Evaluation and comparison of AERMOD and HPDM models

Zeng Quan Yao
Research Institute of Environmental Protection for Electric Power
10 Pudong Road, Pukou, Nanjing 210031, China
zqyao@990.net

Abstract

Two air quality modeling systems AERMOD and HPDM have been evaluated with Nanticoke database. The fractional bias, absolute fractional bias and composite performance measure for 1-hour and 24-hour components were used. The bootstrap procedure was applied to determine the significance of the difference in composite performance between the models. It indicates that HPDM is seriously under-prediction on stable conditions, AERMOD is only a little bit under-prediction. On unstable conditions, Case A of AERMOD among four cases gives the best performance.

1 Introduction

In order to develop a new generation of air quality modeling system applied in electric power industry of China, some new modeling systems developed recent years in other countries have been examined. It is found that HPDM developed by Hanna & Paine and AERMOD developed by Perry et.al. are two best candidates because of their state of the art techniques. Performance evaluation has been carried out. Nanticoke database was used in the evaluation.

2 Model description

Two modeling systems both contain two modules: meteorology data processor and dispersion processor. Owing to the fact that available observed SO2 data
were produced for elevated source, only elevated source module has been evaluated.

2.1 Meteorology data processor

The same meteorology data processor is used for two model systems in this evaluation even their original processors have a little bit difference.

2.1.1 Sensible heat flux

The sensible heat flux $H$, friction velocity $u^*$, temperature scale $\theta^*$ and Monin-Obukhov length $L$ are required in dispersion processor. When the measurement data were unavailable, they would be calculated from routine meteorological data. Two approaches were used according to the available data: profile approach and flux approach.

When observed wind speed at one height, air temperature at one height or temperature difference at two heights were available, $u^*$, $\theta^*$ and $L$ could be estimated by iteration from wind velocity and temperature profiles (van Ulden & Holtslag\(^5\)). Then the sensible heat flux was calculated with the equation $H=-\rho C_p u^* \theta^*$. This procedure is called as profile approach.

Flux approach is based on the conservation equation of heat flux $Q_L = H + LE + G$, where $Q_L$ is net radiation, $LE$ is latent heat flux and $G$ is soil heat flux. The sensible heat flux over the land in daytime was estimated applying Holtslag & van Ulden\(^4\) approach. Then the friction velocity $u^*$ and Monin-Obukhov length $L$ were calculated.

2.1.2 Boundary layer height

When measured boundary layer height was unavailable, it would be calculated using diagnostic (for stable conditions) or prognostic (for unstable conditions) models.

Nieuwstadt\(^5\) model was applied on stable conditions: $h_{id}/L = (0.3 u^-/fL)/(1+1.9 h_{id}/L)$. Avoiding the instantly change of $h$ with $u^*$ and $L$, a time evolution was considered using $dh/dt = (h_{id}-h)/\tau$, where $\tau$ is a time scale.

The prognostic model of Batchvarova & Gryning\(^6\) was used on unstable conditions. Comparison with observation indicated the model gave quite reliable prediction.

2.1.3 Wind velocity and temperature profiles

The velocity on the plume height was estimated with the profile relation (van Ulden & Holtslag\(^3\)). The potential temperature gradient at plume height was decided with a method introduced by Perry et.al.\(^2\).

2.1.4 Turbulence

The wind velocity deviations $\sigma_v$ and $\sigma_w$ are necessary in calculation of dispersion parameter $\sigma_y$ and $\sigma_z$. If $z/z_i \leq 1.0$, Panofsky et.al.\(^7\) and Hicks\(^8\) models
were used to calculate standard deviations of the wind velocities on unstable conditions, otherwise, the wind velocity deviations were taken as half of the surface values. Because the turbulence intensity on crosswind direction is impossible to reduce to a very small value (meso-scale effect), an adjustment to estimated $\sigma_{wj}$ might be made as $\sigma_v = \max (0.5 \text{ m/s}, \sigma_{wj})$. On stable conditions, the ground values of $\sigma_\infty$ and $\sigma_w$ were decided first as $\sigma_\infty = \max (0.5 \text{m/s}, 2.0u_\infty)$ and $\sigma_w = 1.2u_\infty$. Then the ground values were extrapolated to obtain the values at the other heights with $\sigma_{wj} = (1-0.5z/h)\sigma_\infty$ for $z<h$ and $\sigma_{wj} = 0.5\sigma_\infty$ for $z>h$. The final $\sigma_v$ was adjusted with $\sigma_v = \max (0.5\text{m/s}, \sigma_{wj})$. The vertical turbulence was calculated with interpolation formula $\sigma_w = [1-(1-C)z/100]\sigma_w$ between the ground and 100 m. At $z=100$ m and above $\sigma_w = C\sigma_w$, where $C$ is constant and depends on time and cloud cover.

2.3 Dispersion models

Two systems use PDF model for unstable conditions and Gaussian model for stable conditions. They both assume the crosswind distribution of pollutant concentration is Gaussian. However, they are different in dealing with the dispersion parameters, plume rise and penetration.

2.3.1 Unstable condition

AERMOD assumes that the contribution comes from three sources on unstable conditions. They are direct source, indirect source and penetrated source. The lateral integrated concentration for direct source has the same form as low buoyant source ($F_\ast \leq 0.1$) model of HPDM. It can be written as:

\[
C_y = \frac{(1-P)Q}{\sqrt{2\pi u}} \sum_{j=1}^{2} \frac{\lambda_j}{\sigma_{wj}} \sum_{m=0}^{\infty} \left\{ \exp\left[ -\frac{(z-h_j-2mz_j)^2}{2\sigma_{wj}^2} \right] \\
+ \exp\left[ -\frac{(z+h_j+2mz_j)^2}{2\sigma_{wj}^2} \right] \right\}. \tag{1}
\]

Where $P$ is penetration ratio and $Q$ is emission rate. For HPDM: $\lambda_1 = 0.6, \lambda_2 = 0.4; \sigma_{w1}/z_i = (0.64F_\ast^{-2/3}X_i^{-4/3} + 0.058X_i^{-2})^{1/2}, \sigma_{w2}/z_i = (0.64F_\ast^{-2/3}X_i^{-4/3} + 0.23X_i^{-2})^{1/2}, h_1 = h_s + (1.6F_\ast^{-1/3}X_i^{-2/3} - 0.35X_i)z_i, h_2 = h_s + (1.6F_\ast^{-1/3}X_i^{-2/3} + 0.4X_i)z_i, F_\ast = F_i/(uw_i^2z_i); F_i = w_i X_i /u_z$. For direct source of AERMOD: the parameters are: $\sigma_{w1}^2 = \sigma_v^2 + (\sigma_v X_i/u)^2, h_f = h_s + \Delta h + w_j X_i/u; \Delta h = (3F_{m_i} X_i^{3}/2\beta^2 u^3 + 3F_{l_i} X_i^{3}/2\beta^2 u^3)^{1/3}$.

AERMOD uses following model for indirect source:

\[
C_{wy} = \frac{(1-P)Q}{\sqrt{2\pi u}} \sum_{j=1}^{2} \frac{\lambda_j}{\sigma_{wj}} \sum_{m=1}^{\infty} \left\{ \exp\left[ -\frac{(z-h_j-2mz_j)^2}{2\sigma_{wj}^2} \right] \\
+ \exp\left[ -\frac{(z-h_j+2mz_j)^2}{2\sigma_{wj}^2} \right] \right\}. \tag{2}
\]
Where $\sigma_0 = \sigma_0 x/u, h_0 = h_0 + \Delta h_e + w_p x/u$, and $\Delta h_e = \Delta h - \Delta h$, $\Delta h$ and $\Delta h_e$ each has two options, so it has totally four combinations. The option 1 of $\Delta h$ is $\Delta h = (3F_m x/\beta^2 u^2 + 3F_b x^2/2\beta^2 u^3)^{1/3}$; the option 2 is $\Delta h = w_p x/u, w_p = u(z_h-h_f)/x_h$. $x_h$ is the solution of $\Delta h = z_h - h_f$ in option 1. $\Delta h_e$ has two options too. The option 1 is $\Delta h_e = 0.64F_3 x^{4/3}/u^2(z_h-h_f)$. The option 2 is $\Delta h_e = [2F_b z/\alpha u R_y]^{1/2} x/u$. The other parameters in equation (2) can be found in reference 2 and other literature. HPDM model has no this part.

AERMOD model assumes that the material penetrated into the inversion will disperse from a source at height $h_{ep}$, where $h_{ep} > z_i$. For this source, the concentration field is assumed to be Gaussian with the dispersion capped or limited at the height $h_{ep}$. Crosswind integrated concentration is given by:

$$C_{py} = \frac{PQ}{2\pi u \sigma_{zp}} \sum_{m=0}^{\infty} \{ \exp\left(-\frac{(z-(2m+1)h_{ep})^2}{2\sigma_{zp}^2}\right) + \exp\left(-\frac{(z+(2m+1)h_{ep})^2}{2\sigma_{zp}^2}\right) \}.$$  

(3)

HPDM assumes that the penetrated plume does not return to the mixed layer anymore. Moreover, HPDM use different model for high buoyant source ($F_* > 0.1$): $C_y(x,0)uh/Q = 0.056X_* /F_* , X_*/F_* < 10$; $C_y(x,0)uh/Q = \exp\left[-(7F_*/X_*)^{3/2}\right], X_*/F_* \geq 10$.

2.3.2 Neutral and stable conditions

On near neutral and stable conditions, the concentration distribution is assumed to be Gaussian. In HPDM model, if the plume height is lower than the boundary layer height, the reflection both on the boundary layer top and the ground is considered. Otherwise, no reflection on boundary layer top is considered. In AERMOD, however, the boundary layer height will be adjusted to the plume height if the plume is above the boundary layer top.

3 Data base and evaluation method

3.1 Data base

The meteorology data and SO$_2$ monitoring data used in this evaluation are 1983 whole year data which came from Nanticoke Complex. The Complex has 11 elevated sources totally. Thermal Generation Station (TGS) gives the largest emission rate (95% of total emission). There are 20 SO$_2$ monitoring stations around Nanticoke Complex and running 24 hours a day, 365 days a year automatically. The distance of stations to the TGS stack varies from 2.5 km to 37 km. The observed SO$_2$ concentration and meteorology data were hourly averaged. The low level wind velocity was measured at 10m height; the
temperature was measured at 2m height. The accuracy of the wind direction was 10 degrees.

3.2 Evaluation methods

Comparison of observed and calculated meteorological data indicated that the meteorology processor could give reliable results. Considering the main purpose of this paper is to evaluate two dispersion models, we will focus our attention to the dispersion model evaluation.

Owing to the specific regulatory purposes of the models and the inherent limitations of the database, a strong emphasis is placed on the ability of models to predict peak concentrations independent of time of occurrence (Cox & Tikvarf). A major limitation of data is lack of accurate information necessary to pinpoint actual transport wind direction. The database is rich in terms of the number of observations but sparse in terms of the size of the monitory network. This fact argues for using peak concentration under the assumption that the data are dense enough in time to define the maximum concentration for different stability at each monitor. The other reason for focusing on peak concentrations is that the model system is used to estimate the impact of sources on ambient air quality standards. Because of the nature of the standards, the models must accurately predict the highest 1-hour, 24-hour average concentration independent of exactly when or where they may occur. For the above reasons, the evaluation is divided into two separate components: evaluation of 1-hour averages during specific meteorology conditions at each monitor and evaluation of 24-hour averages independent of meteorological condition or spatial location. The fractional bias was used as the fundamental measure of discrepancy between the measurement-based and prediction-based test statistic. The bootstrap procedure was used to estimate the uncertainty in the performance measurement.

Because peak concentrations were highly variable, a robust test statistic was calculated using information contained in the upper end of the distribution of concentrations. The statistics referred as the robust highest concentration (RHC) is preferred to the actual peak value because it mitigates the undesirable influence of unusual events. The RHC based on the tail exponential estimator can be expressed as $\text{RHC} = X(N) + \left[ X_m - X(N) \right] \log(3N-1)/2$, where $X_m$ is average of the N-1 largest values; $X(N)$ is the Nth largest value; N is the number of values exceeding the threshold value (N≤25).

Fractional bias (FB) was used as the basic measure of the model performance. General expression for the fractional bias is given by $\text{FB} = 2(O-P)/(O+P)$, where O = RHC calculated using the observed data; P = RHC calculated using the predicted data. The values of the fractional bias range between -2.0 (extreme over-prediction) and +2.0 (extreme under-prediction). Over-prediction by a factor of two would be indicated by FB value of -0.67 and under-prediction by a factor of two would be indicated by FB value of +0.67. As
a reference, a ratio of predicted RHC to observed RHC was also given by 
\[ R = \frac{P}{O} = \frac{(2-FB)}{(2+FB)}. \]

The evaluation for 1-hour average concentrations requires a separate fractional bias calculation for each meteorology condition and each monitor. The meteorology conditions are selected arbitrary as four categories according to the wind velocity and stability. They are two stability: stable and unstable, and two velocity categories: \( u \leq 4 \text{ m/s} \) and \( u > 4 \text{ m/s} \). The evaluation for 24-hour average concentrations was determined by comparing the largest measurement and predicted-based RHC statistic from all monitors in the network.

The individual fractional bias components (1-hour and 24-hour) were averaged to produce a composite result. For the 1-hour average component, the composite bias was computed by averaging the separate results for each meteorological category and station combination. Considering the positive and negative components might be offset each other when calculating the composite bias, the absolute fractional bias (AFB) was used to calculate the composite fractional bias. The algebraic expression for the composite performance measure is
\[ C = 0.5 \left\{ \frac{\sum \text{AFB}_i(i,j)}{N_s N_m + \text{AFB}_{24}} \right\}, \]
where \( \text{AFB}_i(i,j) \) is AFB for 1-hour under the \( i \) meteorological category at the \( j \) station; \( \text{AFB}_{24} \) is AFB for 24-hour; \( N_s \) is the total number of monitoring station (\( N_s = 20 \) in our evaluation); \( N_m \) is the total number of meteorological categories (\( N_m = 4 \)).

The smaller the composite performance measure (\( C \) value), the better is the overall performance of the model. The difference between the composite performance of one model and another is referred as the model comparison measure. It can be used to judge the significance of the apparent superiority of one model over another. The expression for the model comparison measure is given by
\[ M(A,B) = C(A) - C(B). \]
Where \( C(A) \) = composite performance measure for Model A; \( C(B) \) = composite performance measure for Model B. A negative difference (\( M \) value) means that Model A is performing better, while a positive value indicates that Model B is performing better (Cox & Tikvarf).

The blocked bootstrap was used to generate the estimates of the sampling error. The original year of data (365 days) was partitioned into four blocks corresponding to the four seasons. Within each season, 3-day pieces were randomly sampled with replacement resulting in approximately 30 sampled units per season. This process was repeated using each of the four seasons to construct a complete bootstrap year. Three-day pieces were chosen to preserve day-to-day meteorology persistence. For each bootstrap year that was generated, the composite performance measure and model comparison measure were calculated. Overall composite performance measure and model comparison measure and their standard deviations were obtained from all of bootstrap data.

### 4 Evaluation results

Besides stable situation, four cases of AERMOD were calculated on unstable conditions. Their combinations are: Case A, option 1 of \( \Delta h \) and option 1 of \( \Delta h \);
Air Pollution

Case B, option 2 of $\Delta h$ and option 1 of $\Delta h_i$; Case C, option 1 of $\Delta h$ and option 2 of $\Delta h_i$; Case D, option 2 of $\Delta h$ and option 2 of $\Delta h_i$.

Scattered figures (not shown) indicate that both models are under-predicted at lower wind speed on stable conditions, but AERMOD gives a little bit better results than HPDM. On unstable conditions, however, it seems that HPDM is over-predicted, Case B and Case D of AERMOD are also over-predicted, while Case A and C are under-predicted. Comparing four cases of AERMOD, we find that both Cases A and C give better results than HPDM does.

In order to explore the quantitative comparison, Table 1 shows fractional bias, absolute fractional bias and average ratio of predicted and observed concentrations. From the table, it can be seen that two models are under-prediction seriously on stable conditions at lower wind speed, especially HPDM model. The prediction by HPDM model is less than 50% of the observation overall at lower wind speed category. The situation is improved much at higher speed category. On unstable conditions, it seems that Case A and Case C give quite good results. If considering the under-prediction on stable conditions, however, a little over-prediction will give overall better performance. So Case D may also be an applicable choice. Therefore, HPDM, Case A, Case C and Case D of AERMOD were run using the bootstrap data.

### Table 1. Fractional bias and ratios for different models

<table>
<thead>
<tr>
<th>Condition</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
<th>HPDM</th>
<th>AERMOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable (U ≤ 4 m/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>0.81</td>
<td>0.47</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stable (U &gt; 4 m/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.75</td>
<td>-0.11</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstable (U ≤ 4 m/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.21</td>
<td>0.52</td>
<td>-1.12</td>
<td>1.12</td>
<td>-0.08</td>
<td>0.58</td>
<td>-0.06</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unstable (U &gt; 4 m/s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.16</td>
<td>0.67</td>
<td>0.06</td>
<td>0.51</td>
<td>-0.30</td>
<td>0.70</td>
<td>0.20</td>
<td>0.57</td>
<td>-0.04</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For estimating sampling error and confidence interval of evaluation parameters (FB and AFB), blocked bootstrap data were used (bootstrap meteorology data were used for model run and bootstrap SO$_2$ observation data were used for comparison). The results are show in Table 2. The composite performance measures are also listed in the table. Either from averaged absolute fractional bias for 1-hour concentration, for 24-hour concentration or from
composite performance measure, it seems that AERMOD (except Case B) is better than HPDM. Among them, Case A is the best overall.

Figure 1 displays the absolute fractional bias for 1-hour average, 24-hour average and the grand composite (right panel) among two averaging periods with 95% bootstrap confidence bounds. It can be seen that AERMOD model is better than HPDM overall. Case A of AERMOD is the best. Case C of AERMOD is also not too bad.

Table 2. Composite performance of HPDM and AERMOD models

<table>
<thead>
<tr>
<th></th>
<th>1-hour average</th>
<th>24-hour average</th>
<th>Composite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AFB</td>
<td>σ_{AFB}</td>
<td>AFB</td>
</tr>
<tr>
<td>HPDM</td>
<td>0.79</td>
<td>0.04</td>
<td>0.64</td>
</tr>
<tr>
<td>AERMOD Case A</td>
<td>0.76</td>
<td>0.03</td>
<td>0.48</td>
</tr>
<tr>
<td>Case C</td>
<td>0.75</td>
<td>0.03</td>
<td>0.51</td>
</tr>
<tr>
<td>Case D</td>
<td>0.78</td>
<td>0.03</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Table 3 lists the comparison measures between HPDM and three cases of AERMOD: Case A, Case C and Case D. It seems that again Case A of AERMOD is the best.

Table 3. Comparison measures between HPDM and AERMOD

<table>
<thead>
<tr>
<th></th>
<th>HPDM - Case A</th>
<th>HPDM - Case C</th>
<th>HPDM - Case D</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>0.09</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>σ_{M}</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Figure 2 shows the difference in absolute fractional bias between HPDM and three cases of AERMOD with 95% confidence bounds. For 1-hour averages, Case C of AERMOD is performing better statistically (the difference is larger). For 24-hour averages, Case A is better. For composite performance, Case A is better overall. However, noticing that all confidence bounds are overlap 0, this indicates that there does not exist a statistically significant difference between HPDM and any case of AERMOD at 95% confidence level. However, there does exist significant difference between HPDM and Case A, HPDM and Case C at 67% confidence level because the confidence bounds do not overlap zero at this confidence level.

5 Conclusions

From this evaluation, it can be find that, generally speaking, Case A of AERMOD model is better than other cases and HPDM model; Case C is another applicable choice. HPDM model is seriously under-predicted on lower velocity situation of stable conditions and a little bit over-predicted on unstable
Figure 1. Performance composition between HPDM and AERMOD with 95% bootstrap confidence bounds.

Figure 2. Difference in performance between HPDM and AERMOD with 95% bootstrap confidence bounds.
conditions. While Case A of AERMOD model is a little bit under-prediction on both stable and unstable conditions. There does exist significant difference between HPDM and Case A, HPDM and Case C at 67% confidence level. But there does not exist a statistically significant difference between HPDM and any case of AERMOD at 95% confidence level.

References


