4.2 APPLICATION OF A GENETIC ALGORITHM-COUPLED RECEPTOR/DISPERSION MODEL TO THE DIPOLE PRIDE 26 EXPERIMENTS

Christopher T. Allen *, Sue Ellen Haupt, George S. Young
The Pennsylvania State University, University Park, PA

1. INTRODUCTION

Because source information is not always available for airborne contaminants, source characterization is an important capability in air quality analysis. The purpose of source characterization is to determine the locations, times, and strengths of one or more emissions. Once a source is characterized, future releases from the source may be prevented, or at the very least modeled to predict their potential harm to society.

The tools currently available for this task include both forward and backward models. Output from forward-predicting transport and dispersion models can be interpreted as an expectation for the subsequent concentration field. In contrast, backwards-looking receptor models can be applied to identify the sources of measured pollutants.

The approach presented here is to couple a forward dispersion model with a backward receptor model using a genetic algorithm (GA). Previous studies have approached this problem by using a genetic algorithm (GA) to optimize the calibration factor coupling the dispersion model with the receptor model. This current work takes a more comprehensive approach using real data, using the GA to determine not just the calibration factor (i.e. source strength) but also source timing and location.

Works such as Holland (1975) and Goldberg (1989) first introduced the GA and its broad range of applications. Cartwright and Harris (1993) used genetic algorithms to calculate calibration factors in a simpler context, while Haupt (2005) demonstrated the effectiveness of this technique in a series of sensitivity studies. Haupt and Haupt (2004) discuss the many different types of GAs, of which the continuous parameter GA is most appropriate for tuning the calibration factors of a dispersion/receptor-coupled model.

Haupt, Young, and Allen (2005 – hereafter referred to as HYA) provide a summary of the continuous parameter GA used in the current study. HYA validated the concept using the Gaussian plume dispersion equation with a chemical mass balance (CMB) receptor model, using synthetic receptor data produced by the Gaussian plume equation. The model has since been upgraded by replacing the Gaussian plume equation with the much more sophisticated SCIPUFF dispersion model (Sykes, et al. 1998; Allen, et al. 2005) as the coupled model's forward component. The impact of upgrading in this way has been validated using synthetic data produced by SCIPUFF itself (Allen, et al. 2006).

The next step is to demonstrate the GA-coupled modeling approach in a real-data application. The goal is to use the model as a source characterization tool in the context of an operational dispersion model and with real data. Such a coupled model could be useful in source characterization for hazardous release events where both monitored pollutant and meteorological data are available.

Real-data runs are conducted with the coupled model using neutrally buoyant tracer concentration data from the Dipole Pride 26 field experiments. These runs gauge the model's ability to correctly characterize pollutant sources despite the stochastic scatter of realizations around the forecast ensemble mean. This work demonstrates the applicability of the coupled model concept and the ability of the genetic algorithm to diagnose calibration factors linking the dispersion model with the receptor data.

2. MODEL METHODOLOGY

The coupled model is inspired by the Chemical Mass Balance (CMB) model, which can be written as:

$$\mathbf{C}_{mn} \bullet \mathbf{S}_n = \mathbf{R}_m \qquad (1)$$

where **C** is the source concentration matrix, representing the expected contribution from each source *n* at the receptor for each observation period *m*, as computed by the dispersion model using assumed sources' emission rates; **R** is the vector denoting the measured concentration of pollutant at the receptor with one row for each time period *m*; and **S** is a vector of the unknown calibration factors linking the two. The GA is used to compute the **S** vector that provides the best fit to the data (HYA). This matrix problem is often poorly conditioned, so simple inversion techniques are typically not able to solve the problem. Haupt (2005), however, has demonstrated the GA's ability to find the solution.

As the forward component of the coupled model, SCIPUFF computes the contributions from each source to fill the elements of matrix **C** in (1). SCIPUFF is an ensemble mean dispersion model designed to compute the time dependent field of expected concentrations resulting from one or more sources. The model solves the transport equations using a second-order closure scheme, and treats plumes as a collection of Gaussian puffs (Sykes, et al. 1986; Sykes and Gabruk, 1997). SCIPUFF can be used to predict expected concentrations of emitted gases, particulates, or hazardous releases. One specific application is as an

* *Corresponding author address*: Christopher T. Allen, The Pennsylvania State University, Department of Meteorology, 503 Walker Building, University Park, PA 16802-5013; email: cta117@psu.edu

aid to health agencies in predicting dispersion of hazardous pollutants (Cox, et al. 1998). SCIPUFF is an appropriate choice for our coupled model because of its ability to compute expected concentrations over predefined time periods for any number of sources. Thus, its output can easily be integrated into (1).

3. FIELD DATA TEST METHODOLOGY

The model is applied to a real-data scenario to see how well it performs for individual realizations, in regards to the characterization of source location, emission time, and strength. The data set used here is from the Dipole Pride 26 (DP26) field experiments.

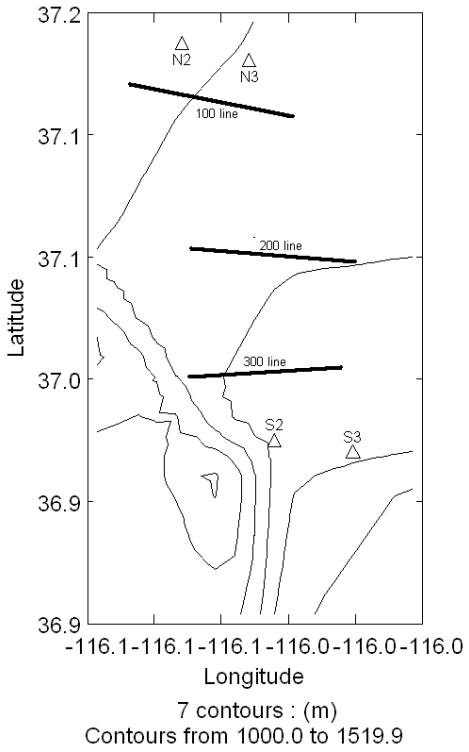

Figure 1. Dipole Pride 26 test domain as represented in coupled model. N2, N3, S2, and S3 are the emission source locations. The thick black lines represent the approximate locations of the receptors (30 along each line). The thin black lines represent terrain contours. Modeled after similar figures in Biltoft (1998) and Chang, et al (2003).

The DP26 field experiments were carried out in November 1996 at the Nevada Test Site (Biltoft 1998). Instantaneous releases of sulfur hexafluoride (SF6) at locations surrounding the receptors were made during the experiments. There were 17 different field tests carried out during the DP26 experiments, although our study only used 14 of these tests, as three were discarded due to missing data. Figure 1 shows the test domain and orientation of sources and receptors. N2, N3, S2, and S3 are the source locations, and the thick black lines show the approximate receptor locations. Further details on these field experiments can be found in Biltoft (1998) and Watson, et al. (1998).

Chang, et al. (2003) used the DP26 data to validate various dispersion models, including SCIPUFF. They found that about 50–60% of SCIPUFF-predicted concentrations came within a factor of 2 of the observations. Most large errors occurred when the modeled puff missed the receptors altogether due to errors in the interpolated wind field. Still, SCIPUFF performed as well as any of the other dispersion models validated by Chang, et al (2003).

As part of the GA-coupled model, we are also using SCIPUFF to model the DP26 experiments. Therefore, we should expect the same discrepancies between expected concentrations and monitored values as found by Chang, et al (2003). Errors in results obtained by the coupled model can perhaps be attributed to these discrepancies. These errors are due in part to the difference between stochastic realizations and ensemble mean predictions. Chang, et al. (2003) attributed the differences mainly to the wind field interpolated by SCIPUFF. The GA-coupled model is designed to help account for these issues by calibrating the dispersion model output to match the receptor data. Also, a future coupled model could tune the wind field to ensure that the modeled puff reaches the receptors.

To use the DP26 data, certain adaptations are required within the coupled model. The main difference between this application and the validation with synthetic data in HYA and Allen, et al. (2006) is that data are now available from multiple receptors, as DP26 includes data from 90 independent receptors during each field test. We should use as much data as possible to maximize accuracy; therefore the overall methodology must be changed (particularly, the cost function) in order to incorporate all of the available data. To do this, the **C** and **R** matrices in (1) were expanded in an extra dimension corresponding to the receptors. The generalized matrix equation is now:

$$\mathbf{C}_{mnr} \bullet \mathbf{S}_n = \mathbf{R}_{mr} \tag{2}$$

where *r* indexes the receptors. As before, *m* indexes the observation periods, and *n* indexes the sources.

The calibration vector **S** is not expanded in this extra dimension, since the goal of the GA is to find a single **S** providing the best fit across all receptors. If the model matches the data perfectly, a single **S** vector would provide an exact fit for all receptors. The basic cost function is the same as in HYA, but modified to sum over all receptors:

$$RMS = \frac{\sum_{r=1}^{90} \sqrt{\sum_{m=1}^{M} (\mathbf{C}_{mnr} \bullet \mathbf{S}_n - \mathbf{R}_{mr})^2}}{\sum_{r=1}^{90} \sqrt{\sum_{m=1}^{M} (\mathbf{R}_{mr})^2}} \tag{3}$$

The denominator represents a normalization scheme elaborated on in Section 4.2.

Several initial runs are made with the coupled model using the DP26 data. The goal of these runs is to characterize the emission locations and times (strength

characterization is the focus of subsequent sections). These runs use the four emission locations (N2, N3, S2, S3) at two times each, for a total of eight sources. Since each field test emitted from only one or two sources, **S** should be equal to 1 at the emitting sources, and 0 for all non-emitters, if all else is perfect. In other words, it should detect which source was the actual emitter for each field test by comparing receptor data with SCIPUFF-predicted contributions from each potential source. **S** values are not restricted to only 0 or 1, but a continuous range, initially set at 0 to 5.

In these initial runs, the correct source and time of emission were identified 64% of the time. In obtaining these results, two changes to the coupled model infrastructure were made to improve performance. One issue is that the model did not characterize sources as non-emitters whose potential plumes dispersed completely outside the receptor domain. To fix this issue, each column in the **C** matrix, representing the pollutant contribution of each source *n*, is summed across all time periods and receptors. These totals are normalized by the maximum contribution from any source *n* to produce a scale factor ranging from 0 to 1:

$$\text{scale}(n) = \frac{\sum_{r=1}^{90}\sum_{m=1}^{M} \mathbf{C}_{mnr}}{\max\left(\sum_{r=1}^{90}\sum_{m=1}^{M} \mathbf{C}_{mnr}\right)} \qquad (4)$$

The scale factor is multiplied by a prespecified upper limit to give the maximum source strength allowed by the GA for each source. Sources that cannot emit into the domain result in scale factors of 0, forcing the GA to limit these sources' $\mathbf{S}_n$ values to 0. This method does not assume any prior knowledge regarding which sources are potential emitters, but does provide objective estimates of each source's potential contribution to the domain.

The range of values allowed by the GA should be increased to counter the inclusion of the scale factor, which may act to narrow the range for potentially correct sources. The range should not be made too large, however, since the run-to-run variability in solutions is proportional to this range. This is elaborated on in Section 4.8.

In addition to the scale factor implementation, the cost function was changed to de-emphasize pollutant magnitudes. Instead of squares of differences, the model uses the base-10 logarithm of the squares of differences:

$$RMS = \frac{\sum_{r=1}^{90}\sqrt{\sum_{m=1}^{M}\log_{10}\left((\mathbf{C}_{mnr}\bullet\mathbf{S}_n - \mathbf{R}_{mr})^2\right)}}{\sum_{r=1}^{90}\sqrt{\sum_{m=1}^{M}\log_{10}\left((\mathbf{R}_{mr})^2\right)}} \qquad (5)$$

causing the cost function value for a particular S vector to depend less on magnitudes and more on whether or not a source is emitting at all. With a logarithmic cost function, overestimations in expected concentration, as often found with this data set, do not so greatly impact the model's performance. For example, if the receptor data value is 500, but the model's predicted concentration is 5000, a logarithmic cost function identifies the value of 5000 as more optimal than a value of 0. This is desired because the main issue in source identification is not strictly the magnitude, but whether a particular source is emitting at all.

4. PERFORMANCE OPTIMIZATION

Now we seek to optimize the performance of the coupled model with the DP26 data set by performing various tests, each designed to determine the impact of tuning a particular parameter value. While the optimization is specific to DP26, many of the general results can be applied to the coupled model for other data sets. The goal is to use the results from the sensitivity studies to produce an automated process that best characterizes the location, time, and strength of the pollutant source(s).

4.1 *GA vs. Random search*

The first test determines if solving the matrix problem requires the GA at all, or if finding the solution by brute force is equally efficient. The GA's performance is compared to the performance of a random search method, which produced random **S** values and evaluated them with the same cost function.

Figure 2 shows the minimum cost for one of the DP26 tests, as found by the GA (dashed) and the random search (solid), taken as an average over 5 runs for 20,000 iterations. Clearly the random search took much longer to find a low cost function value. In fact, out to 20,000 iterations, the random search never caught up to the GA. This shows that a random search is inefficient, and that more sophisticated optimization methods such as a GA are required. HYA and Haupt (2005) tested other methods such as matrix inversion, but since the matrix problem is often poorly conditioned, these techniques did not perform well.

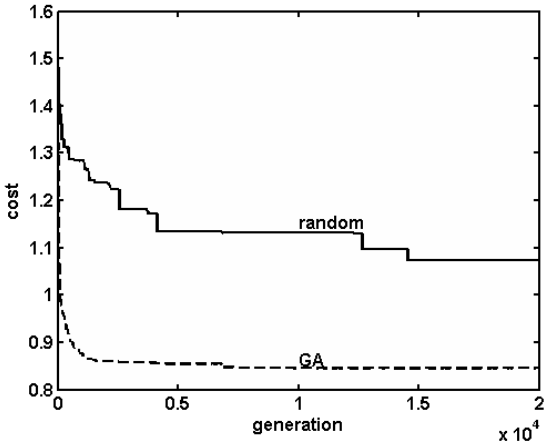

Figure 2. Minimum cost function value as a function of iteration number for the GA (dashed) versus a random search method (solid), carried out to 20,000 iterations.

4.2 *Cost functions*

As mentioned earlier, the cost function was modified to consider logarithms of squared differences, resulting in enhanced performance. We now extend this study with four different cost functions. Each of the four cost functions is either logarithmic or non-logarithmic, and uses one of two normalization schemes.

Normalization can occur before or after summing each receptor's contribution to the cost function. Thus far, the coupled model has weighted each receptor's cost function value proportional to its residual, as in (5), by normalizing after summing the contributions from each receptor. Another option is to weight each receptor equally by normalizing the contributions from each receptor prior to the summation. In terms of (5), this means removing the summation operators from the numerator and denominator and placing a new summation operator in front of the entire ratio:

$$RMS = \sum_{r=1}^{90} \frac{\sqrt{\sum_{m=1}^{M} \log_{10}\left((\mathbf{C}_{mnr} \bullet \mathbf{S}_n - \mathbf{R}_{mr})^2\right)}}{\sqrt{\sum_{m=1}^{M} \log_{10}\left((\mathbf{R}_{mr})^2\right)}} \quad (6)$$

HYA showed that a square is the optimal power of difference to take in the cost function. Therefore, all cost functions considered here involve some form of a squared difference.

The coupled model is run with data from each DP26 field test for each cost function. Since the run-to-run variability in solutions was quite small, long Monte Carlo simulations are not necessary. This also applies to the sensitivity studies that follow.

Success scores are given based on the ability of the GA to correctly characterize the source. For each field test, scores from 1 (worst) to 4 (best) are awarded based on the GA's performance, with separate scores given for location, emission time, and strength characterization. These success scores are then added up across all field tests to produce a total success score for location, time, and strength characterization, as well as overall performance (the sum of those three). The higher the success score, the better the relative performance.

Table 1 shows the success scores for each cost function. While the choice of cost function did not affect the ability of the GA to identify the correct release time, the logarithmic, post-normalization cost function (5) used up to this point scored highest in identifying the release location. Using this cost function adds 25% to the runtime when compared to a non-logarithmic cost function, but the improved performance in correct source identification is worth the extra computational cost, as the logarithmic cost function correctly identified 3 more sources than did the non-logarithmic cost function.

Cost function	Source ID	Time ID	Strength	Total
Log, post-norm	29	34	29	92
Log, pre-norm	27	34	26.5	87.5
No log, post-norm	24	33	25.5	82.5
No log, pre-norm	15	33	18.5	66.5

Table 1. Success scores given to four cost functions in identifying the correct emission source, time, and strength, and the sum of those three, across the 14 Dipole Pride 26 field tests. A higher score indicates more success. The cost functions include either a logarithmic squared difference or just a squared difference, and either of two normalization schemes.

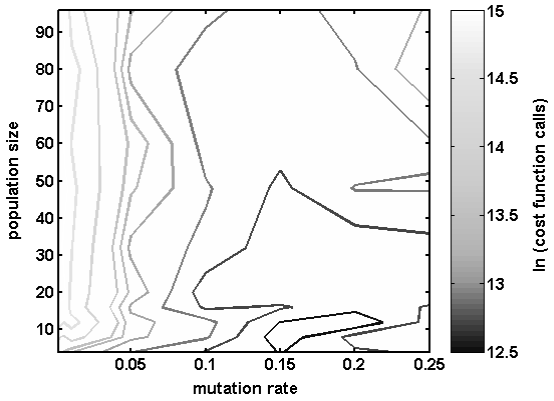

Figure 3. Contour plot of average number of cost function calls versus mutation rate and population size for the coupled model. Darker contours correspond to fewer cost function calls. The number of cost function calls has been normalized by the natural log for viewing purposes.

4.3 *Population sizes and mutation rates*

Haupt (2005) performed a sensitivity study on GA population sizes and mutation rates using synthetic data and found that two combinations of sizes and rates were optimal: high population sizes coupled with low mutation rates, and low population sizes coupled with high mutation rates. A similar sensitivity study is made using the DP26 data set using 5 of the 14 field tests. The goal is to find an optimal combination of population size and mutation rate in terms of a minimum number of cost function evaluations, and therefore minimum computing time.

Figure 3 shows the number of cost function calls for 80 combinations of population sizes and mutation rates, averaged across 5 runs for each field test. The optimal mutation rate was 0.15, and the optimal population size was in the range of 4 to 12. Previous DP26 runs used 0.20 and 8, respectively, so we were not far off the optimal case already. The values used do not affect the

ability of the GA to find the optimal solution, only the number of iterations needed to obtain it; therefore there is no need to re-run previous tests with the optimal mutation rate.

4.4 *Output intervals*

The DP26 data set provides receptor data every 15 minutes. SCIPUFF can also output values every 15 minutes; however, unlike the SCIPUFF output, the DP26 receptor data are not instantaneous concentrations every 15 minutes, but rather time averages over 15-minute periods. Shortening the output interval in SCIPUFF may improve the GA's performance by producing a better match to the 15-minute averages.

Four output intervals are tested: 15, 5, 3, and 1.5 minutes. Expected concentrations produced by SCIPUFF are averaged into 15-minute time periods for inclusion into (2).

Table 2 shows the success scores for this test. Source identification skill is not affected by the output interval, but the 5-minute interval resulted in the best performance for time and strength identification. It is surprising that smaller output intervals did not yield better performance, as shortening the output interval beyond 5 minutes did not appreciably improve accuracy.

Timestep (minutes)	Source ID	Time ID	Strength	Total
15	21	18	12.5	51.5
5	23	21.5	12.5	57
3	22	25.5	18.5	66
1.5	23	23.5	12.5	59

Table 2. Scores given to four timesteps in identifying the correct emission source, time, and strength, and the sum of those three, across 7 of the 14 Dipole Pride 26 field tests. A higher score indicates more success. Receptor data is given every 15 minutes, so smaller timesteps averaged several expected concentrations within each larger 15-minute timestep.

4.5 *Receptor sets*

The DP26 data set includes data for 90 different receptors. Thus far the model has used all 90 at the expense of computing time. The cost function evaluations take less time when only using one receptor, decreasing computing time by a factor of 14. Because of this large discrepancy, it is useful to investigate whether or not using a subset of the receptors results in similar (or even better) performance.

Six different receptor subsets are tested: all 90, the 200 line (see Figure 1), the middle receptor (referred to as #215), the highest-magnitude receptor (in terms of measured concentration), the highest-magnitude receptor on the 200 line (Figure 1), and the highest-magnitude range of 5 adjacent receptors. The latter three are unique for each field test and were determined by objectively analyzing the data.

Table 3 shows the success scores. The best source and time identifications are found when including all 90 receptors. However, the field-test-specific highest-magnitude receptor sets result in the best strength identification. A possible reason is that many receptors recorded concentrations near 0 for the majority of sampling periods during the field experiments, and including these receptors in the analysis lowers the GA-calculated strength.

While using a subset of the receptors within the centerline of the plume gives a more exact strength characterization, it is not always best for source identification. The modeled plume may be slightly off-center of the actual, and thus outside this subset of receptors. Subsequent analyses use all 90 receptors in order to optimize source and time identification.

Receptor set	Source ID	Time ID	Strength	Total
All 90	44	57	18.5	119.5
200 line	41	56	22.5	119.5
#215	29	48	19	96
Highest single	42	49	29	120
Highest in 200 line	31	50	25	106
Highest range of 5	44	51	28.5	123.5

Table 3. Scores given to six receptor sets in identifying the correct emission source, time, and strength, and the sum of those three columns, across the 14 Dipole Pride 26 field tests. A higher score indicates more success. The receptor sets include "All 90" (all receptors included), "200 line" (Figure 3), "#215" (the middle receptor in the 200 line), "Highest single" (highest-magnitude receptor with respect to its values), "Highest in 200 line" (same as highest single, but restricted to the 200 row in Figure 3), and "Highest range of 5" (highest-magnitude range of 5 adjacent receptors' values).

4.6 *Specific source and time characterization*

So far, the coupled model's source array has only included sources N2, N3, S2, and S3 (Figure 1) at two times each within each field experiment. We now use a source array that does not assume a particular release location or time by using a grid of locations and multiple times. This study is split into three parts.

The first part uses a .06°-by-.06° grid of points with .02° resolution as the source array. The grid is located on the side of the domain where the release was known to originate. Each source specifies the previously known correct release time, so that the coupled model is only characterizing the location within the grid. Since there are 16 sources and 14 sampling periods, this appears to violate the requirement of the matrix problem that the number of sources be less than the number of sampling periods; however, if all 90 receptors are used, there are effectively 14×90 sampling periods. The coupled model was able to pinpoint the correct emission grid location for 6 of 14 field tests, while an additional 3 field tests came within one grid point.

The second part of the study considers 16 different emission times separated by 10 minutes each. This portion of the study only characterizes the release time, as each source specifies the previously known correct location. The release time was pinpointed exactly for 9 of 19 emissions made in the 14 field tests, while for 14 of 19 emissions, the GA-produced time was within 30 minutes of the actual. Some of the field tests had major issues related to missing or faulty data (both receptor and surface meteorological data); model runs invoking these field tests have performed poorly in every study thus far, but are still included in the overall analysis.

The first two parts of this study each use some previously known information, either the location or the time of emission. The final portion of the study did not assume either, but attempted to pinpoint the location and time concurrently, using four locations on the corners of a .04°-by-.04° grid, and four emission times separated by 40 minutes each.

The coupled model was, however, unable to diagnose upwind location and release time simultaneously. Figure 4 provides an illustration, with a receptor X and two sources A and B each upwind of the receptor. Suppose source A emits at time 0, and source B emits at time 3. Since both puffs would reach the receptor X at time 6, the coupled model would be unable to apportion the correct amount of pollutant to each source because the puffs have merged. Source A may be the correct emitter in this theoretical example, but since the matrix problem is rank deficient, the GA is likely to attribute some random combination of pollutant to both A and B.

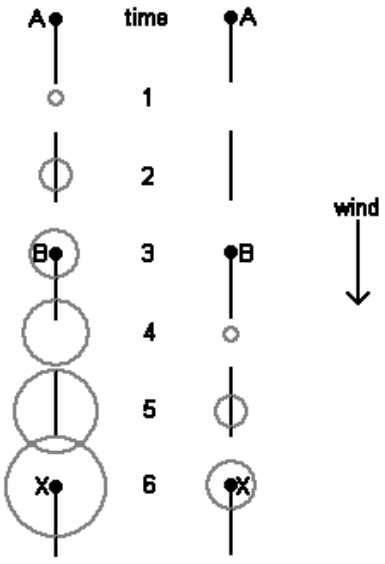

Figure 4. Demonstration of the coupled model's difficulty in diagnosing upwind location and release time simultaneously. If source A emits a puff at time 0 and source B emits at time 3, each puff reaches receptor X at time 6, and the coupled model would be unable to properly attribute the correct amount of pollutant to each source.

Using multiple receptors in the crosswind direction does not solve this problem because there is little (if any) separation between each source's puff. Thus, other receptors off the centerline in Figure 4 have the same problem as those along the centerline. Even though the puffs from sources A and B are of different sizes, the model still transports some of each source's puff to off-center receptors, and the same problem exists. The model performs better if only a single row of potential sources inside the correct range is included in the candidate source array. Such a source array provides the basis for the first stage of the multi-stage process described in Section 5.

4.7 *Source strength characterization*

In an effort to improve source strength characterization, the next set of coupled model runs takes the locations and release times found above as given and reruns the coupled model to tune only the source strength. The best-fit source locations and times were used instead of the actual documented emissions in order to minimize the impact of SCIPUFF's dispersion and transport errors on the source strength analysis. The coupled model has thus far underestimated the strengths when incorporating all 90 receptors, but performance should improve when the coupled model is forced to attribute all of the pollution to only one or two sources. Different receptor sets are re-tested to see if these results agree with the analysis in Section 4.5.

Unlike the case for a larger source array, using all 90 receptors with just the correct sources produces a better strength characterization than the other receptor sets tested (highest single receptor, and highest adjacent range of 5). In cases where SCIPUFF's centerline does not match the data, SCIPUFF's predicted concentrations within smaller receptor sets (which are within the actual puff, not the modeled puff) are low and the GA compensates by greatly over-approximating the source strength. This does not happen in larger receptor sets because all receptors are included in the analysis, including those the modeled puff passes through. Therefore, receptors that the modeled puff envelops, but not the actual puff, act to decrease the strength characterization to a more realistic value.

Using all receptors, source strengths came within an order of magnitude of the actual strength reported by Biltoft (1998) for only 9 of 19 emission scenarios. While performing well in regards to location and emission time, the coupled model's strength specification has thus far not been as precise.

4.8 *Other analyses*

Other sensitivity studies are performed, either involving new model runs or additional analysis of previous studies. One such study determines the effect of upper-air data and inclusion of additional wind profiles on coupled model performance. It is found that changes in upper-air data have little effect on the source strength characterizations, and no effect on locations or times.

Thus, errors in source characterizations cannot be blamed on insufficient upper-air data, as perturbations in the upper-air data or additions of new data were found to have little to no impact on the solutions. This suggests that the plume centerline errors are the result of flows not resolved by the DP26 observations.

Another study analyzes the coupled model's performance for the seven field tests with multiple emissions. For these field tests, the coupled model is able to characterize both releases correctly only 42% of the time (3 out of 7). Additional analysis is done for the other seven field tests which have only one emission, for which the coupled model produces a high-confidence single-emission solution (i.e. the solution did not indicate there were multiple emissions) for 5 of 7 field tests. Because of the issue discussed earlier regarding the possibility of puffs within a multiple-release field test missing the receptor domain, it is not surprising that less than half of the two-emission tests result in a correct characterization of each.

As mentioned earlier, the run-to-run variability in solutions is proportional to the range of values allowed by the GA. This range corresponds to the possible source strengths. To obtain the most precision, this range should be minimized while still encompassing all possible strengths. However, the range must be increased beyond the presupposed range of possible strengths because the scale factor (4) effectively shrinks this range. A study is performed to determine how much this range must be expanded to compensate for the scale factor.

Some correct sources are found to have scale factors as low as 0.1. Thus, to ensure that the range of actual strengths is realized, the maximum allowable source strength must be an order of magnitude larger than the presupposed maximum possible strength. Increasing the range beyond one order of magnitude adds to the run-to-run variability in solutions while not including any additional solutions.

5. MULTI-STAGE PROCESS

Because different candidate source arrays work better for characterizing source location, time, or strength, a combination of methods offers possible improvement in at least some of the characterizations. The goal is to combine the best features of the runs described above to produce a multi-stage process involving multiple model runs to progressively determine the correct source, time, and strength characterization. This multi-stage process goes as follows:

The first stage starts with a coarse grid designed to yield an initial estimate of how many different emission sources are present, and when they emitted. We start with 32 sources – 8 locations at 4 times. The locations are shown in Figure 5. The coupled model calculates how much pollutant should be attributed to each of the four times, and if it is an appreciable amount (within 1/3 of the maximum, since 1/3 was found to be a good benchmark in previous analyses), the time is included in the next stage. This stage allows for a maximum of four emissions (one at each of the four times).

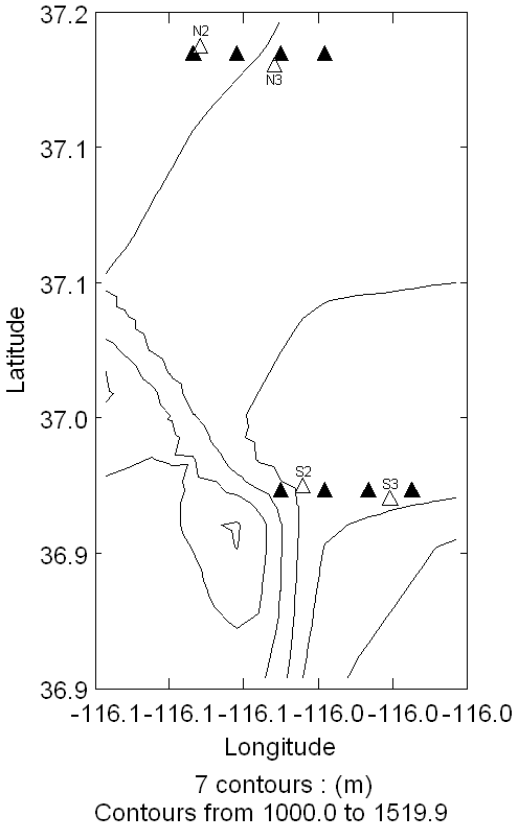

Figure 5. Source configuration for the coupled model consisting of both location and time uncertainty. The filled-in triangles represent the locations tested by this setup, and the hollow triangles represent the actual emission locations. Each source location was included in the source array for each of four release times separated by 40 minutes.

The second stage performs a separate coupled model run for each time output from the first iteration. The goal is to find each time period's most probable location using a .06°-by-.06° grid with .02° resolution (as done before). The location returning the maximum **S** value at each time is then carried to the next stage.

The third stage runs the coupled model once with the locations found in the second stage in the source array, using release times separated by 10 minutes to refine the release time characterization. The goal is to pinpoint the most probable time for each release by comparing the **S** values.

The final stage performs one more model run with the location/time combinations found in the third stage to specify the source strength. After the GA determines the best-fit **S** vector, we filter out any source whose strength was found to be less than 1/3 of the maximum for any source. If any sources are eliminated, this step is repeated without those sources. The final result is a list of emission locations, times, and strengths.

This four-stage process was run once on each of the 14 field tests to mimic a real-world attempt to

characterize the locations, times, strengths, and number of emissions. Source characterization for one of the field tests was perfect in finding the correct location, time, and number of emissions, while the strength characterization for this field test was a factor of 2 too low. Locations determined for other field tests were typically close, while the release times were generally within 40 minutes of the actual. While the majority of strength characterizations were still underestimated, they were closer than in previous studies. Since the final stage only considers the "best fit" sources found by the previous stages, these "best fit" sources resulted in higher strength specifications.

The model determined the correct number of sources in 6 of 14 field tests, while overestimating the number of sources on 5 of the other 8 field tests. The remaining three field tests each involved receptor data where one puff likely missed the receptor domain, and thus resulted in underestimations of the number of sources.

To improve the performance of the multi-stage process, the factor of 1/3 used to filter out non-emitting sources could be optimized further; however, the specific value is subjective and may not directly apply to other data sets. We considered whether the last stage might benefit from removing the logarithm from the cost function (5), but found that this change did not improve strength characterization.

This process could surely be fine-tuned, but as used here shows the real-world applicability of this coupled model towards source characterization. The process can be applied to other data sets, although specifics such as the 1/3 filtering threshold would need to be re-optimized and likely changed. When the coupled model is applied to other data sets, it is expected that prior information regarding some aspects of the sources would improve the ability of the coupled model to characterize their remaining aspects. Examples of useful prior information include a range of possible sources, times, and strengths, as used in this study. It is not currently known exactly how much uncertainty the coupled model can cope with in determining the correct source characterization. The results from Allen, et al. (2006) and HYA, however, suggest that if noise is on the same order of magnitude as the signal, correct characterization is still possible.

6. CONCLUSIONS AND RECOMMENDATIONS

Validation of the GA-coupled model using data from the Dipole Pride 26 field experiments demonstrates the model's potential applications. The coupled model performs well in identifying the correct emission locations and times, to as good a degree of accuracy as can be expected given the disparity between the ensemble nature of the dispersion model and real data resulting from a single realization of turbulent flows. The model struggles somewhat when working with multiple-emission field tests, as one emission may have missed the domain and therefore not been detected by the receptors. This also complicates the coupled model's strength characterization.

Multiple GA runs can be used with different candidate source arrays to produce the best location/time/strength characterization. The multi-stage process described in Section 5 can in theory be tuned to apply to other data sets. The specifics of the multi-stage process are project-dependent, but the multi-stage process as a whole should apply to any data set or situation. If general information regarding the nature of the sources is known (e.g. an idea of when, where, how many, and how much), the coupled model is more robust in finding a solution. This is partly because the coupled model cannot properly differentiate between some groups of parameters, such as upwind range to the source and time since release.

The coupled model can be applied as a tool for identifying approximate source locations and times in real-world scenarios. Given a set of receptor data, one could apply the coupled model in identifying the possible emission source(s). For the Dipole Pride 26 data set, the coupled model was able to pinpoint the source and time of emission to a good approximation for a majority of the field tests, despite the discrepancies between the data and SCIPUFF-predicted output as discussed by Chang, et al (2003). Thus, one can expect similar success with other data sets.

One advantage of the Dipole Pride data set is the availability of meteorological surface data every 15 minutes at several points within the domain. Surface data of this resolution in time and space is seldom available in other applications. A lack of sufficient meteorological data is likely to hinder the coupled model's performance in source characterization. It is recommended that the coupled model be used to help locate the emission source(s), given sufficient meteorological data and at least a rough initial estimate of the location and time of emission. Work with the coupled model is ongoing, including plans to use receptor data generated by more complex fluid dynamics models that are more compatible with SCIPUFF's ensemble-mean predictions, and to use the GA approach to find errors in meteorological data.

7. ACKNOWLEDGEMENTS

This work was supported by internal research and development funds from the PSU Applied Research Laboratory. We thank Joseph Chang for providing us the Dipole Pride 26 data, and Ian Sykes for helpful information on SCIPUFF.

8. REFERENCES

Allen, C. T., S. E. Haupt, and G. S. Young, 2005: Validation of a Genetic Algorithm-Coupled Receptor/Dispersion Model Incorporating SCIPUFF. Presented at the 9th Annual George Mason University Conf. on Atmospheric Transport and Dispersion Modeling, Fairfax, VA, July 2005.

Allen, C. T., S. E. Haupt, and G. S. Young, 2006: Source Characterization With a Genetic Algorithm-Coupled Receptor/Dispersion Model Incorporating SCIPUFF, to be submitted to *J. Appl. Meteor.*

Biltoft, C. A., 1998: Dipole Pride 26: Phase II of Defense Special Weapons Agency transport and dispersion model validation. DPG Doc. DPG-FR-98-001, prepared for Defense Threat Reduction Agency by Meteorology and Obscurants Divisions, West Desert Test Center, U.S. Army Dugway Proving Ground, Dugway, UT, 76 pp.

Cartwright, H. M., and S. P. Harris, 1993: Analysis of the distribution of airborne pollution using GAs. *Atmos Environ.,* **27A**, 1783-1791.

Chang, J.C., P. Franzese, K. Chayantrakom, and S. R. Hanna, 2003: Evaluations of CALPUFF, HPAC, and VLSTRACK with Two Mesoscale Field Datasets. *J. Appl. Meteor.*, **42**, 453-466.

Cox, R. M., J. Sontowski, R. N. Fry, C. M. Dougherty, and T. J. Smith, 1998: Wind and Diffusion Modeling for Complex Terrain. *J. Appl. Meteor.*, **37**, 996-1009

Goldberg, D.E., 1989: *Genetic Algorithms in Search, Optimization, and Machine Learning*, New York: Addison-Wesley.

Haupt, 2005: A Demonstration of Coupled Receptor/Dispersion Modeling with a Genetic Algorithm, submitted to *Atmospheric Environment*.

Haupt, R. L. and S. E. Haupt, 2004: *Practical Genetic Algorithms, 2nd edition with CD.* John Wiley & Sons, New York, NY, 255 pp.

Haupt, S. E., G. S. Young, and C. T. Allen, 2005: Validation of a Receptor/Dispersion Model Coupled with a Genetic Algorithm Using Synthetic Data, submitted to *J. Appl. Meteor.*

Holland, J.H., 1975: *Adaptation in Natural and Artificial Systems,* Ann Arbor: The University of Michigan Press.

Pasquill, F., 1961: The estimation of the dispersion of windborne material. *Meteor. Mag.*, **90**, 33–49.

Sykes, R. I., W. S. Lewellen, and S. F. Parker, 1986: A Gaussian plume model of atmospheric dispersion based on second-order closure. *J. Appl. Meteor.,* **25,** 322–331.

Sykes, R. I., R. S. Gabruk, 1997: A Second-Order Closure Model for the Effect of Averaging Time on Turbulent Plume Dispersion. *J. Appl. Meteor.,* **36**, 1038-1045.

Sykes, R. I., S. F. Parker, D. S. Henn, C. P. Cerasoli, and L. P. Santos, 1998: PC-SCIPUFF version 1.2PD technical documentation. ARAP Rep. 718, Titan Research and Technology Division, Titan Corp., Princeton, NJ, 172 pp.

Watson, T. B., R. E. Keislar, B. Reese, D. H. George, and C. A. Biltoft, 1998: The Defense Special Weapons Agency Dipole Pride 26 field experiment. NOAA Air Resources Laboratory Tech. Memo. ERL ARL-225, 90 pp.