

The Effect of the CALMET Surface Layer Weighting Parameter R1 on the Accuracy of CALMET at Other Nearby Sites: a Case Study

32

Russell F. Lee

Russell Lee, *Meteorologist*, 5806 Prosperity Church Road, Suite A2-128, Charlotte, NC 28269

Jesse L. The

Lakes Environmental Software, Inc., 450 Phillip Street, Suite 2, Waterloo, ON, Canada N2L 6C7

ABSTRACT

The use of CALMET, the meteorological preprocessor for the CALPUFF air dispersion model, requires judgement on the values to be applied to various parameters. Two such parameters are those defining the relative weighting to be applied between the first guess wind field and the observations for surface (R1) and upper air (R2). The initial wind field may be optionally defined by output from a prognostic meteorological model such as MM4 or MM5. In this study, the output from the meteorological prognostic model, MM5, defines the initial wind field. CALMET is run using MM5 only (without observed data), with MM5 and surface data using various settings of the parameter R1, and without MM5 data, but adding observed upper air observations as well as surface observations. Half of the surface data sites are held as a “control” set of surface observations. These are not used as input to CALMET, but are used to compare with the CALMET output to determine, for this case, whether accounting for surface observations in varying degrees (using various settings of R1) improves or degrades the accuracy of CALMET in “predicting” the wind flow at the “control” set of surface stations. In this case study, increasing the weight given to surface observations resulted in some degradation of the accuracy of predictions of wind direction at the “control” set of surface stations.

INTRODUCTION

The U.S. Environmental Protection Agency has recently promulgated the CALPUFF¹ air quality model as a recommended model for long-range transport (50km to 200km) and on a case-by-case basis for complex wind situations. CALMET², the meteorological preprocessor for CALPUFF model, has numerous parameters that must be set by the user. The Interagency Workgroup on Air Quality Modeling (IWAQM), a joint project of the U.S. EPA, the National Park Service, the U.S. Forest Service and the U.S. Fish and Wildlife Service, has published a set of recommended values for many of these parameters³. However, the values used for several parameters are left to professional judgement. Two such parameters are those defining the relative weights given to the first guess wind field, optionally defined by a prognostic meteorological model such as MM4⁴ or MM5⁵, versus observed data, both for surface observations (the parameter R1) and for

upper air observations (the parameter R2). In this study, we evaluate the effects of the parameter R1 on the predicted wind field. The modeler must decide whether the initial field (MM5 data, in this case) or the observed data is of better quality and more representative of the wind flow over the modeling regime. If the MM5 data is judged to be the more representative, then R1 should be set small, so that the observed data has little impact on the results. On the other hand, if the observed data is judged more representative than the modeled data, R1 should be set large, so that the observed data is given more weight. There is little public information available to give the model user guidance on this matter. This case study will provide some evidence that can help the user make that judgement.

METHODOLOGY

Experiment Design

To accomplish the goal of this study, we selected six cases to compare. The first case was to run the CALMET using MM5 data only, with no observed data (Run 0). The next four cases were run for values of R1 ranging from 3 km to 100 km (Runs 1 through 4). The last case was run without MM5 data, which required the addition of upper air stations (Run A1). An area was selected that had relatively uncomplicated terrain, i.e., no major mountain ranges and no large bodies of water, large enough to encompass eight to twelve reliable surface observing stations but not so large as to exceed maximum array sizes in CALMET.

The southwest corner of the grid is at UTM northing of 4090 km, and easting of 230 km, in Zone 15. This is approximately 37° N and 96° W, along the eastern section of the boarder between Oklahoma and Kansas. An area approximately 560 km east to west by 570 km north to south was selected, which encompassed ten NWS surface observing stations. It also allowed a 3-km grid spacing without exceeding the array size limits of the extra large CALMET executable.

Six vertical layers are used in all runs with top faces at 20, 50, 100, 500, 2000 and 3300 meters above ground. The center of the lowest layer, at 10 meters, should provide a reasonable comparison with anemometer winds at the surface weather stations.

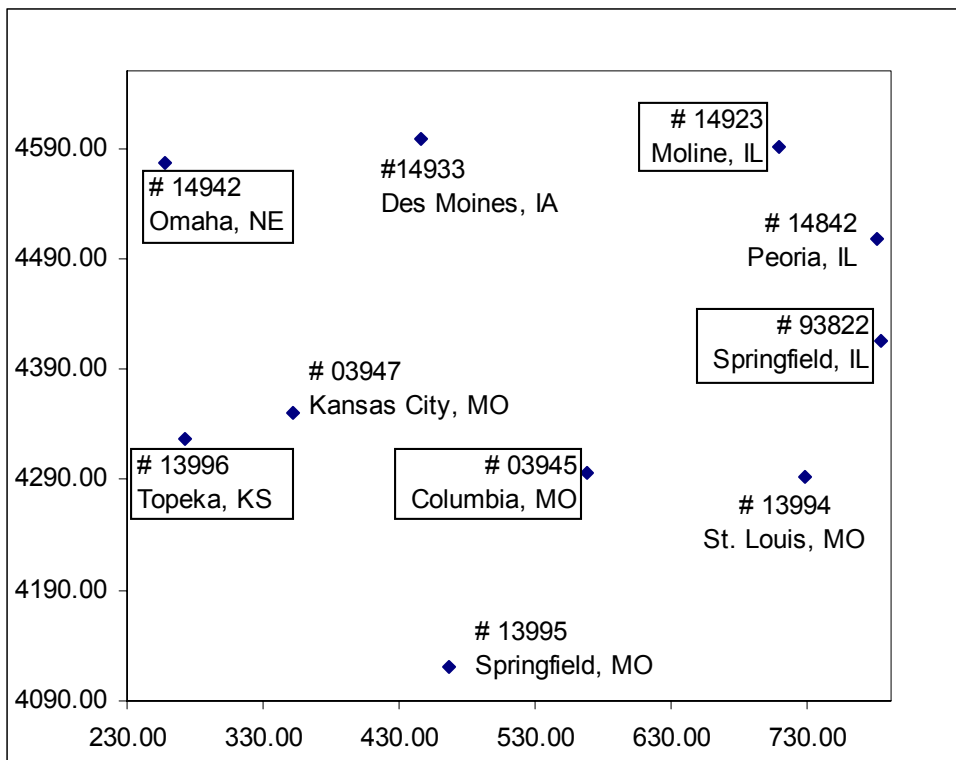
The period of April 1 through April 9, 1996 was selected for this exercise. The only criteria that were used in the selection of the dates were to avoid extreme winter and summer weather, and to have the required meteorological data available. The 1996 MM5 data output⁶ was available, as well as the surface meteorological data from the HUSWO National Climatic Data Center CDROM.⁷ This MM5 output covers the entire contiguous United States with a grid spacing of 36 kilometers for calendar year 1996.

The beta version of CALMET dated 7/11/2003 was used in this study. In addition, we used the Lakes CALPUFF View[®]⁸ interface to automate downloading of geographic and some of the meteorological data, and to provide visualization to aid the quality checking of the project.

Meteorological Data

The ten surface meteorological stations in the region are divided into two groups. One group provides input data for the CALMET model. The other group is used to compare with CALMET predictions, in order to determine which whether the use of surface stations, using the various values of R1, results in more accurate prediction of winds across the area of interest. The relative locations of each of the stations in the two groups are plotted in Figure 1. The stations that were used to provide input to the CALMET Runs 1 through 4 and Run A1 are indicated with boxes.

Figure 1. Locations of the surface meteorological stations. Names of stations used in CALMET computations are enclosed in boxes. Numbers on the axes are UTM coordinates in kilometers.



The MM5 data set provided the upper air data used in Runs 0 through 4. Since Run A1 did not use the MM5 data, it requires the use of at least one upper air station. There are five upper air stations in the region of interest; however, four of them had data issues that would require some time to resolve. The remaining station, the Central Illinois National Weather Service Forecast Office (NWFO) in Lincoln, IL, is used to provide upper air data for Run A1. The Central Illinois NWFO is located about 40 kilometers northeast of Springfield, IL. This data set was downloaded from the WebMet website.⁹

Geographic Data

One-degree digital elevation model (DEM) data and land use/land cover (LULC) data from the U.S. Geological Survey were used in this project. They were downloaded from the WebGIS website¹⁰ from within the CALPUFF View interface.

Specific Details of the Six CALMET Runs

In general, CALMET parameters recommended in the IWAQM Phase II report⁵ are followed, except as noted below. Some parameters are not specified in the Phase II recommendations. These include the radius of influence of terrain features which we set to 10 kilometers (TERRAD=10) for all runs.

Run 0: MM5 data only; no surface or upper air observations

Since the IWAQM Phase II recommendations are based on runs using only observed data, some input parameters are necessarily different than the recommended. Significant settings of specific parameters, especially if they differ from the Phase II recommendations, are indicated in parentheses.

This run is made with only MM5 data, and no surface or upper air observations (NOOBS=2). The MM5 data is input as the initial guess field (IPROG=14). The 3D temperature field is derived from MM5 data (ITPROG=2). The gridded cloud option, ICLOUD, was inadvertently set to 3 (calculated from prognostic relative humidity) instead of the recommended 0. This is not expected to significantly affect the results of the wind analysis.

Run 1: MM5 data plus five surface stations, R1 = 10 kilometers

This run differs from Run 0 in that surface observing stations are introduced in addition to the MM5 data (NOOBS=1). ICLOUD was set to 0, as appropriate. The 3D temperature field is derived from surface observations (since they are available for this run), but uses MM5 data for the upper air (ITPROG=1).

Run 2: MM5 data plus five surface stations, R1 = 3 kilometers

This run differs from Run 1 only in the setting of R1, which is set to 3 kilometers.

Run 3: MM5 data plus five surface stations, R1 = 30 kilometers

This run differs from Run 1 only in the setting of R1, which is 30 kilometers.

Run 4: MM5 data plus five surface stations, R1 = 100 kilometers

This run differs from Run 1 only in the setting of R1, which is 100 kilometers.

Run A1: No MM5 data; five surface stations and one upper air station, R1 = R2 = 100 kilometers

In this run, the MM5 data was eliminated and CALMET was run entirely with data from five surface and one upper air observing stations (NOOBS=0, NUSTA=1, ITPROG=0). The absence of MM5 data also requires that IPROG be set to 0. An additional parameter, R2, is required to define the weighting for the upper air observations. 204 precipitation stations were added, since, in the other runs, precipitation was calculated from MM5 data, which is not used in this run.

The maximum radius for surface data (RMAX1) was expanded from 300 kilometers to 350 kilometers to be certain that at least two surface stations would be used over most of the area. The maximum radius for the upper air station (RMAX2) was set to 700 kilometers to be sure it would be accounted for over the entire domain.

ANALYSIS AND DISCUSSION

In this analysis we are considering the ability of CALMET to predict the wind direction and wind speed at five meteorological observing stations that are not used as input to CALMET. Specifically, we will consider the mean error in predicting wind direction, the standard deviation of the prediction error of the wind direction, and similar comparisons with wind speed. Measured calm winds were removed from the data before making the analysis, since wind direction for calm cases is undefined.

Testing the Mean Error in Predicting Wind Direction

Table 1 shows the mean error in the predicted wind direction produced by CALMET. At four of the five stations, the error is quite small, less than five degrees. However, station 13994 shows a consistent prediction of the wind direction about 15 degrees to the right of the observed. All of stations show a very slight improvement in accuracy as the effects of the measured data in the model is reduced, with the most accurate results appearing in Run 0, which is based entirely on MM5 data. However, is the difference enough to be statistically significant?

Table 1. Mean error of CALMET wind direction relative to the observed wind direction.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	0.43	0.65	0.48	2.77	2.12	3.19
13994	14.36	14.37	14.36	16.13	15.77	16.55
13995	-1.77	-1.80	-1.78	-2.10	-4.14	-9.59
14842	-0.81	-0.88	-0.80	-1.32	-2.53	-2.00
14933	3.40	3.53	3.41	4.55	1.71	5.26

Table 2 shows the Student's t test applied to this data. It should be noted that, since the frequency distribution of wind direction differences is truncated at ± 180 degrees, that the t-test may not be strictly appropriate. However, the authors believe the standard deviations are small enough that this truncation effect should not cause a significant problem. As will be seen (Table 3), the standard deviations range from about 25 to 55 degrees. At worst, this effect would likely result in calculated standard deviations that are slightly too small for the test, so that a case may test marginally significant when it should not be. To avoid any problems, we will use a stricter 0.01 confidence level rather than the 0.05 level. As seen in Table 2, none of the errors tests significantly different from Run A1 values.

Table 2. Student's t test for significant difference relative to RUN-A1. (See text for caution) At the .01 confidence level, $t = 1.96$.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	0.73	0.67	0.72	0.11	0.25	0.00
13994	0.48	0.48	0.48	0.09	0.15	0.00
13995	-1.71	-1.71	-1.71	-1.64	-1.21	0.00
14842	-0.38	-0.36	-0.38	-0.22	0.17	0.00
14933	0.41	0.38	0.41	0.16	0.78	0.00

Standard Deviation of the Error in Predicting the Wind Direction

The standard deviation of the errors in predicting the wind direction is a good indicator of correlation between the prediction and observed values. If the correlation is perfect, the standard deviation will be zero. Table 3 shows the standard deviations of the errors of CALMET-predicted wind directions relative to the observed wind direction for each run, at each meteorological station. For four of the five stations, the standard deviation of the error increases as the influence of measured data is given more preference over the MM5 data.

Table 3. Standard deviation of the error of CALMET wind direction relative to the observed wind direction.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	28.43	28.63	28.50	31.20	40.23	46.63
13994	36.39	36.62	36.42	37.92	47.80	55.22
13995	36.53	36.47	36.53	36.13	35.06	53.89
14842	26.14	25.94	26.10	24.93	27.25	34.23
14933	46.37	46.14	46.36	45.74	45.88	44.06

Statistical significance of these differences is tested using the common F-test, with the results shown in Table 4. For the degrees of freedom in the vicinity of 200 (they range from 189 to 211, depending on the meteorological station), the F value for the 0.05 significance level is 1.26, and for the 0.01 significance level is 1.39. All of the runs for four of the five stations showed standard deviations of the errors that are significantly different than (and better than) Run A1. At these four stations, adding the MM5 data and reducing the weight of the observed data resulted in improved performance.

Table 4. F-test of significance comparing standard deviation of errors relative to Run A1. **Bold** values are significant at the 0.05 level, while the underlined values are significant at the 0.01 level.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	<u>2.69</u>	<u>2.65</u>	<u>2.68</u>	<u>2.23</u>	1.34	1.00
13994	<u>2.30</u>	<u>2.27</u>	<u>2.30</u>	<u>2.12</u>	1.33	1.00
13995	<u>2.18</u>	<u>2.18</u>	<u>2.18</u>	<u>2.22</u>	<u>2.36</u>	1.00
14842	<u>1.72</u>	<u>1.74</u>	<u>1.72</u>	<u>1.89</u>	<u>1.58</u>	1.00
14933	1.11	1.10	1.11	1.08	1.08	1.00

A Look at Wind Speed Errors

Table 5 shows the mean wind speed errors in the CALMET predictions relative to the observed data. This shows no trend in accuracy with changes in R1.

Table 5. Mean wind speed errors.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	-0.57	-0.57	-0.57	-0.58	-0.44	-0.25
13994	-0.62	-0.61	-0.62	-0.55	-0.05	0.62
13995	-0.46	-0.46	-0.46	-0.45	-0.38	0.29
14842	-0.05	-0.04	-0.05	0.03	0.39	0.70
14933	-0.22	-0.22	-0.22	-0.22	-0.17	0.13

Table 6 shows the standard deviations of the wind speed errors for the various stations for each case. There is also no clear trend apparent in this tabulation.

Table 6. Standard deviations of the wind speed errors of CALMET versus observed.

STATION	RUN-0	RUN-2	RUN-1	RUN-3	RUN-4	RUN-A1
	MM5 ONLY	R1=3KM	R1=10KM	R1=30KM	R1=100KM	NO MM5
03947	1.23	1.22	1.23	1.23	1.38	1.63
13994	0.87	0.87	0.87	0.87	1.07	1.54
13995	1.58	1.58	1.58	1.59	1.72	2.38
14842	1.25	1.25	1.25	1.22	1.25	1.34
14933	1.57	1.57	1.57	1.57	1.59	1.70

CONCLUSIONS

In this case study, only the standard deviations of the error of CALMET's prediction of wind directions showed statistical significance. The best performance was shown when the observed data sites were given the least weight. It would not be appropriate to generalize this to all situations. It is clear, however, that giving weight to high quality measured data may still result in degradation in overall model performance.

REFERENCES

1. Scire, J.S.; Fernau, M.E.; Yamartino, R.J. *A User's Guide for the CALPUFF Dispersion Model (Version 5)*; Earth Tech, Inc.: Concord, MA. 2000.
2. Scire, J.S.; Robe, F.R.; Fernau, M.E.; Yamartino, R.J. *CALMET Meteorological Model (Version 5)*; Earth Tech, Inc.: Concord, MA. 2000.
3. *Interagency Workgroup on Air Quality Modeling (IWAQM) Phase 2 Summary Report and Recommendations for Modeling Long-Range Transport Impacts*; U.S. Environmental Protection Agency. U.S. Government Printing Office: Washington, D.C. 1998; EPA-454/R-94-019.
4. Anthes, R.A.; Hsie, E.Y.; Kuo, Y.H. *Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4)*; National Center for Atmospheric Research, Boulder, CO; 1995. NCAR Tech. Note, NCAR/TN-282+STR.
5. Grell, G.; Dudhia, J.; Stauffer, D. *A description of the fifth generation Penn State/NCAR Mesoscale Model (MM5)*; National Center for Atmospheric Research, Boulder, CO; 1995. NCAR Tech. Note, NCAR/TN-398+STR.
6. Olerud, D.; Alapaty, K; Wheeler, N. *Meteorological Modeling of 1996 for the United States with MM5*; MCNC—Environmental Programs, Research Triangle Park, NC; 2000.
7. The Hourly U.S. Weather Observations (HUSWO) CDROM is available from the National Climatic Data Center, Asheville, NC.
8. Calpuff View® is a product of Lakes Environmental Software, Waterloo, ON, Canada.
9. WebMet website is at <http://www.webmet.com>
10. WebGIS website is at <http://www.webgis.com>

KEY WORDS

CALMET, CALPUFF, air dispersion model, air quality model