High-resolution modeling and evaluation of ozone air quality of Osaka using MM5-CMAQ system

SHRESTHA Kundan Lal*, KONDO Akira, KAGA Akikazu, INOUE Yoshio

Division of Sustainable Energy and Environmental Engineering, School of Engineering, Osaka University, Osaka 5650871, Japan.
E-mail: kundan@ea.see.eng.osaka-u.ac.jp

Received 06 June 2008; revised 10 September 2008; accepted 23 September 2008

Abstract

High-resolution modeling approach is increasingly being considered as a necessary step for improving the monitoring and predictions of regional air quality. This is especially true for highly urbanized region with complex terrain and land-use. This study uses Community Multiscale Air Quality (CMAQ) model coupled with MM5 mesoscale model for a comprehensive analysis to assess the suitability of such high-resolution modeling system in predicting ozone air quality in the complex terrains of Osaka, Japan. The 1-km and 3-km grid domains were nested inside a 9-km domain and the domain with 1-km grid covered the Osaka region. High-resolution Grid Point Value-Mesoscale Model (GPV-MSM) data were used after suitable validation. The simulated ozone concentrations were validated and evaluated using statistical metrics using performance criteria set for ozone. Daily maxima of ozone were found better simulated by the 1-km grid domain than the coarser 9-km and 3-km domains, with the maximum improvement in the mean absolute gross error about 3 ppbv. In addition, 1-km grid results fared better than other grids at most of the observation stations that showed noticeable differences in gross error as well as correlation. These results amply justify the use of the integrated high-resolution MM5-CMAQ modeling system in the highly urbanized region, such as the Osaka region, which has complex terrain and land-use.

Key words: air quality model evaluation; ozone; Community Multiscale Air Quality; MM5; statistical analysis

DOI: 10.1016/S1001-0742(08)62341-4

Introduction

Urban atmospheric pollution is a worldwide problem that has become increasingly severe with the rise in global urbanization. We have been facing the challenging task of attaining the air quality standards for air pollutants such as ozone. With the development of third-generation air quality models (AQM), we have a great scope of using these models to guide the development of emission control strategies in non-attainment areas. These models have been tested against very limited observational data for limited numbers of episodes in Japan. Hence, a more systematic and comprehensive model evaluation is needed to assess model reliability.

Air quality in megacities is of great concern because of continuing industrialization and migration toward urban centers. Besides health and social problems, air pollution has serious regional and global environmental ramifications (Akimoto, 2003). In Japan, two big metropolitan cities, Tokyo and Osaka, are the typical cases that show the urgency required for exhaustive study and analysis of air pollution in these regions. In Osaka, the largest metropolitan area in western Japan, photochemical air pollution is of crucial concern (Itano et al., 2006). It is possible to match key characteristics of an ozone episode through the manipulation of model input parameters in ways that may be somewhat arbitrary. The result in such cases indicates that the model can match a particular episode quite well but has poor predictive power when applied to other episodes (Sailor and Dietsch, 2007). Although this approach is quite resource intensive to implement for a large number of places and episodes, the use of third-generation air quality models, coupled with state-of-art meteorological models, using parallel cluster machines proves to be practically beneficial.

Regional pollutant and precursor transport, sea breeze, and heat island effect play important roles in the generation of complex wind patterns in the highly irregular mountainous terrains of Osaka. Palau et al. (2005) and Fenger (2003) have suggested that in a complex-terrain coastal site, because of the strong effect of the meteorological interactions between the different scales on the integral advection and the turbulent dispersion of pollutants, using an inadequate scale to solve the meteorology can result in a very big gap in the simulation of lower-layer pollutant behavior at urban scales.

Globally speaking, many 3-D AQMs like Community Multiscale Air Quality (CMAQ) (Byun and Ching, 1999), CAMx, SAQM, and so on have been coupled with some
prominent mesoscale models like MM5 (Dudhia et al., 2005) and RAMS (Pielke et al., 1992). The different aspects of modeling ozone, its precursors, particulate matters, etc. have been investigated with a strong emphasis on the evaluation of both the AQMs and the mesoscale models that drive them (Hogrefe et al., 2001a, 2001b; Zhang et al., 2006b, 2006c; Biswas et al., 2001; Palau et al., 2005).

In Japan, AQMs coupled with meteorological models have had a few applications (Uno et al., 1996). An example is the use of RAMS (Regional Atmospheric Modeling System) meteorological model with CMAQ (Community Multiscale Air Quality) model (Zhang et al., 2004, 2006a; Sugata et al., 2001). There has been some studies regarding the comparison of different meteorological models coupled with CMAQ (Hara et al., 2005), but the extensive simulations in a complex and urban geographical region over a long period has not been carried out. We cannot only rely on episode data sets to describe the local meteorology and ozone problems. Because we can get a better performance of the model on longer time scales, modeling periods should be longer than the duration of a single episode to increase confidence in the regulatory modeling process (Hogrefe et al., 2001a). The grid resolution also significantly influences the chemistry and transport processes affecting the ozone air quality so that the choice of grid size should be carefully selected for the ozone simulation purposes (Jang et al., 1995a, 1995b; Arunachalam et al., 2006; Jimenez et al., 2006). This study used CMAQ air quality model coupled with MM5 mesoscale model for a comprehensive analysis to assess the suitability of this high resolution modeling system in predicting ozone air quality in the complex terrains of Osaka, Japan. A Pentium-4 Linux cluster containing 9 nodes is used for the various simulations. Thus the simulated ozone concentrations are evaluated against the observed data by means of different quantitative statistical metrics using performance criteria set for ozone. The results are presented and the suitability of using the integrated high-resolution MM5-CMAQ modeling system for simulating ozone air quality for a highly urbanized region with complex terrain and land-use is discussed.

1 Methods

1.1 Modeling region

The Osaka prefecture region was selected as the target area for evaluating the MM5-CMAQ system (Fig. 1). The area around Osaka was used as the finest domain area with a grid size of 1 km (Domain 3). This domain was nested inside a coarser domain with a grid size of 3 km (Domain 2). These domains were further nested in the coarsest domain of 9 km grid size covering nearly all of Japan (Domain 1).

The MM5 domains has 118 × 118, 118 × 118, and 91 × 91 grids for Domain 1, 2, and 3, respectively. For CMAQ model, the boundary was trimmed in all the three domains to minimize the effect of boundary effects from the meteorological models. The CMAQ domains have 105 × 105, 105 × 105, and 78 × 78 grids for Domain 1, 2, and
the 9-km domain (Domain 1) and 3-km domain (Domain 2), the national emission data with the spatial resolution of 0.125° by 0.0833° were used. For 1-km domain (Domain 3), the data covering the Kansai region with 0.0125° by 0.00833° resolution were used. These emission data are available for July, 2002. Another feature of this data is the availability of weekday and weekend mobile emission data. Similarly, the biogenic emission has also been incorporated as the average of July.

1.4 Air quality modeling

Community Multi-scale Air Quality (CMAQ) model is a Eulerian-type air quality model that simulates concurrently the atmospheric and land processes affecting the transport, transformation, and deposition of air pollutants and their precursors, on both regional and urban scales (Byun and Ching, 1999). It is based on a “one atmosphere” philosophy that tackles the interactions not only among multiple atmospheric pollutants, but also between regional and urban scales. CMAQ is configured with the chemical mechanism of Statewide Air Pollution Research Center (SAPRC-99) mechanism. CMAQ version 4.5 was used as the air quality model in the present work. For the post-processing of the CMAQ results, Climate Data Analysis Tools (CDAT) software was used.

1.5 Observation stations

For the evaluation of simulated meteorological and air quality variables, the observation data were obtained from the “Air Pollution Continuous Monitoring Network Data Files” provided by Environmental Pollution Control Center, Osaka Prefecture. The air pollution monitoring network of Osaka covers the Osaka prefectural area with 20 air monitoring stations, and 11 exhaust monitoring stations. Similarly, there are 94 stations distributed across various cities of Osaka. Figure 2 shows the station points used in this study for the purpose of evaluating the model performances.

1.6 Statistical metrics

In order to evaluate the simulation of ozone air quality in the MM5-CMAQ system, several quantitative statistical performance measures were used collectively in the assessment. These statistical measures are mathematically expressed Eqs. (1)–(5) (Yu et al., 2006; Zhang et al., 2006b). The traditional metrics like normalized mean bias (NMB) and normalized mean error (NME) have been used to evaluate the model performance for simulating ozone.

Fig. 2  Observation stations of Osaka Prefecture. The station codes are provided for each observation station according to the codes given in Osaka Prefecture’s online database.
and its precursors. NMB and NME have the advantage of avoiding inflated values since they employ relative difference method of normalization. Similarly, mean absolute gross error (MAGE) is useful for observing the spread of the departure between the simulated and observed data. Similarly, the root mean square error (RMSE) provides a quantitative estimate of the uncertainty in the predictions, as compared to a given set of observational data. All the data analyses were done in the Python programming environment, and the R statistics package (R Development Core Team, 2007) was used to calculate the statistical environment, and the R statistics package (R Development Core Team, 2007) was used to calculate the statistical

decision method of normalization. Similarly, mean absolute
error (MAGE) is useful for observing the spread of
departure between the simulated and observed data. Similarly, the root mean square error (RMSE) provides a
quantitative estimate of the uncertainty in the predictions,
as compared to a given set of observational data. All
the data analyses were done in the Python programming
environment, and the R statistics package (R Development
Core Team, 2007) was used to calculate the statistical

\[ \text{CCOF} = \frac{\sum_{i=1}^{N} (M_i - \bar{M})(O_i - \bar{O})}{\left( \sum_{i=1}^{N} (M_i - \bar{M})^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{N} (O_i - \bar{O})^2 \right)^{\frac{1}{2}}} \] (1)

Normalized mean bias (range \(-1 \sim +1\))

\[ \text{NMB} = \frac{\sum_{i=1}^{N} (M_i - O_i)}{\sum_{i=1}^{N} O_i} \] (2)

Normalized mean error (range \(0 \sim +\infty\))

\[ \text{NME} = \frac{\sum_{i=1}^{N} |M_i - O_i|}{\sum_{i=1}^{N} O_i} \] (3)

Mean absolute gross error (range \(0 \sim +\infty\))

\[ \text{MAGE} = \frac{1}{N} \sum_{i=1}^{N} |M_i - O_i| \] (4)

Root mean square error (range \(0 \sim +\infty\))

\[ \text{RMSE} = \left( \frac{1}{N} \sum_{i=1}^{N} (M_i - O_i)^2 \right)^{\frac{1}{2}} \] (5)

where, \(M_i\) and \(O_i\) are measured and observed data at \(i\) respectively, and \(N\) is the total number of points and time steps. Similarly, \(\bar{M}\) and \(\bar{O}\) refer to the average values of \(M_i\) and \(O_i\) respectively.

2 Results and discussion

Among the different stages involved in the MM5-CMAQ simulation, the meteorological modeling, or MM5 simulation, and CMAQ simulation account for most of the simulation time. Using a Pentium-4 Linux cluster containing 9 nodes, MM5 simulation took 5 d of user time, and CMAQ took 45 h, 45 h, and 5 d of user time for domain 1, 2, and 3, respectively. Since the time step used in the finest domain (Domain 3) was 3 s for MM5, user time in the MM5 simulation is comparable to the Domain 3 simulation in CMAQ. As compared to the single-processor simulation, the cluster simulation took nearly 1/5th of the user time of single-processor simulation.

2.1 Meteorological variables

The meteorological variables such as air temperature, solar radiation and wind speed are important for accurate simulation of surface ozone. The near surface air temperature was well-captured by the MM5 simulation with very high correlation coefficient (approx. 0.8), low mean absolute gross error (approx. 1.5 K), and a faithful reproduction of diurnal amplitude (Fig. 3a). The solar radiation was also well-captured with generally low radiation periods on cloudy days (Fig. 3d). The low radiation days like 10 July, 2002 have been faithfully reproduced by the present model. The horizontal components of surface wind were also well simulated with good correlation coefficient (approx. 0.65) and mean absolute gross error (approx. 1.5 m/s). Since four dimensional data assimilation (FDDA) was not used in the modeling, a significant loss in the contribution of diurnal variations in wind was avoided, and the local wind circulations were fairly well simulated at fine resolution of 1 km (Fig. 3b). The wind direction was also well-captured as compared to the observed values (Fig. 3c). It should be noted that the observed wind direction is available at only 16 equally spaced discrete directions.

2.2 Quantitative analysis of ozone

The temporal variation in near surface ozone concentration is well reproduced by the MM5-CMAQ simulation (Fig. 3e). The prevalent low ozone periods in the first half of the simulation period and the relatively high ozone periods in the second half of the simulation are fairly well predicted. The diurnal variation of surface ozone is the most faithfully reproduced characteristics of the ozone time series. It is quite clear from the time series of solar radiation (Fig. 3d) that the second half of the simulated month has higher incoming solar radiation, and hence the diurnal ozone is generally higher than the first half of the simulated month. Thus, solar radiation is one of the factors affecting ozone production. Importantly, the correlation between the meteorological variables like air temperature (Fig. 3a) and wind speed (Fig. 3b) with ozone can also be clearly inferred from the results. Having noted that the second half of the simulated period does not have rainy days, the high wind periods are found to be typically marked by low ozone episodes such as on 25 and 26 of July, 2002. These two days are also conspicuously marked by wind direction between northerly and easterly winds (Fig. 3c). The synoptic wind with high speed prevalent in these days seems to have contributed toward low ozone. On the other hand, low wind periods such as on 23 and 29 of July, 2002 are characterized by expected high ozone days. On these days, weak synoptic winds and sea breeze effect contribute towards high ozone periods. Overall, the ozone time series are well predicted by the MM5-CMAQ system along with the meteorological variables like air temperature and wind speed that affect the amplitude and variation of ozone time series.

Table 1 shows the statistical measures for ozone with cut-off value of 40 ppbv using the hourly ozone data, and also, the metrics for the daily maxima of ozone. In the first
Fig. 3  Time series of near surface air temperature (a), wind speed (b), wind direction (c), solar radiation (d), and ozone (e) at station No. 107 (Kokusetsu-osoaka), which is a typical example of a station at the city center of Osaka.

part, pairs of data with the observation values less than 40 ppbv are not used to calculate the statistical metrics. In these kinds of evaluation, the observation-simulation pairs of data can be paired or unpaired in space and time. For the most stringent evaluation and analysis we can use the paired data over the whole length of simulation time and over all the stations used in the evaluation. Therefore, paired data have been used in this study. Although USEPA (2007) has not recommended any fixed criteria for model performance in its latest report, performance criteria
No. 6  High-resolution modeling and evaluation of ozone air quality of Osaka using MM5-CMAQ system

Table 1  Quantitative statistical performances for surface ozone with 40 ppbv cut-off concentration, and daily maximum ozone

<table>
<thead>
<tr>
<th>Metrics</th>
<th>O$_3$ (40 ppbv cut-off)</th>
<th>Daily maximum O$_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>D1</td>
<td>D2</td>
</tr>
<tr>
<td>Count</td>
<td>4559</td>
<td>4559</td>
</tr>
<tr>
<td>CORCF</td>
<td>0.39</td>
<td>0.40</td>
</tr>
<tr>
<td>NMB</td>
<td>–0.03</td>
<td>–0.06</td>
</tr>
<tr>
<td>NME</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>RMSE</td>
<td>23.72</td>
<td>22.94</td>
</tr>
<tr>
<td>MAGE</td>
<td>17.54</td>
<td>17.26</td>
</tr>
</tbody>
</table>

recommended in its previous reports are used in this study for the quantitative analysis of model performance. For 40 ppbv cutoff ozone, USEPA criteria limit for NMB is $\leq 15\%$, and $\leq 30\%$ for NME. Since the NMB and NME results (Table 1) are well within these criteria limits, the present modeling system can be considered to possess a high predictive performance and skill. For 40 ppbv cut-off ozone, all errors show better predictions in Domain 3 than the coarser domains. NMB shows a small amount of underprediction in the ozone simulation. This small underprediction in 40 ppbv cutoff ozone suggests that the daytime ozone production chemistry may not be sufficient but the overall overprediction in total ozone time series (not shown in graph) suggests some overprediction in the nighttime ozone production. Both RMSE and MAGE reduce by nearly 1 ppbv when 1-km grid (Domain 3) is used instead of 9-km grid (Domain 1). On the other hand, normalized statistics, NMB and NME, and correlation show that the skill and performance of the MM5-CMAQ system do not appreciably alter in the 9-km, 3-km, and 1-km ozone air quality. This implies that the overall performance of the present modeling system is acceptable in terms of the aforementioned USEPA criteria, and also, the 1-km grid reduces the magnitude of the overall error with a small amount of underprediction at daytime, although the nighttime ozone production chemistry has still scope for improvements to reduce overpredictions at ozone minima.

2.3 Quantitative analysis of daily maxima of ozone

For the accuracy in the simulation of daily maximum surface ozone, the criteria limit for NMB is set at $\leq 20\%$. In this respect, the 1 km grid (Domain 3) result (Table 1) is within this limit, which is an improvement over the 9-km grid result. Similarly, NME, RMSE and MAGE also show some improvement when the grid size progressively reduces from 9-km to 1-km. Overall, the daily maximum surface ozone is found to be improved in the high-resolution 1-km grid, if correlation, NMB, NME, RMSE and MAGE are taken into consideration.

The highest difference observed in MAGE between Domain 1, 2, and 3 is approximately 3 ppbv. In order to analyze the stations with appreciable difference in MAGE, the differences greater than 2 ppbv are shown in Fig. 4, and here, the 1-km grid (Domain 3) is found to be better than the coarser grids in almost all the observation stations (11 out of 14 stations). The differences in MAGE values among the grid resolutions are generally attributed to topography and elevation, large NOx point emission sources, and representation of NOx titration. In this case also, most of the stations showing improvement in MAGE with the grid resolution seem to be particularly influenced by the NOx point emission sources in the Sakai industrial region (Fig. 2), which is one of the major parts of the Osaka Bay region.

The largest improvement in correlation coefficient at any observation station is approximately 0.165 (Fig. 5). Neglecting very small values, the 1-km grid (Domain 3) shows the best correlation in most stations (8 out of 13 stations). Even though the improvements are small, the overall picture suggests a better skill of the model to simulate the ozone air quality. The observation stations
The daily maximum of ozone is systematically overpredicted in the low ozone periods in the first half month (Fig. 6), which contributes to the overall positive bias in the daily maxima. The 1-km grid seems to especially fare better than the coarser grids in these low ozone periods. As expected, the bias increases systematically in the coarser grids, which may be partially attributed to the higher dilution of NO\textsubscript{x} emissions in the coarser grids.

### 2.4 Ozone precursors

The precursors of ozone, NO\textsubscript{2} and NO show a tendency toward overall underprediction in the MM5-CMAQ system used in this research. NO\textsubscript{2} has sufficiently acceptable levels of NMB values of nearly –0.05 although the progressive increase in grid resolution does not decrease the absolute gross error appreciably suggesting that there is still scope for improvement in the emission modeling of NO\textsubscript{x} at high resolution of around 1-km grid.

### 3 Conclusions

An integrated MM5-CMAQ high-resolution modeling system was tested for simulating ozone air quality for a highly urbanized region with complex terrain and land-use for a month long period in July, 2002. The Osaka prefecture region was selected as the target area for evaluating the MM5-CMAQ system. The area around Osaka was used as the finest domain area with a grid size of 1 km (Domain 3). The high-resolution Japanese GPV-MSM meteorological data were used in running MM5 for deriving meteorological parameters for the three nested domains and then CMAQ model was run using the meteorological input from MM5 and the emission data from the Japanese emission database.

From this study, we can conclude that these results demonstrate the practical applicability as well as the validity of using high-resolution MM5-CMAQ simulation for
studying air quality of a highly urbanized region with complex terrain and land-use. The results of correlation, further lead us to conclude that the noticeable improvements in the prediction of daily ozone maxima generally occurs in the finest domain having grid size of 1 km.

Acknowledgments

We are grateful to Japan Clean Air Program (JCAP) of Japan Petroleum Energy Center (JPEC), Tokyo, Japan for providing the Japanese emission data used in this work.

References


Itano Y, Wakamatsu S, Hasegawa S, Ohara T, Sugata S, Hayasaki M et al., 2006. Local and regional contributions to springtime ozone in the Osaka metropolitan area, estimated from aircraft observations. Atmospheric Environment, 40(12):

2117–2127.


