Meteorological Modeling Protocol
For Application to PM2.5/Haze/Ozone Modeling Projects

Kirk Baker

December 12, 2004

Lake Michigan Air Directors Consortium
Midwest Regional Planning Organization
Des Plaines, Illinois

INTRODUCTION

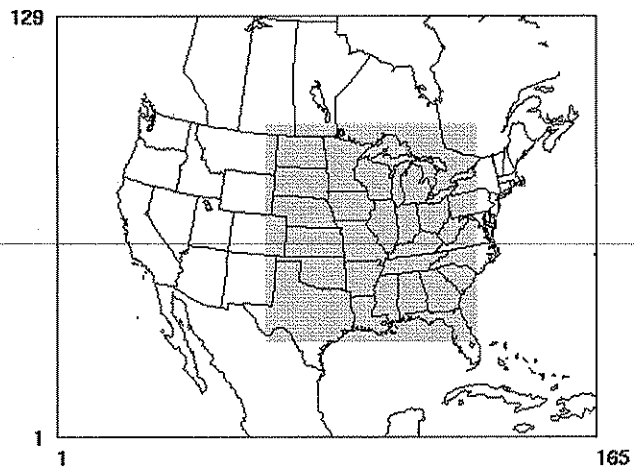
The purpose of this document is to outline the configuration and application of MM5 to support photochemical and emissions modeling projects. All information provided in this document is relevant to NCAR's 5th generation Mesoscale Model version 3.6.3 Dudhia, 1993 and Grell et al, 1994). The computing platform supported by LADCO/Midwest RPO is the Red Hat version 7.X Linux operating system and the Portland Group Fortran compiler. MM5 consists of the Mesoscale model MM5 and a suite of pre-processors including PREGRID, REGRIDDER, RAWINS, LITTLE R, INTERPF, INTERPX, and TERRAIN.

The model parameterizations and physics options outlined in this document are based on a series of sensitivity runs that indicate an optimal configuration for the Upper Midwest (Johnson, 2003). The model configuration and parameterizations outlined in this document describe recent MM5 applications. Evolving science in meteorological modeling and photochemical modeling necessitate that this document change to reflect the state of the science.

The annual 2002 36 km MM5 simulation was completed by Matthew Johnson at Iowa DNR. The 36/12 km 2-way nested simulation for the Summer of 2002 was conducted jointly by Steven King at Illinois EPA and Kirk Baker at LADCO.

TERRAIN

The TERRAIN processor defines the horizontal grid of the MM5 application. The 24 category USGS 10 minute (~19 km) data is used for the National RPO 36-km domain, and 5 minute (~9 km) data for 12-km domains. The National RPO grid is a Lambert conic projection centered at coordinates -97, 40 with first and second true latitudes at 33 and 45 degrees (See Figure).



The 36 km grid contains 165 x 129 km grid cells and the 12 km has 193 x 199 grid cells (See Table 1). The 12 km grid is two-way nested within the mother grid to allow fine grid feedback into the coarse grid. Additional options are set to allow generation of data to support the Pleim-Xu land surface module. Variables LSMDATA and IEXTRA are both set equal to TRUE.

Domain ID	Grid	X Cells (East-West)	Y Cells (North-South)	Cell Size (km)	Mother Domain ID	Lower Left X,Y of Nest
1	National RPO	165	129	36	1	1, 1
2	Upper Midwest	193	199	12	1	66, 30

PREGRID

The PREGRID processor converts meteorological analyses data such as NCEP or Eta to an intermediate data format that the REGRIDDER processor can utilize. Eta/AWIP 3D and SF analyses data (ds609.2) will be used to initialize the REGRID processor. Snow cover will be estimated from water equivalent snow depth. The input analyses will be processed 3 hourly (10,800 seconds). The AWIP grib definition tables will be used to map Eta data into MM5. The ETA skin temperature is used as the source of sea surface temperature. The Eta analysis files with the extension ".tm12" are not used since they are the "cold start" global analysis files.

REGRIDDER

The REGRIDDER processor takes the data extracted from analyses fields and interpolates the data to user specified pressure levels and to the user specified horizontal grid.

LITTLE R

The RAWINS and LITTLE R processors perform objective analysis on the output from REGRIDDER using surface and upper air observation data. Since these observations are incorporated into the Eta analysis fields this step is considered redundant. Sensitivity tests where Eta 3 hourly analysis was used to initialize with and without RAWINS objective analysis showed no difference in model performance (Baker, 2002).

Even though this step is redundant, LITTLE R is applied to enable surface nudging of soil moisture and temperature in the Pleim-Xu land surface module. NCEP ADP surface (ds 464.0) and upper air (ds 353.1 and ds 353.4) data are the appropriate data to input into LITTLE R and/or RAWINS.

k(MM5)	sigma	press.(mb)	height(m)	depth(m)
34	0.000	10000	14662	1841
33	0.050	14500	12822	1466
32	0.100	19000	11356	1228
31	0.150	23500	10127	1062
30	0.200	28000	9066	939
29	0.250	32500	8127	843
28	0.300	37000	7284	767
27	0.350	41500	6517	704
26	0.400	46000	5812	652
25	0.450	50500	5160	607
24	0.500	55000	4553	569
23	0.550	59500	3984	536
22	0.600	64000	3448	506
21	0.650	68500	2942	480
20	0.700	73000	2462	367
19	0.740	76600	2095	266
18	0.770	79300	1828	259
17	0.800	82000	1569	169
16	0.820	83800	1400	166
15	0.840	85600	1235	163
14	0.860	87400	1071	160
13	0.880	89200	911	158
12	0.900	91000	753	78
11	0.910	91900	675	77
10	0.920	92800	598	77
9	0.930	93700	521	76
8	0.940	94600	445	76
7	0.950	95500	369	75
6	0.960	96400	294	74
5	0.970	97300	220	74
4	0.980	98200	146	37
3	0.985	98650	109	37
2	0.990	99100	73	36
1	0.995	99550	36	36
0	1.000	100000	0	--SURF--

INTERPF

The INTERPF processor takes the REGRIDDER/LITTLE R output that is at standard pressure levels and interpolates that data to the vertical grid defined by the user. The vertical grid is defined in terms of sigmas, where 1 is the surface and 0 is the top of the model atmosphere. The top of the MM5 simulation is 100 millibars, which is approximately 15 kilometers above ground level.

The vertical atmosphere is resolved to 34 layers, with thinner layers in the planetary boundary layer. This is to capture the important diurnal variations in the boundary layer while also having layers in the upper troposphere to try and resolve convective activity. Output from the INTERPF processor is ready for input into MM5.

INTERPX

The INTERPX processor is used to extract the soil temperature and soil moisture data from MM5 output files and overwrite the soil temperature and moisture fields on the MMINPUT file for the next 5 day simulation block. This allows soil moisture and temperature to be carried over to subsequent modeling simulations.

For example, to simulate 20 days in 4 blocks of 5 days, the first block of 5 days would use the standard MMINPUT to run MM5, and the subsequent 3 blocks of 5 days would take the MM5 output and extract soil temperature and moisture data for the next 5 day block.

This option was only used for the 36 km annual simulation of 2002. This option is no longer recommended since it has been shown to introduce a cold bias for the temperature field, particularly in the winter months (Olerud, 2003).

MM5

The output from INTERPF, LITTLE R, and TERRAIN processors were used to run MM5. These files must be in the "./MM5/Run" directory and have the generic filenames given directly out of these processors. 3D analysis nudging for the wind field, temperatures, and moisture were applied above the boundary layer only. Analysis nudging was not performed on the rotational wind field. In addition, the observation nudging flag was turned off. This type of nudging is appropriate usually when you have a very dense set of observation data from a field study, which this application lacked. The default nudging weighting factors were used for all simulations: 2.5×10^{-4} for wind fields and temperatures and 1.0×10^{-5} for moisture fields.

Table 3

Configuration	36km and 12km Domains
Explicit Moisture	Mixed Phase (Reisner I)
Cumulus	Kain-Fritsch 2
PBL	Pleim-Chang (ACM)
Radiation	RRTM
Multi-Layer Soil Model	Pleim-Xu
Shallow convection	No
4-D Data Assimilation	Analysis nudging on above PBL
Moist Physics Table	No

Table 3 outlines the model configuration used for MM5 modeling up to the date of this document. All simulations use the mixed phase moisture rather than simple ice so that all four phases of water will be explicitly output by MM5. This is important since the photochemical model is applied for an annual basis and correctly characterizing the phase of water is important for several physiochemical processes.

Atmospheric radiation is calculated every 15 minutes in the model. Vertical moisture and temperature advection are set to use linear interpolation. Other important variables switched to ON include: moist vertical diffusion in clouds, temperature advection using potential temperature, diffusion using perturbation temperature, 3D coriolis force, and upper radiative boundary condition. Sea surface temperature and snow cover are set to vary with time.

The Pleim-Xu land surface module requires that 3 additional variables be set in the MM5 deck: ISMRD, NUDGE, and IFGROW. ISMRD is set to use soil moisture fields from the Eta analyses. NUDGE is set to nudge soil moisture data to the analyses fields. IFGROW is set to option 2, which takes vegetative growth into account based on vegetative fraction data from the TERRAIN file.

MODEL EXECUTION

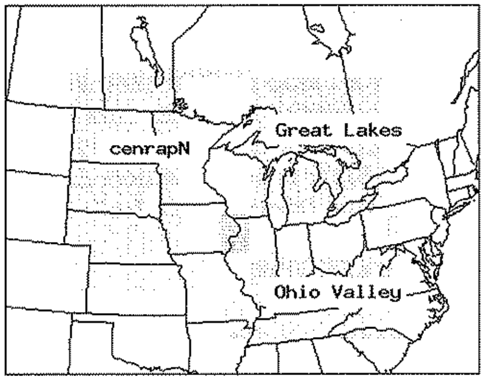
MM5 was executed in 5 day blocks (7200 minute simulation) with a 90 second time step. Model results are output every 60 minutes and the model output files are written out (i.e. split) every 24 hours to accommodate post-processing utilities. The start and end times are 12Z. Only 4 days from each block will be used for input to a photochemical model: 5Z to 5Z, which means the first 17 hours of the MM5 simulation are ramp-up and there are 5 hours at the end of the simulation that are not used.

The 2002 annual simulation was initiated at 12Z December 16, 2001 and was run through 12Z January 1, 2003. The 5 day blocks are evenly divided between these 2 dates. Alternatively, the following table illustrates how each month from the 2001 summer and 2003 annual simulations will be modeled using 5 day blocks. This standardized approach to simulating each month reduces post-processing burden.

MM5 Run	start date	start time	end date	end time	total days	total hours	ramp up hours	extra hours
1	MM/01/YY	12Z	MM/06/YY	12Z	5	120	17	5
2	MM/05/YY	12Z	MM/10/YY	12Z	5	120	17	5
3	MM/09/YY	12Z	MM/14/YY	12Z	5	120	17	5
4	MM/13/YY	12Z	MM/18/YY	12Z	5	120	17	5
5	MM/17/YY	12Z	MM/22/YY	12Z	5	120	17	5
6	MM/21/YY	12Z	MM/26/YY	12Z	5	120	17	5
7	MM/25/YY	12Z	MM/30/YY	12Z	5	120	17	5
8	MM/29/YY	12Z	NEXTMM/02/YY	12Z	5	120	17	5
PCM Run	start date	start time	end date	end time	total days	total hours		
1	MM/02/YY	5Z	MM/06/YY	5Z	4	97		
2	MM/06/YY	5Z	MM/10/YY	5Z	4	97		
3	MM/10/YY	5Z	MM/14/YY	5Z	4	97		
4	MM/14/YY	5Z	MM/18/YY	5Z	4	97		
5	MM/18/YY	5Z	MM/22/YY	5Z	4	97		
6	MM/22/YY	5Z	MM/26/YY	5Z	4	97		
7	MM/26/YY	5Z	MM/30/YY	5Z	4	97		
8	MM/30/YY	5Z	NEXTMM/01/YY	5Z	4	97		

MODEL PERFORMANCE

The performance of MM5 will be analyzed qualitatively by comparing output surface fields of temperature, winds, convective activity, and cloud cover to 12 hourly UNISYS surface weather maps and satellite cloud cover images. Vertical sounding plots of temperature, humidity, wind speed, and wind direction will be analyzed for select upper air stations in the Midwest. Vertical sounds plots will show model predictions against Forecast Systems Laboratory (FSL) / National Climatic Data Center (NCDC) Radiosonde data archive (RAOBS) upper air data at 0Z and 12Z.



Model performance will be assessed quantitatively with the METSTAT tool from Environ (Emery et al, 2001). The metrics used to quantify model performance include mean observation, mean prediction, bias error, gross error, root mean square error (including systematic and unsystematic components), and index of agreement. These metrics will compare model predictions to Techniques Data Laboratory U.S. and Canada surface hourly observations (NCAR dataset ds472.0).

The MM5 model outputs approximately 15 meter predictions while observations are at 10 meters. METSTAT applies micro-meteorological adjustments to the MM5 estimates to approximate 10-m values. MM5 outputs near-instantaneous values (90 second time step) as opposed to the values with longer averaging times taken at monitor stations, so that should be kept in consideration when interpreting model performance metrics and making qualitative comparisons to satellite maps.

Model performance metrics will be applied to sub-regions of the Upper Midwest, meaning the metrics are hourly spatial averages of multiple monitor locations. This will be done to gain a better understanding of MM5 performance for 2 geographically and meteorologically diverse regions: Great Lakes and Ohio Valley. All metrics are calculated for all sites within the specified model performance region for an hourly and daily time period (0Z to 23Z). Mean wind direction is estimated by averaging the U and V wind vector components and converting those averages to an average wind direction in compass degrees.

MPE Region	**Great Lakes**	**Ohio Valley**
X Coordinate SW (km)	200	450
X Coordinate NE (km)	1500	1600
Y Coordinate SW (km)	100	-472
Y Coordinate NE (km)	1200	300
~ NX 36km cells	36	32
~ NY 36km cells	31	22
Surface Met Stations (jan 2000)	273	179
Surface Met Stations (july 2001)	283	210

Additional analysis of rainfall will be done on a seasonal basis. Rainfall totals in each grid cell by season will be compared to the corresponding seasonal totals at observation sites.

Annual simulations present a challenge in terms of adequately assessing model performance so photochemical modelers will know the strengths and weaknesses of the meteorological inputs. A report will be compiled with the following elements: select vertical sounding plots at Upper Midwest sounding stations, daily metrics output by METSTAT, rainfall analysis results, and any qualitative comparisons between model output and UNISYS plots.

PERFORMANCE METRICS

The *bias error* (bias) is the degree of correspondence between the mean prediction and the mean observation, with lower numbers indicative of better performance. Values less than 0 indicate under-prediction. The *gross error*, or mean absolute error, is the mean of the absolute value of the residuals from a fitted statistical model. Lower numbers indicate better model performance.

Root Mean Square Error (RMSE) is a good overall measure of model performance. The weighting of (*prediction-observation*) by its square tends to inflate RMSE, particularly when extreme values are present. With respect to a good model the root mean square error should approach zero. RMSE can be divided into a systematic and unsystematic component by least-squares regression. Since differences described by systematic RMSE can be described by a linear function, they should be relatively easy to dampen by a new parameterization of the model. Unsystematic RMSE can be interpreted as a measure of potential accuracy or noise level (Emery et al, 2001). With respect to a good model the systematic difference should approach zero while the unsystematic difference approaches RMSE.

Index of Agreement is a relative measure of the degree of which predictions are error-free. The denominator accounts for the model's deviation from the mean of the observations as well as to the observations deviation from their mean. It does not provide information regarding systematic and unsystematic errors. The index of agreement approaches one when model performance is best.

POST PROCESSING FOR PHOTOCHEMICAL MODELS

The meteorological fields output by MM5 are prepared for use by the photochemical models with processing utilities. These programs translate certain meteorological parameters from the MM5 grid to the photochemical grid. Additionally, these processors must estimate parameters that are not explicitly output by MM5. Cloud cover is not output by MM5 and must be diagnosed based on moisture ratios. Vertical mixing is based on vertical diffusivity coefficients. This is a key variable not output by MM5 using the configuration outlined in this protocol.

The vertical diffusivities are calculated inside the CMAQ model based on the PBL height output by MM5. CAMx4 and REMSAD have vertical diffusivities based on the O'Brien 1970 vertical diffusivity algorithm. This scheme takes the PBL height output by MM5 and creates a well-mixed atmosphere inside the PBL. The minimum vertical diffusivity coefficient for CAMx4, REMSAD, and CMAQ are 0.1 m^2/s. The CMAQ coefficient is established (and modified from the default of 1.0 m^2/s) in the model code. A processing utility was applied to the vertical diffusivity coefficient input files for CAMx4 and REMSAD. A minimum vertical diffusivity coefficient of 1.0 m^2/s is assigned to all grid cells with an urban land use fraction up to 350 meters above ground (model layer 5). This is done to better represent the greater vertical mixing overnight in urban areas. Since the meteorological processing programs for each model not only translate data, but also diagnose certain key parameters, this step must be scrutinized to achieve optimal model results.

The vertical resolution used in MM5 consists of 34 sigma layers that represent the terrain following atmosphere up to 100 millibars. The table below displays each vertical layer in terms of sigma level, pressure (millibars), height above ground level (meters) and layer thickness (meters). The relationship to the layer structure used in the photochemical models is also shown. The photochemical model layer structure avoids layer collapsing in the lower boundary layer to better resolve the mixing depth. A compromise in the upper troposphere is met by employing layer collapsing to reduce computational effort and still maintain some upper troposphere resolution for long-range transport.

k(MM5)	sigma	p(mb)	depth(m)	k(PCM)	depth(m)
34	0.000	100	1841	16	5597
33	0.050	145	1466		
32	0.100	190	1228		
31	0.150	235	1062		
30	0.200	280	939	15	2549
29	0.250	325	843		
28	0.300	370	767		
27	0.350	415	704	14	2533
26	0.400	460	652		
25	0.450	505	607		
24	0.500	550	569		
23	0.550	595	536	13	1522
22	0.600	640	506		
21	0.650	685	480		
20	0.700	730	367	12	634
19	0.740	766	266		
18	0.770	793	259	11	428
17	0.800	820	169		
16	0.820	838	166	10	329
15	0.840	856	163		
14	0.860	874	160	9	318
13	0.880	892	158		
12	0.900	910	78	8	155
11	0.910	919	77		
10	0.920	928	77	7	153
9	0.930	937	76		
8	0.940	946	76	6	151
7	0.950	955	75		
6	0.960	964	74	5	148
5	0.970	973	74		
4	0.980	982	37	4	37
3	0.985	987	37	3	37
2	0.990	991	36	2	36
1	0.995	996	36	1	36
--SURF--	1	1000	0	--SURF--	--SURF--

It is difficult to establish an optimal vertical grid resolution for ozone and PM applications, so it should meet certain scientific criteria outlined in the WRAP modeling protocol (Tonnesen et al, 2001). The layer structure chosen for a modeling application should be capable of adequately resolving the diurnal variations in the boundary layer growth and mixing, long-range transport processes, wind shear, as well as transport to and from the free troposphere.

REFERENCES

Baker, K. MM5 Sensitivity for the National RPO Grid, Presented at the Ad-Hoc Meteorological Modeling Meeting, Des Plaines, IL, 2002.

Dudhia, J., A nonhydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front, *Mon. Wea. Rev.*, 121, 1493-1513, 1993.

Emery, C., Tai, E., Yarwood, G., Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes, Final report submitted to Texas Natural Resources Conservation Commission, 2001.

Grell, G. A.; J. Dudhia, and D. R. Stauffer, *A description of the Fifth Generation Penn State/NCAR Mesoscale Model (MM5),* NCAR Tech. Note, NCAR TN-398-STR, 138 pp., 1994.

Johnson, M., Meteorological Modeling Protocol: IDNR 2002 Annual MM5 Application, 2003.

Olerud, D., K. Alapaty, and N. Wheeler, *Meteorological modeling of 1996 for the United States with MM5*, Final report submitted to OAQPS, U. S. EPA, 2000.

Olerud, D., VISTAS Meteorological Modeling, Presented at the Ad-Hoc Meteorological Modeling Meeting, Des Plaines, IL, 2003.

Tonnesen, G., Wang, Z., Lents. J., Gemmill, D., Morris, R., Modeling Protocol and Quality Assurance Plan for the WRAP Visibility Modeling Project, 2001.