solution

GA Assimilated Solution
Comparison

Wind Direction

Puff Centroid
The Shallow Water Assimilation: TusseyPuff

2-D shallow water model

Gaussian Puff model

Wind field

Concentration field

Anke Beyer-Lout
GA-Var Results

Source Strength

Sigma

Location

Time (s)

Magnitude of Location Error [m]
• Determining uncertainty measures for dispersion
• Assimilating sensor data for computing transport and dispersion
• Characterizing source and meteorological data given field sensor measurements (6 journal articles published)
Fundamental Question: How do we predict transport and dispersion of hazardous contaminant given:

- inexact source information
- inexact meteorological conditions
- inherent uncertainties of turbulent diffusion
- realization vs. ensemble average of models
- time variability of wind field
Need to Compute Source and Meteorological Parameters

- Source parameters may not be known, but required to model
- Meteorological parameters must represent local flow correctly
- Correlated parameters make a complex solution space
- Met data may not be available on site or not representative
- One way to address this is to include the Met data in the search variables

Genetic Algorithm: a robust tool to solve highly nonlinear optimization problems
The Genetic Algorithm

- starting population
- cost function
- done?
  - yes
  - local optimizer
  - no
  - mutated population
  - parents offspring
  - parents mating pool
  - natural selection
  - keep discard
Parameters required to Predict Transport and Dispersion:

- **Source parameters**
  - 2D location (x,y)
  - Height
  - Strength
  - Time of Release

- **Meteorological modeling parameters**
  - Wind direction
  - Wind speed
  - Stability class
  - Boundary Layer Depth

- **Sensor Characteristics**
Given these puff locations, where is the source?
What meteorological conditions exist?

<table>
<thead>
<tr>
<th>Grid Size</th>
<th>Grid Spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 x 2</td>
<td>8000 m</td>
</tr>
<tr>
<td>4 x 4</td>
<td>4000 m</td>
</tr>
<tr>
<td>6 x 6</td>
<td>2667 m</td>
</tr>
<tr>
<td>8 x 8</td>
<td>2000 m</td>
</tr>
<tr>
<td>16 x 16</td>
<td>1000 m</td>
</tr>
<tr>
<td><strong>32 x 32</strong></td>
<td><strong>500 m</strong></td>
</tr>
<tr>
<td>64 x 64</td>
<td>250 m</td>
</tr>
</tbody>
</table>

Kerrie Long
### Noiseless Model Results: $\theta=180^\circ$

<table>
<thead>
<tr>
<th>Grid Size</th>
<th>Found $\theta$ (°)</th>
<th>Strength (Kg s$^{-1}$)</th>
<th>$(x,y)$ (m,m)</th>
<th>Release Time (s)</th>
<th>Speed (m s$^{-1}$)</th>
<th>Cost Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual Solution</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>1.0e-3</td>
</tr>
<tr>
<td>GA Alone 2x2</td>
<td>179.12</td>
<td>2.94</td>
<td>(120,-730)</td>
<td>172</td>
<td>7.6</td>
<td>2.0e-1</td>
</tr>
<tr>
<td>Hybrid GA 2x2</td>
<td>178.10</td>
<td>5.05</td>
<td>(290,-760)</td>
<td>184</td>
<td>7.9</td>
<td>1.0e-1</td>
</tr>
<tr>
<td>GA Alone 4x4</td>
<td>179.69</td>
<td>1.30</td>
<td>(40,-40)</td>
<td>26</td>
<td>5.2</td>
<td>4.1e-3</td>
</tr>
<tr>
<td>Hybrid GA 4x4</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,-20)</td>
<td>0</td>
<td>5.0</td>
<td>2.0e-9</td>
</tr>
<tr>
<td>GA Alone 6x6</td>
<td>179.91</td>
<td>1.69</td>
<td>(10,80)</td>
<td>28</td>
<td>5.0</td>
<td>2.2e-3</td>
</tr>
<tr>
<td>Hybrid GA 6x6</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>3.2e-9</td>
</tr>
<tr>
<td>GA Alone 8x8</td>
<td>179.18</td>
<td>1.90</td>
<td>(80,170)</td>
<td>39</td>
<td>5.0</td>
<td>6.0e-3</td>
</tr>
<tr>
<td>Hybrid GA 8x8</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>3.1e-9</td>
</tr>
<tr>
<td>GA Alone 16x16</td>
<td>179.96</td>
<td>1.35</td>
<td>(0,40)</td>
<td>13</td>
<td>5.0</td>
<td>1.6e-3</td>
</tr>
<tr>
<td>Hybrid GA 16x16</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>3.4e-9</td>
</tr>
<tr>
<td>GA Alone 32x32</td>
<td>180.07</td>
<td>1.39</td>
<td>(-10,40)</td>
<td>13</td>
<td>5.0</td>
<td>1.8e-3</td>
</tr>
<tr>
<td>Hybrid GA 32x32</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>3.6e-8</td>
</tr>
<tr>
<td>GA Alone 64x64</td>
<td>179.96</td>
<td>1.45</td>
<td>(0,80)</td>
<td>18</td>
<td>5.0</td>
<td>2.1e-3</td>
</tr>
<tr>
<td>Hybrid GA 64x64</td>
<td>180.00</td>
<td>1.00</td>
<td>(0,0)</td>
<td>0</td>
<td>5.0</td>
<td>3.0e-9</td>
</tr>
</tbody>
</table>

Kerrie Long
Sensor Constraints

Detection Levels

Saturation Levels

Data Points

Concentration (kg m\(^{-3}\))

Detection Levels:
- 1.00E-04
- 1.00E-08
- 1.00E-12
- 1.00E-16

Saturation Levels (%):
- 100
- 50
- 10
- 1

Luna Rodriguez
Gaussian Puff with Thresholds

A
No Thresholds

B
Sat Level 100% & Det Level 1e-16

Sensor Location (km)
Sensor Location (km)
Skill Score Results

Detection Level at 1e-16

Detection Level at 1e-12

Detection Level at 1e-8

Detection Level at 1e-4

Luna Rodriguez
The ability to retrieve boundary layer depth depends on domain size and atmospheric stability.

Reflected contaminants will not impact surface concentration values if the domain is too small and if the atmosphere is too stable.

Inverting Turner's approximation gives the horizontal distance from the source where reflected contaminants significantly impact surface concentration values.

This distance is given by:

\[ x = \exp \left\{ \frac{1}{K} \ln \left( \frac{2 z_i - z_{source}}{2.15} \right) - \frac{I}{K} + \left( \frac{J}{2K} \right)^2 \right\} - \frac{J}{2K} \]
Six Parameter – noise 1 m release

- Retrieval with stability classes A and D very robust
- Stability Class B similar to stability class A, and stability class C similar to stability class D
- Stability classes C and D not as robust for high boundary layer depths, but retrieval still successful for most of the SNRs tested
Differences with Stability

Gaussian Plume - Stability A

Gaussian Plume - Stability D

Gaussian Plume - Stability C

Gaussian Plume - Stability F
Optimizing 3 Meteorological Parameters

- **Seek**
  - Wind speed (continuous: [0, 20])
  - Wind direction (continuous: [0, 360])
  - Stability class (integer: [1, 6])

- **Solutions close to “exact”**

Results for 10 runs of a Mixed Integer Genetic Algorithm

<table>
<thead>
<tr>
<th></th>
<th>Wind Spd (m/s)</th>
<th>Wind Dir (°)</th>
<th>Stability Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>5.000</td>
<td>180.000</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>4.990</td>
<td>180.026</td>
<td>6</td>
</tr>
<tr>
<td>Median</td>
<td>4.987</td>
<td>180.028</td>
<td>6</td>
</tr>
<tr>
<td>Stand Dev.</td>
<td>0.226</td>
<td>0.014</td>
<td>0</td>
</tr>
</tbody>
</table>
Several groups have developed algorithms for the purposes of assimilating sensor information and determining chemical or biological agent source locations and strengths.

FFT07 was designed for evaluation of such algorithms.
• The Eulerian/semi-Lagrangian (EULAG) numerical model uses an LES approach to solve the partial differential equations governing turbulent fluid flow.

• For a realistic risk assessment of contaminant releases, EULAG is coupled with a global climatology analysis tool (GCAT).

• CFD data allows us to understand the details of the flow and eliminates unknowns due to noise.

<table>
<thead>
<tr>
<th>Institution</th>
<th>NCAR/RAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>GCAT/EULAG</td>
</tr>
<tr>
<td>Sensors:</td>
<td>100</td>
</tr>
<tr>
<td>Domain:</td>
<td></td>
</tr>
<tr>
<td>Horizontal:</td>
<td>2240 m x 2240 m</td>
</tr>
<tr>
<td>Vertical:</td>
<td>810 m</td>
</tr>
<tr>
<td>Time:</td>
<td>300 s</td>
</tr>
</tbody>
</table>
FFT70 Simulated Release T + 0 seconds

Luna Rodriguez
Kerrie Long
• Second-order Closure Integrated Puff Model
  – Gaussian puff-based transport model
  – Ability to account for turbulence, terrain, and weather effects

• Inputs
  – Fixed wind field
    • Wind speed $3.5 \text{ m s}^{-1}$
    • Wind direction $293.50^\circ$
  – Single puff ‘instantaneous’ release
  – Release Mass $0.74 \text{ kg}$
Experimental Setup

Wind Direction

Sensor Network: ~100 m spacing

SCIPUFF Domain

GA Search Space

~1.7 km

~2.2 km

~5.1 km

~7.3 km

Kerrie Long
Data and Forecast

FFT70 Simulated Release T + 0 seconds

Horizontal Slice at z = 0.0m
C7F14 at Day 00 00:00:10 (10.0 sec)

Kerrie Long
Release Location & Strength as Determined by the Lowest Cost Function Value

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error in [m]</td>
<td>227</td>
<td>6</td>
</tr>
<tr>
<td>Search Space [m]</td>
<td>1700</td>
<td>2200</td>
</tr>
<tr>
<td>% Error</td>
<td>13%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>
GA is useful for environmental problems

Assimilating sensor observations can greatly improve AT&D predictability

Process involves:
1. Determine the realization characteristics
2. Assimilate the data into that current prediction

Can recover wind information given concentration observations

Produces better concentration predictions

Can recover unknown source and meteorological Parameter
Now to more detailed: Rock Springs Site
WRF-ARW Setup

- Five grid nests
  - 36 km
  - 12 km
  - 4 km
  - 1.33 km
  - 444 m
- One-way interface from coarse to fine
WRF-ARW Setup

- 43 Vertical layers
- 5 layers in lowest 10 m with 2 m spacing
- FDDA
- [http://www.meteo.psu.edu/~wrfrrt/](http://www.meteo.psu.edu/~wrfrrt/)
• Reynolds Averaged Navier Stokes
• Finite Element Flow Solver
• Second-order accurate in space and time
• Can demonstrate details of flow about structures
• Domain: 2.7Km x 2Km by 1Km
• Mesh size: 200x200x100 = 4e6 nodes
• Spatial resolution: 1.5 m in the transverse directions
  1 m near wall spacing
Case Study

Assimilate Mesoscale Model Inflow Data

Velocity Magnitude set to constant 10 m/s

Velocity obtained from Mesoscale model. Velocity Magnitude varies from 0 to 50 m/s

Constant Inflow Case

Specified Inflow Case
Case Study - Acusolve Flow Visualization With Mesoscale Assimilation

- **Constant Inflow Case**: Velocity Magnitude set to constant 10m/s
- **Specified Inflow Case**: Velocity obtained from Mesoscale model. Velocity Magnitude varies from 0 to 50 m/s
So - Where are GAs Going?

Everywhere
Questions?