

Determining the Effectiveness of the Superensemble for Atmospheric NO_2 Concentration Prediction

Frank Gronwald*, Shouu-Yuh Chang

North Carolina A & T University

*Corresponding author: fsgronwa@aggies.ncat.edu

Received July 01, 2024; Revised August 02, 2024; Accepted August 09, 2024

Abstract The need for government regulators and industrialized sources to determine the level of air pollution is essential. Air dispersion models are often used to determine the concentration of a pollutant. However changing conditions and several assumptions made by the models limit their accuracy at various times. This research was performed by combining four different air dispersion models (Gaussian Plume, Variable K Theory, Box, and AFTOX) into a superensemble. Since the superensemble is typically more accurate than its member models, the calculated result should be a more accurate prediction under any condition. One of the key parameters in the formation of the superensemble is whether the superensemble calculations for that range are to be fixed or continuous throughout. In the interest of evaluating performance, the change in accuracy for each member model and superensemble was measured through determining RMSE.

Keywords: Superensemble, Air Dispersion Modeling, NO_2 , Brier Score, RMSE

Cite This Article: Frank Gronwald, and Shouu-Yuh Chang, "Determining the Effectiveness of the Superensemble for Atmospheric NO_2 Concentration Prediction." *Journal of Atmospheric Pollution*, vol. 10, no. 1 (2024): 1-8. doi: 10.12691/jap-10-1-1.

1. Introduction

One of the drawbacks of using any air dispersion model is the amount of error in the estimation of the pollutant concentration that gets produced. This error mostly arises from the inherent difficulty of using a relatively simple modeling framework to determine the concentration profile which is influenced by very complex and variable stimuli [1]. In the process of calculating the concentration of a pollutant there are a multitude of opportunities for error to become a factor. Some of the more prominent sources of error for a particular model include: the development of mathematical formula used by the model, the quality of the initial data parameters used in the calculation, the number of pollutant sources, and the complexity of the atmosphere [2].

Another prominent source of error is commonly found during the gathering and measurement of air samples. The concentration data obtained from air samples obtained at the same testing site but during different time periods could vary widely. The sampled data could differ by 10 or even 100 times as much due to changes in the wind direction, atmospheric stability, emission level, reactivity of the pollutant, or the pollutant existing in two or more phases [3].

However, there are also many sources of error that end up being too costly or impractical to obtain either full or partial estimations. The error sources include: missing some or all of the data necessary for calculating the

magnitude of imprecision from the measurements, having the available data be insufficient, or not knowing how exactly it is related to the uncertainty calculation [4]. One example of this type of error may stem from the concentration of nitrogen. The highest being at the fuel source inside the smokestack. Since the fuel source typically is not fully heterogeneously mixed, the concentrations will often vary at various times [5]. These nitrogen concentrations may also be influenced through the various manufacturing processes being run. Human error and unpredicted events are also sources of unknown error. The main corrective solution to minimize the error is to create adjustments in the model calculations.

2. Superensemble Algorithm

By calibrating the model against known concentrations, it is hoped that the error may be minimized. However, it is well known that some bias may be introduced within the process [5]. A possible workaround is to use several models with the objective that some of the residual error will cancel itself out [6]. The superensemble has the capability to provide significant error reduction through its design. This is accomplished by using a training phase where the forecasts of the various models can be analytically reconstituted through applying multiple linear regressions [7]. This means that the error, bias, and skill are implicitly included and optimized into the superensemble forecast in the prediction phase.

Each air model has a weakness that hinders their

accuracy under some condition or time of day. Thus, it is theorized that it may be beneficial to combine the various air dispersion models. This research study attempted to do this using a superensemble. The superensemble uses the first part of each model calculation and compares it to the actual conditions. This helps it to choose which model result to give the most weight in the calculation of the superensemble result [8]. It is hoped that through reducing the influence of the less accurate models and giving more weight to the best models, the final prediction will be more accurate.

The superensemble was created due to the need for enhancement of the capabilities in ensemble mean forecasting [9]. An ensemble mean is essentially just a straight average of all its member models. All ensemble mean members are given the same weight of one. The superensemble is an extension of bias-corrected ensemble mean forecasting which removes the bias of each member model. The superensemble assigns higher and lower weights to account for biases and better or worse model performances [10]. The weights used are found through a multiple linear regression method which regresses the member model forecasts against the observed state. This creates a reward system which gives the better performing models a higher weight [11]. The only drawback to this method is that a model's previous forecasts do not always predict how accurate the model will be in the current period especially if the data has changed [12].

Because not all models are equally reliable at different points in space and time, it can be expected that the probability associated with the resulting deterministic forecast would not be the same for both the superensemble and the multi-model ensemble [13]. The probability of a forecasted event from an ensemble system is based on the fraction of ensemble members predicting the event. After finding the probability using the superensemble model, a comparison of the two forecasts should be made [14].

The superensemble was first developed to improve the Atlantic tropical cyclone forecasts [15]. It was designed to use previous forecasts from a pool of member models to correct the biases in their forecasts and rank each model's relative strength. By using this method, the new tropical cyclone predictions were more accurate than any of the models in the pool [16]. The superensemble method was next applied to climate forecasting and global precipitation forecasting [17]. The Briar Skills Score was introduced to the superensemble process to assist in the process of increasing forecast accuracy [18].

Up to this point, the superensemble had been used to predict hurricanes in the Atlantic. Other applications began to emerge, starting with the adaptation for the prediction of the surface temperature in several US regions [19]. The superensemble was also adapted to predict tropical storms and typhoons in regions such as the Pacific [20], Korea [21], Southeast Asia [22], South America [23], and China [24]. Other applications developed for the superensemble application include prediction of precipitation [25,26], and [27]; evaluating planetary boundary schemes [28]; determining the ocean surface temperature [29]; and modeling the diurnal cycle [30] and [31].

[32] first came up with the idea of using several of the available numerical weather prediction (NWP) models in

an ensemble to predict air quality measurements. That was followed by the combination of both NWP and air quality models [33]. [34] developed an ensemble of several forecasts which were input into a Lagrangian air dispersion model. The next breakthrough appeared with the idea to create a mean ensemble to predict the ozone concentration [35,36], and [37]. In 2007, [38] developed Polyphemus, an ensemble consisting of two Gaussian and two Eulerian models. To date there has been no published research that combines the superensemble method with air dispersion modeling.

The superensemble is composed of two parts: a training phase and a forecast phase. The training phase entails the previous forecasts from the member models and the corresponding observed state [39]. The parameters consistent with all the models in the pool are regressed against the observed state using a linear regression technique. This linear regression technique is performed through a minimization function which acts to limit the spread between the parameters and the observed state [40]. The minimization function is given by this equation:

$$J = \sum_{t=0}^{\text{length}} (S(t) - O(t))^2 \quad (1)$$

where J is the error term which needs to be minimized for obtaining the statistical weights, length is the length of the training period dataset, S(t) is the superensemble prediction, and O(t) is the observed state [41]. The weights are then sent to the forecast phase where the superensemble forecast is created. The weights are based solely on each model's performance in the training phase and can be negative if that model's result is negatively correlated with the result of the other models in the superensemble [42] and [43]. The weights can also be modified in consideration of the regional variation of skills of each model. These varied weights make the forecast more accurate compared to those from the biased-removed ensemble forecasts [46]. The length of the training period is an important factor in obtaining accurate forecasts. Prior superensemble research has found that between 50 and 75 cases are needed to obtain an acceptable forecast [47]. One of the limitations of the training period is the changing of member models. The amount of data sets which can be used is limited by the last change of each model [48].

The prediction phase of the superensemble is found through the following equation:

$$S(t) = \bar{O} + \sum_{i=1}^N a_i (F_i(t) - \bar{F}_i) \quad (2)$$

where $\bar{O}$ is the mean of the observed value in the training phase, N is the number of models in the superensemble, a_i is the regression coefficient or weight of each model, $F_i(t)$ is the forecast made by each model I, and $\bar{F}_i$ is the mean of all forecasts over the training period.

3. Methodology

The superensemble algorithm used in this research was developed through the renovation of the algorithm used in the Amorphe project [38]. Their superensemble was

adapted from one which interfaced with rainfall precipitation models to one which could be applied to the air dispersion models discussed previously. Since the air dispersion models were run under steady state conditions, further alterations were performed to change the measurement variable for the training period. The effect of the alteration shifted the paradigm from one of elapsed time (days) to one of distance traveled (meters). This is how a distance based superensemble was created.

In this research four different air dispersion models (Gaussian Plume, Variable K Theory, Box, and AFTOX) were used individually and then were combined into a superensemble. The models were discussed in greater detail in [49,50,51,52], and [53].

Two different sets of pollutant concentration profiles were used in the creation and evaluation of the superensemble algorithm. The first data set was used in the training phase as the input data needed for the superensemble creation [54] and [55]. The remaining data set was used in the prediction phase for evaluation purposes. Unless otherwise specified, the intervals for both the superensemble training and prediction phases were set at 500m. This meant that over the course of the emitted distance several superensemble calculations were made. The data output from the superensemble will give the calculated pollutant concentration and the weights given to each model [56]. The accuracy of the superensemble and each model can then be obtained by comparing their results to the measured profile.

This research was applied to a case study which used common procedures. These procedures were uniformly performed and included superensemble formation and performance evaluation using both the Root Mean Square Error (RMSE) [57,58], and [59] and the Brier Skill Score (BSS) [60] and [61]. [47] and [62] provided the guidelines for applying these methods for air dispersion models.

4. Superensemble Formation

The creation of the superensemble was performed through the execution of a multi-step process. The first step was to run each member model using a given set of initial conditions unique for each data set. The concentration profile was calculated for each distance away from the point source found within the data set. The concentrations obtained by all the models in the pool were then regressed against the measurements of the data set using a linear regression technique [63]. This linear regression technique was performed through applying a non-linear minimization program written using LINGO 13.0 to solve the minimization function. This function acts to limit the spread between the parameters and the observed state and is commonly solved through both matrices and the Gauss-Jordan Elimination Method [64]. The minimization function used is similar to Equation 1 with the adaptation of using downwind distance as opposed to time elapsed due to steady state conditions and is given below:

$$J = \sum_{x=0}^{length} (S(x) - O(x))^2 \quad (3)$$

where J is the error term that needs to be minimized to obtain the statistical weights, x is the length of the downwind distance away from the point source, S(x) is the superensemble prediction of the pollutant concentration, and O(x) is the observed value of the pollutant concentration. The need to find the necessary weights for each model made it necessary for Equation 3, the training phase equation, to be combined with Equation 4 the prediction phase equation. This new equation is given below.

$$S(x) = \bar{O} + \sum_{i=1}^N a_i (F_i(x) - \bar{F}_i) \quad (4)$$

In the prediction equation $\bar{O}$ is the mean of the observed value in the training phase, N is the number of models in the superensemble, a_i is the regression coefficient or weight of each model, $F_i(x)$ is the forecast made by each model i, and $\bar{F}_i$ is the mean of all forecasts over the training period. When combined these two equations form Equation 5.

$$J = \sum_{x=0}^{length} \left(\left(\bar{O} + \sum_{i=1}^N a_i (F_i(x) - \bar{F}_i) \right) - O(x) \right)^2 \quad (5)$$

The value of J was then minimized through the previously mentioned LINGO program designed to simulate the Gauss-Jordan elimination method others have successfully used [65,66,67], and [68].

Once the model weights were calculated for each model, the superensemble was then created. The next step was the completion of the training phase. After that, the prediction phase occurred. This meant that the superensemble and the member models were applied to predict the pollutant concentration values from a completely different data set. The member models and superensemble were run using the new initial conditions relevant to the new data set to obtain the predicted values [69]. Once the calculated values are obtained, the next logical step is to evaluate how well the member models and superensemble were able to perform. The purpose of this case study is to determine the effectiveness of the superensemble for NO_2 emission prediction.

5. Baldwin NO_2 Case Study

In this case study the Westvaco data set was used for the superensemble training phase and the Baldwin data set was used for the prediction phase [47]. The initial parameters required for the operation of each model were obtained from the initial conditions given in each data set. These initial values are given in Table 1. The NO_2 concentration data that had been collected from both the Westvaco site and Baldwin site was organized into three separate files. These files contained the concentration measurements (WVA01, WVA03, and WVA24) and (BAL01, BAL03, and BAL24) originating from slightly different time periods. Concentration profiles of the three

Westvaco measurement runs are given in Figure 1 and the three Baldwin measurement runs are given in Figure 2. The data from all three of these Westvaco measurement runs as well as their initial conditions was then averaged out to simulate a single collection obtained using steady state conditions [70].

However, during the Baldwin measurement runs, the data from the BAL01 run was found to be significantly different from the other two runs. The best solution to this was to treat BAL01 as an outlier. This meant that there were only two measurement runs. Their initial conditions were averaged to simulate a single collection under steady state conditions.

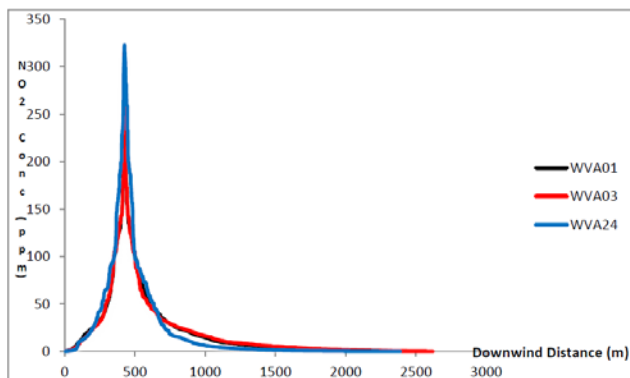


Figure 1. NO_2 concentration profile for Westvaco data set

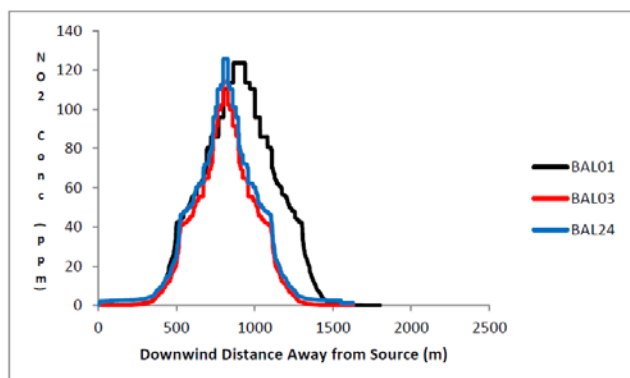


Figure 2. NO_2 concentration profile for Baldwin data set

Table 1. Initial Conditions for Member Model Calculations

Location	WVA01	WVA03	WVA24	Mean WVA	BAL01	BAL03	BAL24	Mean BAL
Emission Rate of Pollutant ($\frac{g}{m}$)	10365	10550	9300	10070	24650	26340	21590	24190
Wind Velocity ($\frac{m}{s}$)	3.67	4.88	3.94	4.15	2.04	1.52	2.35	1.97
Stack Height (m)	184.8	184.8	184.8	184.8	252.2	252.2	252.2	252.2
Stack Diameter (m)	5.94	5.94	5.94	5.94	3.38	3.38	3.38	3.38
Stack Gas Exit Temperature ($^{\circ}\text{K}$)	425	421	373	406	4.52	391	440	427
Ambient Temperature ($^{\circ}\text{K}$)	271	272	275	273	284	281	279	282
Exit Velocity ($\frac{m}{s}$)	28.9	31.7	37.6	32.7	51.2	57.8	42.7	46.0
Stability Category	F	F	F	F	F	F	F	F
Date	1/24/83	1/24/83	1/24/83	1/24/83	5/8/91	5/8/91	5/9/91	5/8/91
Time	10:15	13:45	16:35	13:30	17:35	19:05	08:20	19:00

Figure 3 displays the profile of the calculated concentrations from the models as well as the concentrations that were measured during the training

phase. The models were calibrated during this period to minimize the difference with the measured concentration. Once the weights were obtained, the superensemble was then created and utilized along with the member models. The purpose of this was to predict the concentrations that occur at the Baldwin location.

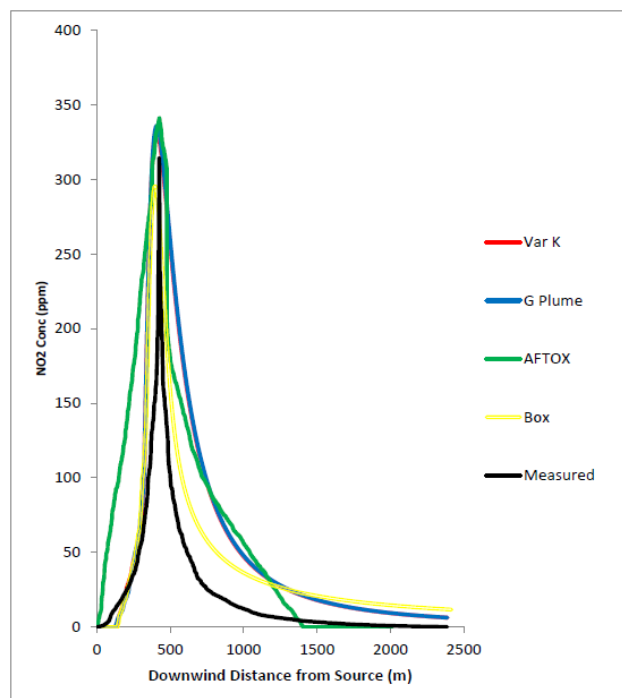


Figure 3. Calculated NO_2 Concentration Profile for Superensemble Training Period

The initial range for the superensemble calculation was arbitrarily set at a value of 500m. Because the total distance of the concentration profile was just under 2500m, five total superensemble calculations were necessary. Table 2 shows the calculated superensemble weights for that were given for each model over all five distance ranges. From the information shown in Table 2, it becomes apparent that the weights for each model vary significantly from range to range.

Table 2. Calculated Superensemble Weights

Distance Range	Variable K Theory Model	Gaussian Plume	AFTOX	Box Model
(m)	Weight Value			
0-500	0.22	0.15	0.59	0.03
501-1000	0.01	0.09	0.65	0.24
1001-1500	0.31	0.29	0.25	0.14
1501-2000	0.50	0.44	0.01	0.05
2001-2500	0.40	0.49	0.01	0.10

The next portion of the study began to explore the effects of the superensemble range. A superensemble was created from the calculated superensemble weights of Table 2. This superensemble will now be called a Fixed Superensemble. One of the observed drawbacks was that the weights for each model may vary at different distances. This means that the superensemble may not have the optimum weight values at any given point within the distance range. A second Superensemble was created to observe the effects. This new superensemble used a

continuous range of 500m and calculated the model weights at every 1m point. This created superensemble will now be called a Continuous Superensemble. Figure 4 shows the calculated concentrations for each member model and both Fixed and Continuous superensembles for the Baldwin data set. Over the course of the entire concentration range, the graph gets very busy. Figure 5 shows a smaller snapshot of the calculated concentrations over an 100m interval.

A further analysis between the superensembles is undergone. Figure 6 shows the calculated concentrations for both Fixed and Continuous superensembles over the course of the entire concentration range. Figure 7 clearly demonstrates the large advantage of determining model weights at each point as it is magnified over an 50m interval. Being able to adjust higher or lower at every point makes such a huge difference in accuracy and precision.

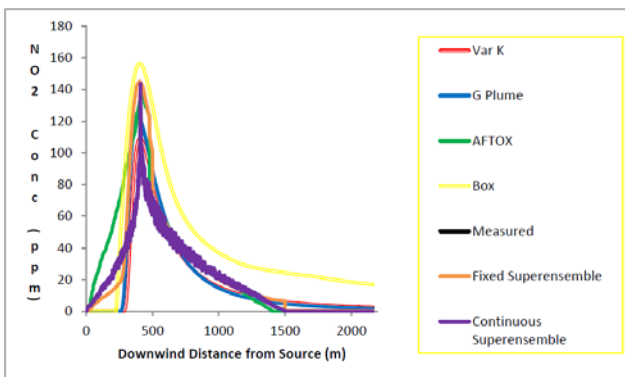


Figure 4. Calculated NO_2 Concentration for Entire Range of Baldwin data set

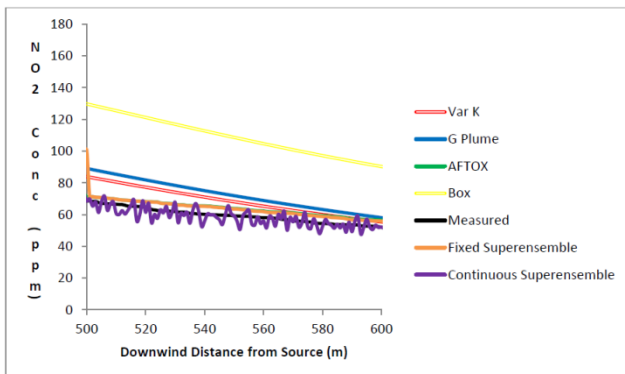


Figure 5. Calculated NO_2 Concentration for 100m Interval of Baldwin data set

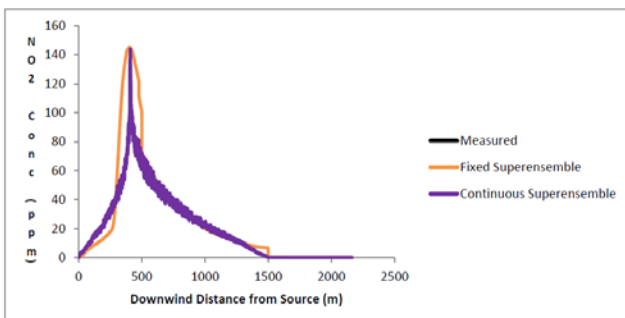


Figure 6. Calculated NO_2 Concentration for Entire Range of Baldwin data set

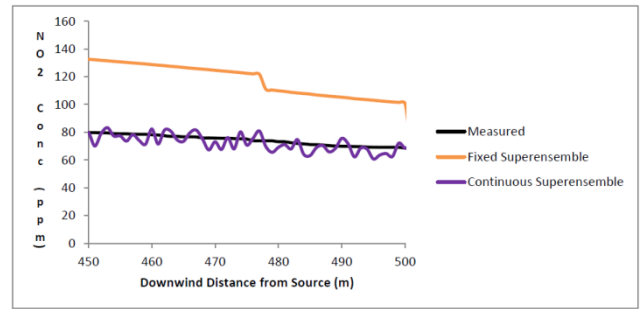


Figure 7. Calculated NO_2 Concentration for 50m Interval of Baldwin data set

6. Error Determination

Root Mean Square Error (RMSE) was used as the method to evaluate error. Figure 8 shows the calculated RMSE for the concentration profile for all member models and both superensembles. As one can see, the highest error generally occurred in the area around the peak. The one clear exception was the AFTOX model which significantly overpredicted concentrations after the peak.

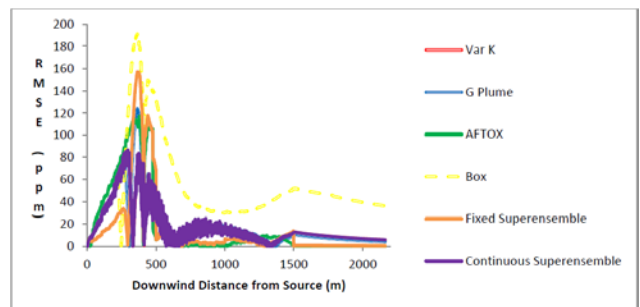


Figure 8. Calculated Root Mean Square Error (RMSE)

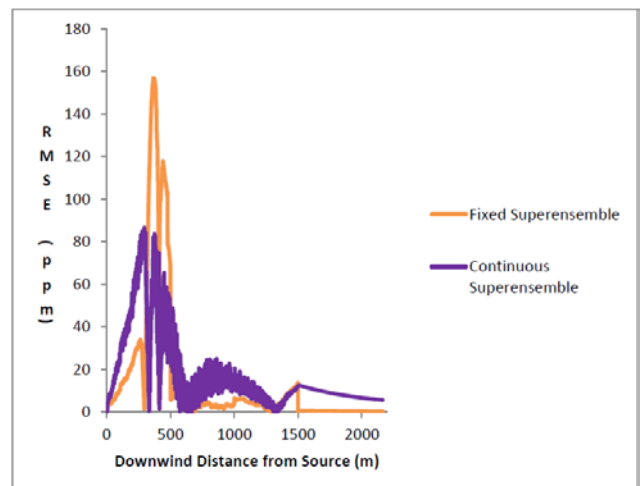


Figure 9. Calculated RMSE over Downwind Profile

The difference between the different superensembles really becomes apparent as shown in Figure 9. While the Fixed Superensemble is slightly better in some stretches of the concentration profile, the Continuous Superensemble is very accurate around the peak where it has very little error overall. This is due to it being able to constantly readjust relative to the smaller training period size. Thus, it can get closer to calculating the measured

value. Most of the stretches where the Fixed Superensemble had a lower RMSE value were characterized by very low measured concentration values so the difference between the two was minimal overall.

7. Conclusion

One of the key limiting factors in the effectiveness of the superensemble is the reliance of using the member models to obtain the weights during the training phase. Because the models are often missing the peak on the concentration profile it makes the RMSE values higher there than at other points of the concentration profile. This also affects the superensemble due to their relationship. In many cases it can be indicated that the superensemble is limited by the performance of the best model. However, by using very low or negative weights on the worst models, slightly better performance is theoretically possible. It is advantageous to start out with very accurate and consistent models. But a perfectly predicting model would negate the need for a superensemble in the first place.

Where the superensemble gains its competitive edge is through the inconsistency of the models. An individual model tends to have regions of above average accuracy as well as regions with poor predictions. By having the superensemble continually near the best performing model, it constantly beats all the member models. This is due to the models losing their effectiveness in certain regions of the concentration profile. This is the reason why the range size is so important towards superensemble accuracy. Further study of the effects of decreasing the range size is planned as well as exploring the full tradeoff of implementing continuous superensemble calculations at each measurement point with the amount of calculation time and computing power that would be required.

Another point of contention is the method of evaluation for the models and the superensemble. One of the issues with using RMSE is the variability of the calculations as one travels away from the source. One model may have the lowest error at one distance and then have one of the highest values further downwind. Using the mean RMSE can give a gauge of accuracy. However, there is the chance of applying the model or superensemble with the lowest mean RMSE and having the highest error at the property line.

Future research will look into incorporating the Briar score, which is dependent on the threshold concentration being exceeded to trigger an event. Thus, it would be able to determine accuracy around the threshold concentration level. The Brier score operates on the assumption that the difference in the number of events will lead to a better prediction of accuracy. Other methods of error evaluation will need to be explored.

References

- [1] Turner, D. B. and R. H. Shulze (2007). Practical Guide to Air Dispersion Modeling. Trinity Consultants. Dallas.
- [2] Khajehnajafi, S., R. Pourdarvish, and H. Shah (2011). "Modeling Dispersion and Deposition Of Smoke Generated from Chemical Fires." *Process Safety Progress*. 30(2): 168-177.
- [3] Banerjee, T., S. C. Barman, and R.K. Srivastava. (2011). Application of Air Pollution Dispersion Modeling for Source Contribution Assessment and Model Performance Evaluation at Integrated Industrial Estate in Pantnagar. *Environmental Pollution*. 159(4): 865-875.
- [4] Pasquill, F. (1974). Atmospheric Dispersion Parameters in Gaussian Plume Modeling. Part II. Possible Requirements for Change in the Turner Workbook Values. EPA-600/4-76-030b, U.S. Environmental Protection Agency.
- [5] Monache, L. D., X. Deng, H. Modzelewski, G. Hicks, T. Cannon, R.B. Stull, and C. Cenzo. (2007). "Air Quality Ensemble Forecast Over the Lower Fraser Valley, British Columbia." *Air Pollution Modeling and Its Application*. Springer US: 399-402.
- [6] Mallet, V. and Sportisse, B. (2006). Ensemble Based Air Quality Forecasts: A Multi Model Approach Applied to Ozone. *J. Geophys. Res.*, 111-119.
- [7] Yun, W. T., L. Stefanova, and T.N. Krishnamurti. (2003). "Improvement of the Multimodel Superensemble Technique for Seasonal Forecasts." *Journal of Climate*. 16(22): 3834.
- [8] Webber, D.M., M.J. Iivings, and R.C. Santon. (2011). "Ventilation Theory and Dispersion Modelling Applied to Hazardous Area Classification." *Journal of Loss Prevention in the Process Industries*. 24(5): 612-621.
- [9] Warner, T., R.S. Sheu, J.F. Bowers, R.I. Sykes, G.C. Dodd, and D.S. Henn. (2002). Ensemble Simulation with Coupled Atmospheric Dynamic and Dispersion Models Illustrating Uncertainties in Dosage Simulations. *Journal of Applied Meteorology*. 41. 488-504.
- [10] Jordan, M. R. (2007). Using the Superensemble Method to Improve Eastern Pacific Tropical Cyclone Forecasting. M.S. Thesis, Florida State University.
- [11] Willford, C.E. (2002) Real-time Superensemble Tropical Cyclone Prediction. Ph.D. Dissertation, Florida State University.
- [12] Warner, T., R.S. Sheu, J.F. Bowers, R.I. Sykes, G.C. Dodd, and D.S. Henn. (2002). Ensemble Simulation with Coupled Atmospheric Dynamic and Dispersion Models Illustrating Uncertainties in Dosage Simulations. *Journal of Applied Meteorology*. 41. 488-504.
- [13] Sportisse, B. (2007). "A review of current issues in air pollution modeling and simulation." *Computational Geosciences*. 11(2): 159-181.
- [14] Chaves, R. R., R. S. Ross, and T.N. Krishnamurti (2005). "Weather and Seasonal Climate Prediction for South America Using a Multi Model Superensemble." *International Journal of Climatology*. 25(14): 1881-1914.
- [15] Yun, W. T., L. Stefanova, A.K. Mitra, and W. Dewar. (2005). "A Multi Model Superensemble Algorithm for Seasonal Climate Prediction Using DEMETER Forecasts." *Tellus: Series A*. 57(3): 280-289.
- [16] Stefanova, L. and T. N. Krishnamurti (2002). "Interpretation of Seasonal Climate Forecast Using Brier Skill Score, The Florida State University Superensemble, and the AMIP-I Dataset." *Journal of Climate*. 15(5): 537.
- [17] Krishnamurti, T. N., S. Basu, J. Sanjay, and C. Gnanaseelan. (2008). "Evaluation of Several Different Planetary Boundary Layer Schemes Within a Single Model, a Unified Model, and a Multi Model Superensemble." *Tellus A*. 60(1): 42-61.
- [18] Krishnamurti, T. N., C. Gnanaseelan, and A. Chakraborty. (2007). "Prediction of the Diurnal Change Using a Multimodel Superensemble. Part I: Precipitation." *Monthly Weather Review*. 135(10): 3613-3632.
- [19] Mo, K. C. (2003). "Ensemble Canonical Correlation Prediction of Surface Temperature over the United States." *Journal of Climate*. 16(11): 1665-1680.
- [20] Kumar, T., J. Sanjay, B. Basu, A. Mitra, D. Rao, O. Sharma, and P. Pal. (2007). "Experimental Superensemble Forecasts of Tropical Cyclones over the Bay of Bengal." *Natural Hazards*. 41(3): 471-485.
- [21] Kim, M.K., I.S. Kang, C.K. Park, and K.M. Kim. (2004). "Superensemble Prediction of Regional Precipitation over Korea." *International Journal of Climatology*. 24(6): 777-790.
- [22] Mahmud, M. and R. S. Ross (2005). "Precipitation Assessment of a Superensemble Forecast over South East Asia." *Meteorological Applications*. 12(2): 177-186.
- [23] Chaves, R. R. and R. S. Ross (2005). "Weather and seasonal climate prediction for South America using a multi-model

- superensemble." *International Journal of Climatology*. 25(14): 1881-1914.
- [24] Krishnamurti, T. N., A. K. Mishra, and A. Rajeevan. (2009). "Improving Global Model Precipitation Forecasts over India Using Downscaling and the FSU Superensemble. Part I: 1–5-Day Forecasts." *Monthly Weather Review*. 137(9): 2713-2735.
- [25] Mahmud, M. and R. S. Ross (2006). "Verification of the Motion Fields over Southeast Asia by the Florida State University (FSU) Superensemble Forecast." *Singapore Journal of Tropical Geography* 27(2): 176-190.
- [26] Cane, D. and M. Milelli (2006). "Weather Forecasts Obtained with a Multimodel SuperEnsemble Technique in a Complex Orography Region." *Meteorologische Zeitschrift*. 15(2): 207-214.
- [27] Cartwright, T. J. and T. N. Krishnamurti (2007). "Warm Season Mesoscale Superensemble Precipitation Forecasts in the Southeastern United States." *Weather and Forecasting*. 22(4): 873-886.
- [28] Krishnamurti, T., A. Sagadevan, A. Chakraborty, A. Mishra, and A. Simon. (2009). "Improving Multi Model Weather Forecast of Monsoon Rain over China Using FSU Superensemble." *Advances in Atmospheric Sciences*. 26(5): 813-839.
- [29] Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, S. Gadgil, and S. Surendran. (1999). "Improved Weather and Seasonal Climate Forecasts from Multi Model Superensemble." *Science*. 285(5433): 1548-1550.
- [30] Chakraborty, A. and T. N. Krishnamurti (2009). "Improving Global Model Precipitation Forecasts over India Using Downscaling and the FSU Superensemble. Part I: Seasonal Climate." *Monthly Weather Review*. 137(9): 2736-2757.
- [31] Chakraborty, A., T. N. Krishnamurti, and C. Gnanaseelan. (2007). "Prediction of the Diurnal Cycle Using a Multimodel Superensemble. Part II: Clouds." *Monthly Weather Review*. 135(12): 4097-4116.
- [32] Vautard, R., Blond, N., Schmidt, H., Derognat, C., Beekmann, M. (2001). Multi Model Ensemble Ozone Forecast over Europe: Analysis of Uncertainty. *Journal of Applied Meteorology*. 40: 326–344.
- [33] Riccio, A., A. Ciamarella, E. Solazzo, and S. Potemski . (2012). "On the Systematic Reduction of Data Complexity in Multi Model Atmospheric Dispersion Ensemble Modeling." *J. Geophys. Res.* 117(5): 5302-5314.
- [34] Straume, A. G. (2001). A More Extensive Investigation of the Use of Ensemble Forecasts for Dispersion Model Evaluation. *J. Appl. Meteor.*, 40, 425–445.
- [35] Mölter, A., S. Lindley, A. Simpson, and R. Agius. (2010). "Modeling Air Pollution for Epidemiologic Research — Part I: A Novel Approach Combining Land Use Regression and Air Dispersion." *Science of the Total Environment*. 408(23): 5862–5869.
- [36] Mohan, M., S. Bhati, and A. Rao. (2011). "Application of Air Dispersion Modelling for Exposure Assessment from Particulate Matter Pollution in Mega City Delhi." *Asia-Pacific Journal of Chemical Engineering*. 6(1): 85-94.
- [37] Peltier, L. J., S. E. Haupt, A. Deng, J.A. Lee, K.J. Long, and A.J. Annunzio. (2010). "Parameterizing Mesoscale Wind Uncertainty for Dispersion Modeling." *Journal of Applied Meteorology & Climatology*. 49(8): 1604-1614.
- [38] Mallet, V., E. Qu., B. Sportisse, A.D. Biasi, M. Debry, E. Korsakissok, I. Wu, L. Roustan, Y. Sartelet, K. Tombette, and H. Foudhil. (2007). The Air Quality Modeling System Polyphemus. *Atmos. Chem. Phys. Discuss.* 7(3): 6,459–6,486.
- [39] Berliner, L. M. and K. Yongku (2008). Bayesian Design and Analysis for Superensemble Based Climate Forecasting. *Journal of Climate*. 21(9): 1891-1910.
- [40] Ke, Z., W. Dong, and P. Zhang. (2008). "Multimodel Ensemble Forecasts for Precipitations in China in 1998." *Advances in Atmospheric Sciences*. 25(1): 72-82.
- [41] Sirdas, S., R. S. Ross, and A. Chakraborty (2007). "Evaluation of the FSU Synthetic Superensemble Performance for Seasonal Forecasts over the Euro-Mediterranean Region." *Tellus: Series A* 59(1): 50-70.
- [42] Challa, V., J. Indranti, and J. Baham. (2008). "Sensitivity of Atmospheric Dispersion Simulations by HYSPLIT to the Meteorological Predictions from a Meso Scale Model." *Environmental Fluid Mechanics*. 8(4): 367-387.
- [43] Krishnamurti, T. N., A. Chakraborty, R. Krishnamurti, W.K. Dewar, and A.C. Clayson. (2006). "Seasonal Prediction of Sea Surface Temperature Anomalies Using a Suite of 13 Coupled Atmosphere–Ocean Models." *Journal of Climate*. 19(23): 6069-6088.
- [44] Szymczak, H. L. and T. N. Krishnamurti (2006). "Skill of Synthetic Superensemble Hurricane Forecasts for the Canadian Maritime Provinces." *Meteorology and Atmospheric Physics*. 93(3): 147-163.
- [45] Krishnamurti, T. N., C. M. Kishtawal, D.W. Shin, and E. Williford. (2000). "Improving Tropical Precipitation Forecasts from a Multianalysis Superensemble." *Journal of Climate*. 13(23): 4217.
- [46] Wilks, D. S., (1995) *Statistical Methods in the Atmospheric Sciences*. International Geophysics Series, 59, 467, 317-324.
- [47] Gronwald, F. and S. Y. Chang. (2018). "Evaluation of the Precision and Accuracy of Multiple Air Dispersion Models." *Journal of Atmospheric Pollution*, 6:(1), 1-11.
- [48] Sharan, M. and M. Modani (2007). "Variable K-theory Model for the Dispersion of Air Pollutants in Low Wind Conditions in the Surface Based Inversion." *Atmospheric Environment*. 41(33): 6951-6963.
- [49] Sharan, M. and A.K. Yadav (1998). "Simulation of Diffusion Experiments Under Light Wind, Stable Conditions by a Variable K-Theory Model." *Atmospheric Environment*. 32(20): 3481-3492.
- [50] Taseiko, O. V., S. V. Mikhailuta, A. Pitt, A.A. Lezhenin, and Y.V. Zakharov. (2009). "Air Pollution Dispersion Within Urban Street Canyons." *Atmospheric Environment*. 43(2): 245-252.
- [51] Labovský, J. and L. Jelemenský (2011). "Verification of CFD Pollution Dispersion Modelling Based on Experimental Data." *Journal of Loss Prevention in the Process Industries*. 24(2): 166-177.
- [52] Kiša, M. and L. Jelemenský (2009). "CFD Dispersion Modelling for Emergency Preparedness." *Journal of Loss Prevention in the Process Industries*. 22(1): 97-104.
- [53] Vijaya, K., T. S. V., T. N. Krishnamurti, M. Fiorino, and M. Nagata. (2003). "Multimodel Superensemble Forecasting of Tropical Cyclones in the Pacific." *Monthly Weather Review* 131(3): 574.
- [54] Shin, D. W. and T. N. Krishnamurti (2003). "Short to Medium Range Superensemble Precipitation Forecasts Using Satellite Products: 1. Deterministic forecasting." *J. Geophys. Res.* 108(D8): 8383.
- [55] Shin, D. W. and T. N. Krishnamurti (2003). "Short to Medium Range Superensemble Precipitation Forecasts Using Satellite Products: 2. Probabilistic forecasting." *J. Geophys. Res.* 108(D8): 8454.
- [56] Mishra, A. and T. Krishnamurti (2007). "Current Status of Multi Model Superensemble and Operational NWP Forecast of the Indian Summer Monsoon." *Journal of Earth System Science*. 116(5): 369-384.
- [57] Ross, R. S. and T. N. Krishnamurti (2005). "Reduction of Forecast Error for Global Numerical Weather Prediction by The Florida State University (FSU) Superensemble." *Meteorology and Atmospheric Physics*. 88(3): 215-235.
- [58] Pagowski, M., G. A. Gell, S.E. Peckham, W. Gong, L. Delle Monache, J. McQueen, and P. Lee. (2006). "Application of Dynamic Linear Regression to Improve the Skill of Ensemble Based Deterministic Ozone Forecasts." *Atmospheric Environment*. 40(18): 3240-3250.
- [59] Brier, G. W., (1950) Verification of Forecasts Expressed in Terms of Probability. *Mon. Wea. Rev.* (78): 1–3.
- [60] Mahmud, M. (2004). "Skill of a Superensemble Forecast over Equatorial South East Asia." *International Journal of Climatology*. 24(15): 1963-1972.
- [61] Murphy, A. H. (1973) A Note on the Ranked Probability Skill Score. *J. Appl. Meteor.* 10. 155–156.
- [62] Krishnamurti, T. N., C. M. Kishtawal, Z. Zhang, T. LaRow, D. Bachiochi, S. Gadgil, and S. Surendran. (2000). "Multimodel Ensemble Forecasts for Weather and Seasonal Climate." *Journal of Climate*. 13(23): 4196-4216.
- [63] Eric-Williford, C., T. N. Krishnamurti, R.C. Torres, and S. Cocke. (2003). "Real-Time Multimodel Superensemble Forecasts of Atlantic Tropical Systems of 1999." *Monthly Weather Review*. 131(8): 1878-1894.
- [64] Krishnamurti, T. N., S. Surendran, D.W. Shin, R.J. Correa-Torres, E. Williford, C. Kummerow, R.F. Adler, J. Simpson, R. Kakar, W.S. Olson, and F.J. Turk. (2001). "Real Time Multianalysis Multimodel Superensemble Forecasts of Precipitation Using TRMM and SSM/I Products." *Monthly Weather Review*. 129(12): 2861-2883.

- [65] Chakraborty, A. and T. N. Krishnamurti (2006). "Improved Seasonal Climate Forecasts of the South Asian Summer Monsoon Using a Suite of 13 Coupled Ocean–Atmosphere Models." *Monthly Weather Review*. 134(6): 1697-1721
- [66] Jordan, M. R., T. N. Krishnamurti, and A.C. Clayson. (2008). "Investigating the Utility of Using Cross-Oceanic Training Sets for Superensemble Forecasting of Eastern Pacific Tropical Cyclone Track and Intensity." *Weather & Forecasting*. 23(3): 516-522.
- [67] Chaves, R. R., A. K. Mitra, and T.N. Krishnamurti. (2005). "Seasonal Climate Prediction for South America with FSU Multi Model Synthetic Superensemble Algorithm." *Meteorology and Atmospheric Physics*. 89(1): 37-56.
- [68] Krishnamurti, T. N., K. Rajendran, S. Lord, Z. Toth, and X. Zou. (2003). "Improved Skill for the Anomaly Correlation of Geopotential Heights at 500 kPa." *Monthly Weather Review*. 131(6): 1082.
- [69] Mutemi, J. N., L.A. Ogallo, A.K. Mishra, T.N. Krishnamurti, and T.S. Vijaya Kumar. (2007). "Multi Model Based Superensemble Forecasts for Short and Medium Range NWP over Various Regions of Africa." *Meteorology and Atmospheric Physics*. 95(1): 87-113.
- [70] Ross, R. S., A. Chakraborty, A. Chen, L. Stefanova, and S. Sirdas. (2007). "Improved Seasonal Climate Forecasts for the Caribbean Region Using the Florida State University Synthetic Superensemble." *Meteorology and Atmospheric Physics*. 98(3): 137-174.



© The Author(s) 2024. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).