Remote sensing data assimilation in WRF-UCM mesoscale model: Madrid case study

R. San José[1], J. L. Pérez[1], J. L. Morant[1] & R. M. González[2]
[1]Environmental Software and Modelling Group,
Computer Science School, Technical University of Madrid (UPM), Spain
[2]Department of Meteorology and Geophysics, Faculty of Physics,
Complutense University of Madrid, Ciudad Universitaria, Spain

Abstract

Data assimilation is a powerful numerical technique that is used to substantially improve numerical meteorological simulations. In this contribution we have used the WRF mesoscale meteorological model (NCAR, US) to show the importance of using remote sensing data (satellite and tower data), the sensitivity of the results and the improvement when compared with observational surface data (wind and temperature). We have used CLC100 m instead of GTOPO 30'', 10 m spatial resolution GIS data of Madrid (Spain) city to produce urban land use types according to the Urban Canopy Model (UCM) (NCAR) approach: airborne temperature (4 m spatial resolution), albedo, anthropogenic heat flux, shadowing in UCM and tower data (wind and temperature). The results show a high sensitivity to all of these parameters. For historical simulations - where in-situ meteorological data is available - data assimilation is a crucial tool to improve the results. The sensitivity of the results to the different high resolution input data is also crucial for the results of the simulation. The correlation coefficient for temperature is improved up to 0.960.
Keywords: remote sensing, data assimilation, mesoscale models.

1 Introduction

Data assimilation methods have been used in meteorological simulations for the last 50 years. The first methods included the so-called Cressman algorithm or

"objective analysis", which uses simple interpolation techniques and they are also known as three dimensional data assimilation approaches. The methods called "nudging" or four dimensional data assimilation (FDDA) are based on simple Newtonian relaxation (the second axiom of Newton). The nudging approach is applied to the dynamical part of the motion equation instead of the static part where the Cressman approach is applied. The basic idea is to add a term to the dynamical part of the equations, which is proportional to the difference between the model forecast and the observational matrix. A mixed data assimilation technique is known as the Kalman filter approach, which is an optical interpolation but applied in a continuous way to take into account the different vector states. Due to the huge dimension of the subsequent matrix to be solved, special KF algorithms have been developed, such as ensemble and reduced-rank KF tools. In fact, FDDA methods converge to KF methods [1, 2].

Mesoscale meteorological models have seen a rapid and extraordinary development during the last 50 years. In the 1980s the first numerical meteorological models appeared in parallel with the extraordinary progress on computational power [3]. The development of the MM5 model [4] during the 1990s and the change to the modern WRF (Weather Research and Forecasting model) model [5] produced a substantial advance on the simulation of the atmospheric dynamics. Numerical techniques have been improved substantially and the advance of computer power has produced a merging of the data assimilation techniques and numerical mesoscale models.

In this contribution, we show the sensitivity results of including more and better data into the numerical mesoscale meteorological model (WRF), which includes a recently added module called the Urban Canopy Model (UCM) that substantially improves the "urbanization" of the mesoscale meteorological model WRF. The UCM includes detailed information of three different types of urban land use classification: low intensity, high intensity and commercial urban land use types. The impact of adding high resolution GIS information (to produce the three different urban land use types already mentioned), detailed urban parameters to describe the Canyon street, anthropogenic heat flux and shadowing techniques is shown to be very important. Finally, the use of FDDA to assimilate tower detailed information is essential to improve the final results.

2   Data and model domains

We have used albedo from the AVHRR-NOAA satellite with 1100 m spatial resolution. Airborne data include: surface reflectance (SR), air temperature (2 m in height) (4 m spatial resolution), June, 28, 2008 21:11 GMT and July, 2008. Also included is high resolution land use data with 14 different land use types: 1. Water (lake), 2. Water (swimming pools), 3. Trees, 4. Green grass, 5. Bright bare soil, 6. Dark bare soil, 7. Roads with asphalt, 8. Other roads and pavements, 9. Asphalt roofs, 10. Tile roofs, 11. Concrete roofs, 12. Metal roofs, 13. Shadows and 14. Border (no data). We have used 10 m GIS data for the whole city of Madrid and CLC 100 m for land use in the mesoscale model combined with the

high resolution land use data mentioned above. The tower data is temperature, humidity and wind. We have also used high resolution emissivity data (4 m).

The WRF/UCM model has been set up by using a Lambert Conformal Conical Projection with three different nests: 1) 119 x 119 grid cells with 16.2 km spatial resolution; 2) 117 x 117 grid cells with 1.8 km spatial resolution and 3) 127 x 127 grid cells with 0.2 km spatial resolution. In all cases they are used with 23 vertical layers up to 100 mb (see Figure 1).

3   WRF/UCM model set-up

The WRF/UCM model requires a set of additional parameters to the control the WRF configuration. The UCM module includes three different urban land use types that have to be calculated as mentioned before. The UCM model includes a complete change of the urban energy fluxes, which produces different values for sensible and heat fluxes and ground and net radiation fluxes (see Figure 2). The inclusion of anthropogenic heat flux is done by downscaling the 0.5ºx 0.5º degree global information produced by Flanner [6] to 200 m spatial resolution in the highest spatial resolution domain over Madrid city (Spain). The shadow effects and the average road width and the height of the buildings are estimated by using the detailed land use information presented before. The details of the

1) 119 x 119 grid cells with 16.2 km spatial resolution;
2) 2) 117 x 117 grid cells with 1.8 km spatial resolution and
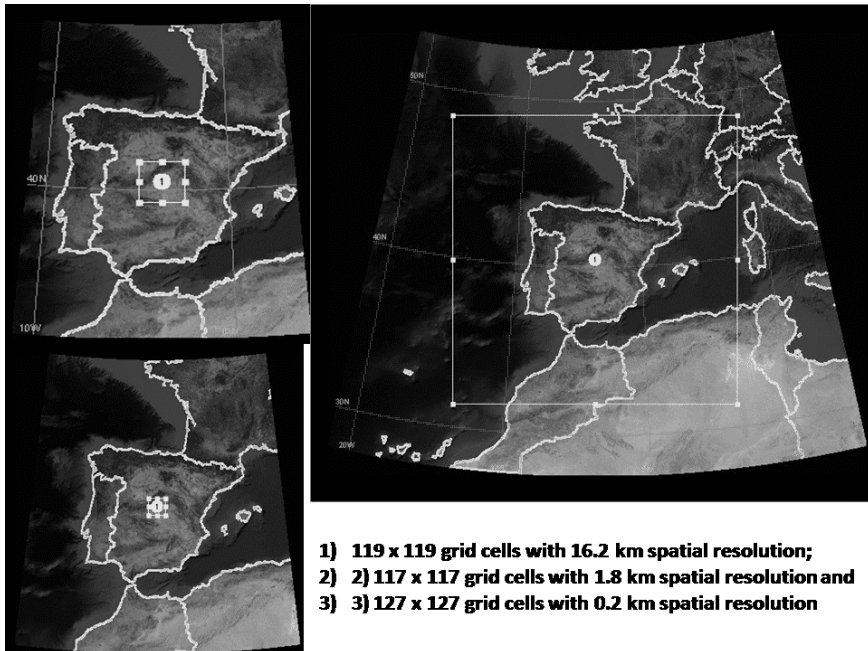3) 3) 127 x 127 grid cells with 0.2 km spatial resolution

Figure 1:   WRF/UCM model domains for this experiment. June, 23, 2008 to July, 7, 2008.
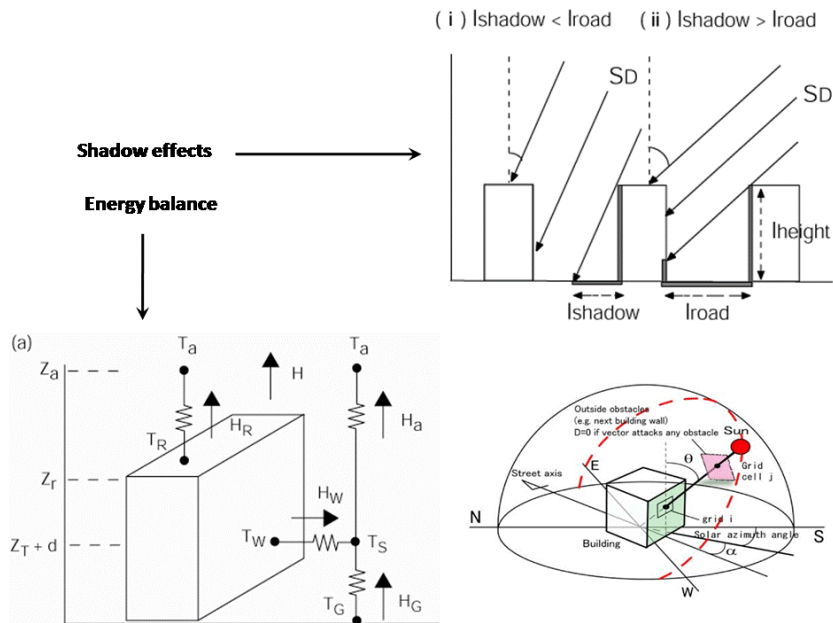
Figure 2:   Scheme of the energy balance and shadow effects as included in the UCM model.

Figure 3:   10 m spatial resolution GIS data set.

shadow model are included in the UCM description [7]. Figure 2 shows a scheme of the energy balance and shadow effects in an ideal canyon street as coded into the UCM. In Figure 3 we see a 3D approach to the shadow effects. Figure 3 also shows the 10 m spatial resolution GIS map of Madrid City.

4   Model results

We have performed several simulations as follows: 1) WRF. This simulation does not include the UCM module; 2) WRF-UCM-USGS24 module, but using only 24 USGS land use types; 3) WRF-UCM-USGS33. This simulation includes three different urban landuse types according to the UCM approach, it also includes high resolution (10 m) topography; 4) WRF-UCM-USGS33-PARAM. This simulation uses the height and width of the roads for the different urban landuse types according to the real 3D morphology (Geographic Information System); 5) WRF-UCM-USGS33-PARAM-AH. This simulation adds the anthropogenic heat flux mentioned before obtained by downscaling the global anthropogenic heat flux data set; 6) This simulation adds the shadow effect described into the UCM module; 7) The last simulation uses the tower data in the Madrid Municipality to improve the simulation data for wind and temperature. The stations used for applying the FDDA data assimilation technique are different to those used for final comparison between observed and modelled data. The simulations include the Iberian Peninsula domain with 16.2 km spatial resolution; the Madrid Community domain with 1.8 km spatial resolution and the Madrid city (Municipality) with 0.2 km spatial resolution. In Table 1 we see a summary of the parameters used for the three different urban landuse types based on the real GIS 3D data set. Figure 4 shows large differences for the ground and sensible heat fluxes when using UCM (WRF-WRF/UCM) (default values) in the Madrid domain. Up to 64 and 42 watt/m2, respectively, are due to the implementation of UCM on June, 26, 2008, 23:00 GMT. Figure 5 also shows large differences between WRF-UCM-USGS33-PARAM and WRF-UCM-USGS33. The parameters shown in Tables 1 and 2 produce differences up to 39 and 26 watt/m2 in ground and sensible heat fluxes.

Table 1:   Real parameters for building heights, roads and roofs in Madrid City obtained from the real GIS instead of using the default values as described in the UCM module for the three different urban landuse types.

	Urban fraction Default	Urban fraction Madrid	Buildings heights Default	Buildings heights Madrid	Roof width Default	Roof width Madrid	Road width Default	Road width Madrid
Low Intensity	0.5	0.97	5.0 m.	8.39 m.	8.3 m.	18 m.	8.3 m.	19 m.
High intensity	0.9	0.99	7.5 m.	14.96 m.	9.4 m.	20 m.	9.4 m.	19 m.
Commercial	0.95	0.94	10 m.	3.73 m.	10 m.	18 m.	10 m.	20 m.

Table 2:   Real parameters for emissivity, albedo and anthropogenic heat flux for the different urban landuse types as described in the UCM module.

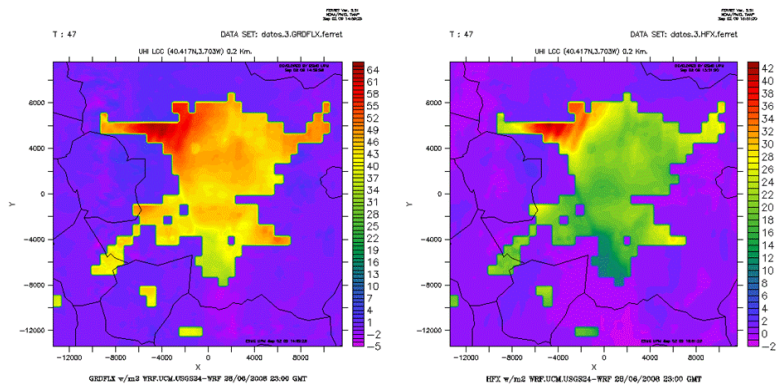|	Urban fraction Madrid	Buildings heights Madrid	Roof width Madrid	Road width Madrid	Frac. Albedo Madrid	Frac. Emissivity Madrid	Anthropogenic
Low Intensity	0.97	8.39 m.	18 m.	19 m.	0.0929	0.96906	12 w/m2
High intensity	0.99	14.96 m.	20 m.	19 m.	0.095435	0.97171	37 w/m2
Commercial	0.94	3.73 m.	18 m.	20 m.	0.095393	0.9677	13 w/m2

Figure 4:   Ground and sensible heat flux differences between WRF and WRF-UCM (default values). June, 26, 2008, 23:00 GMT.
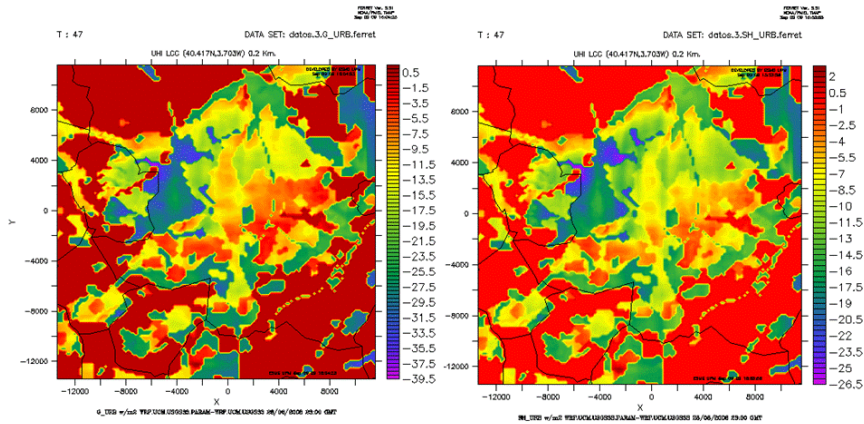

Figure 5:   Ground and sensible heat flux differences between WRF-UCM-USGS33-PARAM – WRF-UCM-USGS33. June, 26, 2008, 23:00 GMT.
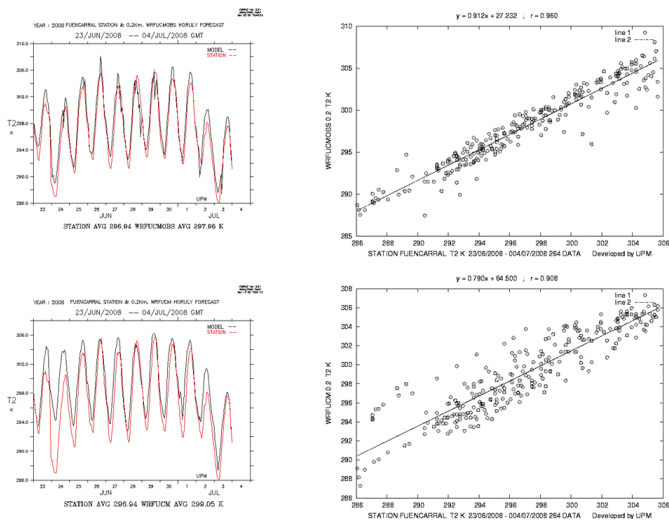
Figure 6: Model and observational data with FDDA (upper figure) and without FDDA (lower figure) in the Fuencarral monitoring station in the Madrid city area with 0.2 km spatial resolution for WRF/UCM and all the additional parameters presented in this work.
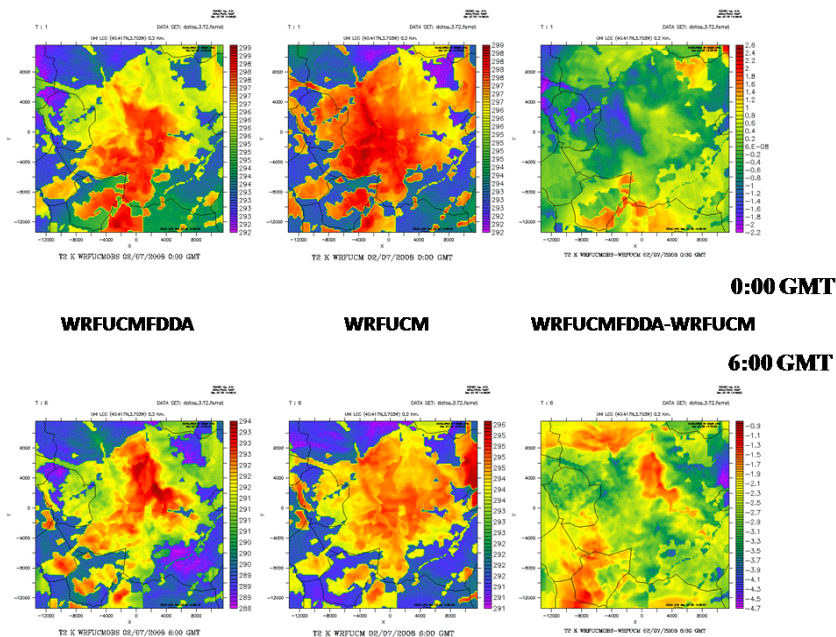

Figure 7: Comparison between temperature ºK maps with the results of WRF/UCM with FDDA and the results without FDDA data assimilation techniques.

Table 3:   Statistical parameters obtained when running the WRF/UCM
             model (with all detailed data sets) with the FDDA technique.

WRFUCMOBS AVG STATION	U10	V10	T2	WSPD
Observed mean	0.13	-0.18	298.43	1.90
Calculated mean/Obs. mean	1.00	1.70	1.00	1.36
Observed STD	1.39	0.78	4.90	0.88
Calculated STD/Obs. STD	1.37	1.85	0.95	1.56
Bias	-0.00	-0.13	0.30	0.69
Absolute Bias	0.87	0.86	0.84	0.94
R2	0.55	0.38	0.94	0.31
RMSE	1.27	1.15	1.22	1.33
RMSE/Observed Mead	9.96	-6.41	0.00	0.70
Percentage within +/- 50	18.94	7.95	100.00	61.74
No of Data Sets	264	264	264	264

Table 4:   Statistical parameters obtained when running the WRF/UCM
             model (with all detailed data sets) without FDDA technique.

WRFUCM AVG STATION	U10	V10	T2	WSPD
Observed mean	0.13	-0.18	298.43	1.90
Calculated mean/Obs. mean	6.84	2.13	1.00	1.71
Observed STD	1.39	0.78	4.90	0.88
Calculated STD/Obs. STD	1.96	2.40	0.87	1.63
Bias	0.75	-0.20	1.31	1.35
Absolute Bias	1.76	1.26	1.56	1.47
R2	0.57	0.33	0.87	0.38
RMSE	2.06	1.58	2.22	1.76
RMSE/Observed Mead	16.07	-8.76	0.01	0.93
Percentage within +/- 50	6.06	7.58	100.00	35.23
No of Data Sets	264	264	264	264

In Figure 6 we see the effect of including some monitoring stations as observational data into the FDDA numerical technique in WRFUCM (complete). The correlation coefficient increases from 0.908 to 0.960. In Figure 7 we see the differences in T ºK for the Madrid Municipality with 0.2 km spatial resolution when using WRF/UCM with the FDDA data assimilation technique and without using it. The increase of temperature at 0:00 GMT can be up to 2.6 ºK and the decrease in T ºK can be up to -4.7 ºK in some specific areas at 06:00 GMT.

5 Conclusions

We have performed several simulations using the WRF model with an increasing degree of detail in the use of topographic, landuse and other parameters included in the mesoscale meteorological model. In addition, we have used the Urban Canopy Model (UCM) and the data needed to perform the simulations has been extracted from detailed observational remote sensing data sets. The use of real 3D urban morphology to extract the real heights, roof widths and road widths has been crucial. The sensitivity analysis between the seven different simulations with increasing complexity shows that sensible and ground heat fluxes can change substantially during the daily cycle. In some cases ground and sensible heat fluxes vary up to a substantial percentage of the total value (20-30% or more). The FDDA data assimilation technique, when using observational tower datasets, shows a substantial improvement in the simulations when compared with observational data that has not been included in the FDDA technique [8]. The results substantially smooth the patterns in a way that large differences between observational and model data are drastically reduced. In most cases, the correlation coefficient is also substantially increased.

When wind speed is assimilated the correlation coefficient gets a little lower, but the patterns look much nicer (smoother). It is necessary to carry out many more sensitivity and assimilation exercises, but it is clear from this experiment that the FDDA technique looks very promising to improve model simulation results.

Acknowledgements

References

[1] Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. Q. J. R. Meteorol. Soc., **112**, 1177-1194.

[2] Todling R. and S.E. Cohn (1994): Suboptimal schemes for atmospheric data assimilation based on the Kalman filter. Mon. Wea. Rev. 122, 2530 – 2557

[3] Pielke, RA. 1984. Mesoscale Meteorological Modeling. 1ª edition, Academic Press, New York

[4] Grell, G. A., S. Emeis, W. R. Stockwell, T. Schoenemeyer, R. Forkel, J. Michalakes, R. Knoche, and W. Seidl, 2000: Application of the multiscale, coupled MM5/chemistry model to the complex terrain of the VOTALP valley campaign. Atmospheric Environment, 34, 1435-1453.

[5] Skamarock W.C., J.B. Klemp, J. Dudhia, D. Gill, D. Barker, W. Wang, J.G. Powers. A description of the advanced research WRF Version 2. NCAR Technical Note NCAR/TN-468+STR (2005)

[6] Flanner, M. G. (2009), Integrating anthropogenic heat flux with global climate models, Geophys. Res. Lett., 36, L02801

[7] Kusaka H., H. Kondo, Y. Kikegawa and F. Kimura (2001). A simple-layer urban canopy model for atmospheric models: comparison with multi-layer and slab models. Boundary-Layer Meteorology, 101, 329-358.

[8] San José, R., J.L. Pérez, J.L. Morant and R.M. González (2008): "Elevated PM10 and PM2.5 concentrations in Europe: a model experiment with MM5-CMAQ and WRF/CHEM" WIT Transactions on Ecology and the Environment, Vol. 116, pp. 3-12. ISSN: 1743-3541 (on-line).