Asociación Argentina
de Mecánica Computacional

Mecánica Computacional Vol XXX, págs. 3529-3542 (artículo completo)
Oscar Möller, Javier W. Signorelli, Mario A. Storti (Eds.)
Rosario, Argentina, 1-4 Noviembre 2011

AN APPROACH FOR USING REMOTE SENSING PRODUCTS AND GROUND OBSERVATIONS IN THE EVALUATION OF A NUMERICAL WEATHER PREDICTION MODEL

Pablo G. Cremades, Salvador E. Puliafito, David G. Allende, and Rafael P. Fernandez

Grupo de Estudios Atmosféricos y Ambientales (GEAA), Facultad Regional Mendoza, Universidad Tecnológica Nacional, Rodríguez 273, 5500 Mendoza, Argentina, pablocremades@gmail.com, http://www.frm.utn.edu.ar/geaa/

Consejo Nacional de Investigaciones Científicas y Técnicas

Keywords: WRF model, MODIS, radiometer, radiosounding, water vapor.

Abstract. A common problem in Air Quality Modeling (AQM) is the lack of meteorological observations in the area of interest for the initialization of Numerical Weather Prediction (NWP) models or evaluation of their results. Even if a significant number of weather stations are available, they cannot provide information above the surface level. Satellite atmospheric products constitute a unique source of information as they can give measurements at every point of the planet, and vertical profiles of some meteorological variables. However, this information should be validated if operational use is intended in air quality studies. This work develops a methodology for spatial and temporal forecast verification using three different sources of information regularly available in Mendoza: MODIS atmospheric products, radiosounding data and radiometer measurements of water vapor. We present comparisons and statistical analysis between Moderate Resolution Imaging Spectroradiometer (MODIS) temperature and dew point profiles versus traditional upper air profiles obtained by soundings and total water vapor content obtained with radiometer. We developed statistical indicators of information quality, as a guide for operational use of satellite atmospheric products for NWP model validation.

1 INTRODUCTION

Air Quality Modeling (AQM) provides very useful information for policy makers, as it gives pollutant concentrations over a region and its evolution over time, and may also provide predictions for near future. However, models require detailed meteorological information input that may be obtained from meteorological stations in the regions of interest, or from Numerical Weather Prediction (NWP) models. The use of meteorological models as input to air quality models is the best approach because they provide description of meteorological variables for every grid point of the modeling domain and at every vertical level, in contrast to ground observations that are sparse in the domain and only provide information at ground level. Accordingly, GEAA has been working with Weather Research and Forecasting (WRF) model to produce high resolution meteorological input data for use in air quality studies (Peckham et al. 2010).

The problem that remains is the validation of NWP models in order to guarantee that results from air quality models are within expectable errors. The validation process requires comparing model results with measurements of meteorological variables in the modeling domain. A secondary problem is the availability of such observations. For example, in the case of Mendoza, Argentina, there are only two meteorological stations. An interesting alternative is the use of satellite atmospheric products. They offer global measurements at a spatial and temporal resolution that depends on the instrument and platform.

In this paper, we discuss the use of Moderate Resolution Imaging Spectroradiometer (MODIS) atmospheric products, radiosounding data and a radiometer for validating WRF results.

2 MODIS ATMOSPHERIC PRODUCTS VALIDATION OVER MENDOZA

MODIS, on board platforms Terra and Aqua, offers atmospheric profile product consisting on the following information: total-ozone burden, atmospheric stability, temperature and moisture profiles, and atmospheric water vapor. All of these parameters are produced day and night for Level 2 at 5km x 5km pixel resolution. Profiles are given for 20 standard pressure levels. There are two MODIS Atmosphere Profile data product files: MOD07_L2, containing data collected from the Terra platform; and MYD07_L2, containing data collected from the Aqua platform. (http://modis-atmos.gsfc.nasa.gov/MOD07_L2/index.html).

We did a statistical analysis of MODIS data over a period of 4 years, from 2007 to 2010, over Mendoza, also using information available from radiosondes, ground observations and a radiometer operated by GEAA, for water vapor determination.

2.1 Radiosondes and Radiometer datasets

In Mendoza, radiosondes are launched only during summer, when heavy hailstorms threaten agricultural production. Table 1 summarizes data from station SAME, 87418: 32º50'S 68º47'W used in this paper for MODIS evaluation.

Year	Data series
2007	January - March
2008	January - March / November - December
2009	January - December
2010	January – March

Table 1: Radiosondes monthly measurement campaigns from 2007 to 2010. Station SAME, 87418.

National Meteorological Service launches one radiosonde a day during those months (at Airport station only). The estimated launch time is 11:00 local time. The information obtained from this station is daily temperature and dew point profile.

On the other hand, water vapor content was obtained from a measurement campaign with a radiometer operated by GEAA from January to March, 2007 (Ortiz and Puliafito, 2008). This instrument operates at 92GHz and allows the estimation of water vapor total column in the troposphere.

2.2 MODIS data processing

Comparing radiosonde temperature and dew point profiles with MODIS profiles requires extracting data for the point at the same coordinates of the radiosonde launching point from MODIS data files. Often, data for that point may not be available, as a consequence of cloud cover or poor quality of the retrieval. To overcome this problem one could take an average profile, taking all the points around the radiosonde launching point, but this approach may lead to errors as a consequence of the lack of representativeness of that ensemble. We did a statistical analysis to find the behavior of standard deviation as a function of the radius of the averaging kernel.

Given the radiosonde R_d for day d, we extracted data from MODIS Atmospheric Profile L2, for that day, for the radiosonde launching point coordinates and all the points within a radius r. r is expressed in points, and each point is equivalent to 5km. We only considered MODIS profiles which had at least 2 valid measures in the lower 5 pressure levels, and at most 6 missing values for the whole profile. All the valid profiles within radius r where averaged at all levels. In case of r=0, there is only one profile. For data sets that met the previous requirement, points were interpolated with a cubic spline. Figure 1 shows a MODIS temperature profile for February 13, 2007, obtained for r=0.

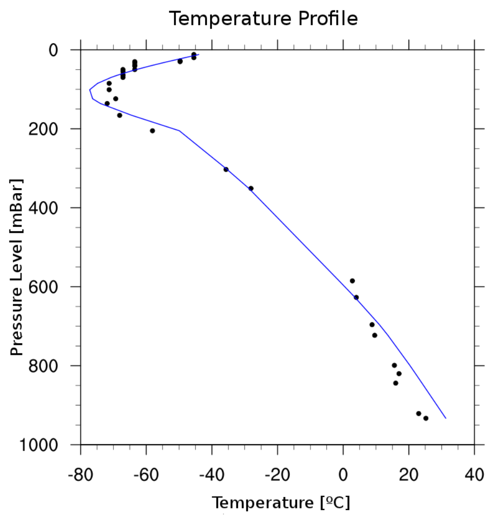

Figure 1: MODIS temperature profile for February 13, 2007 (blue line), and radiosonding data retrieved temperature (black dots).

Radiosonde and MODIS profiles were then compared at the same pressure levels where radiosondes measurements were defined, which vary from set to set. This approach guarantees that no errors are further introduced by interpolating radiosonding data points. We took the difference between MODIS values and radiosonde at every level and for all the days in the period. The differences were then grouped in ranges and the standard deviation was computed for each range (representative level), and for each year. Figure 2 depicts the procedure for

2007 dataset, and for an averaging kernel radius of zero.

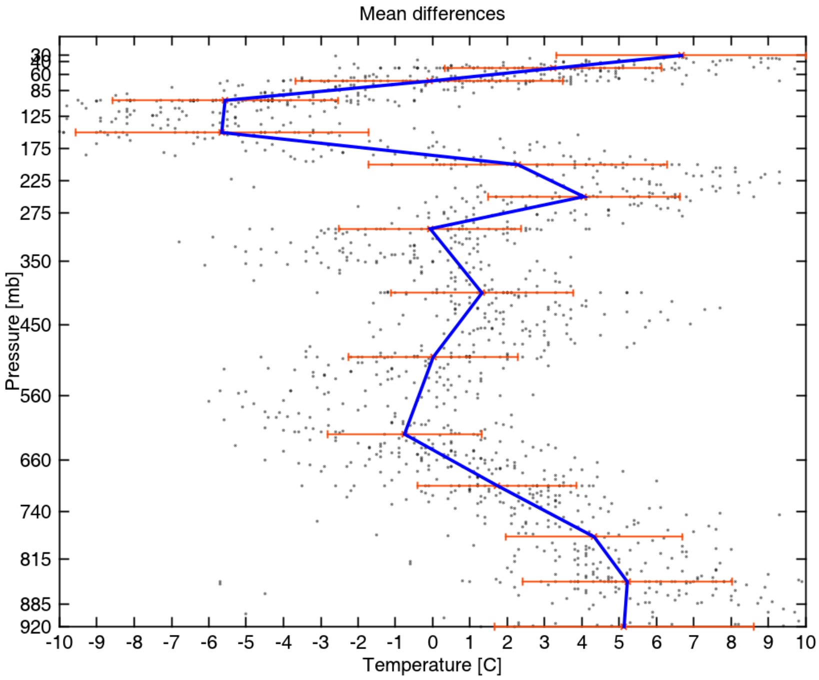

Figure 2: Mean differences (blue line) between MODIS and radiosonde 2007 profiles, and standard deviation (orange bars). Black dots represent absolute differences. Averaging kernel radius r=0.

The procedure was repeated for averaging radius from 0 to 5 points. We computed the number of days for which it was possible to obtain a MODIS profile, relative to the total number of days for which we had radiosondings (RP). We also computed the number of points used in building the average profile, relative to the maximum possible number of valid points, in case there are no missing values (RDP). With this information, we built an indicator that allowed us to find the optimum averaging radius, for which the probability of finding a MODIS profile for any given day is maximum, with minimum deviation from the true value. This indicator is the standard deviation (STD) of the differences between radiosondes and MODIS weighted with the inverse of the number of valid profiles, and the inverse of the relative number of not missing values:

$$WSTD_i(r)=\frac{1}{RP(r)}\times\frac{1}{RDP(r)}\times STD_i(r) \tag{1}$$

In order to summarize the information, pressure levels were grouped in three ranges indicated by the subscript i. Then, WSTDi(r) is the standard deviation of the differences between radiosondes and MODIS profiles within a distance r, weighted with the inverse of the number of valid profiles, and the inverse of the relative number of not missing values.

STD represents the long term index of agreement between MODIS and the local radiosonde. RP(r) accounts for the probability of finding a MODIS profile for any given day, averaging all the points within a radius of r from the actual radiosonde launching point. Finally, RDP is a measure of the statistical accuracy of the cubic interpolation and the averaging kernel.

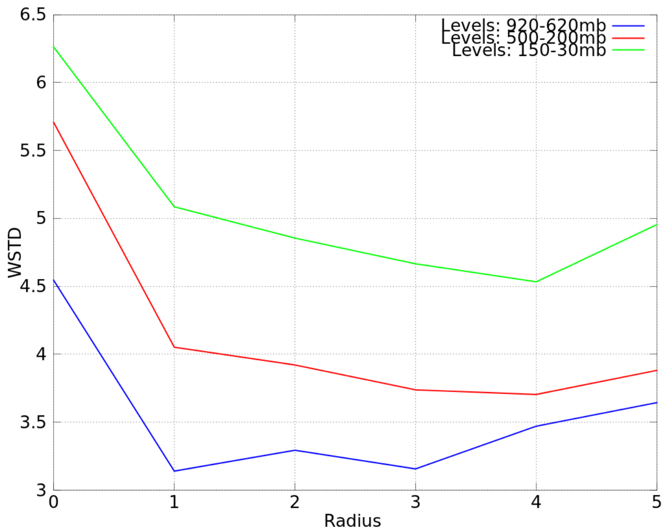

Figure 3: WSDT versus kernel averaging radius. Each line represents a pressure range.

Figure 3 shows the behavior of WSTD with respect to the averaging radius r. We grouped data in 3 pressure level ranges to analyze the tendency for lower, mid, and high pressure levels independently. All the functions show a minimum between r=3 and r=4, which turns to be the optimum radius for averaging. All the curves peak at r=0, mainly as a result of RP, which means that the probability of finding a MODIS profile at and exact location for any day is very low. This result enforces the idea of taking an average profile. RDP also grows as r does, making the cubic interpolation more precise. Above r=4, the RP and RDP approaches a constant value, and STD increases as a consequence of the lack of representativeness of points far from the radiosonde.

Figure 4 shows mean differences between MODIS and radiosonde temperature profiles for 2008 dataset, for and averaging kernel radius r=4. Standard deviation is slightly greater than that shown in Figure 2 for 2007, this phenomenon being a consequence of the error introduced by the averaging kernel. Nonetheless, for a total of 204 radiosondes in the period it was possible to obtain 132 MODIS profiles, while for r=0 only 79 MODIS profiles were valid. In other words, for r=4 we obtained 67% more profiles to compare than with r=0.

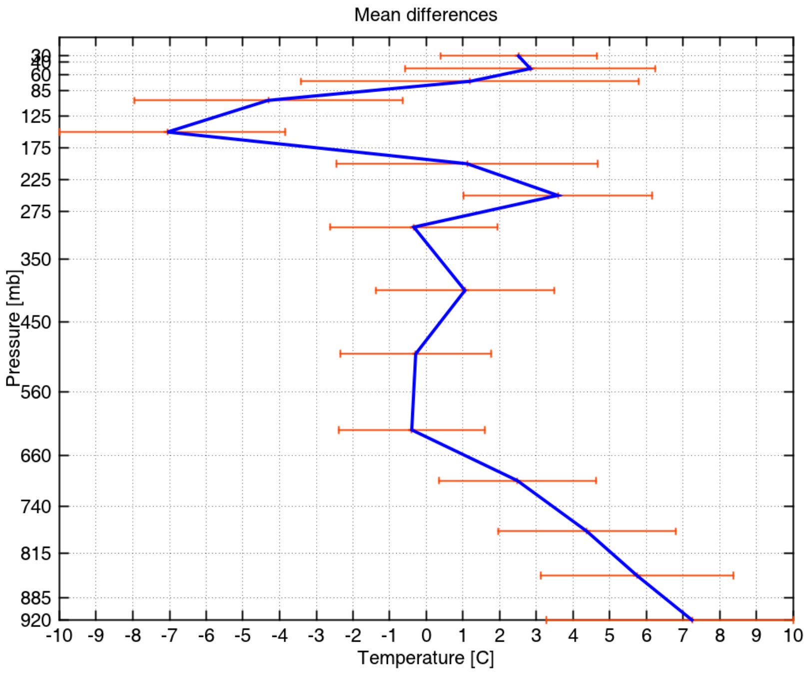

Figure 4: Mean differences (blue line) between MODIS and radiosonde 2008 profiles, and standard deviation (orange bars). Averaging kernel radius r=4.

The big differences found in lower troposphere can be a consequence of the time of satellite retrieval. While radiosondes are launched at 14:00 UTC, Terra passes over the region between 13:00 and 16:00 UTC.

In upper layers, the behavior seems to be the same for different periods, showing a systematic error, that requires further investigation.

Dew point showed much larger differences, in the order of -20ºC to -40ºC (not shown). This result may be a consequence of the poor vertical resolution of MODIS profiles, mainly in lower levels, where water vapor content is grater. Furthermore, lower level MODIS retrievals are often missing, increasing the error. On the other hand, for upper levels, differences in dew point are in the order of -5ºC.

2.3 Water vapor content

Ortiz and Puliafito (2008) showed the results obtained with a radiometer for water vapor content in Mendoza. Datasets from that measurement campaign (January to March, 2007) were used to validate MODIS water vapor retrievals.

The radiometer integrates the column of water vapor density along the instrument line of sight; a complete description of the radiometer can be found in Puliafito et al. (1995); results for 2007 measurement campaign can be found in Ortiz and Puliafito (2008). We simulated the line of sight of the instrument when extracting MODIS dew point values, in order to better represent the measurement process of the radiometer. Figure 5 shows the correlation between MODIS, radiometer and radiosonde.

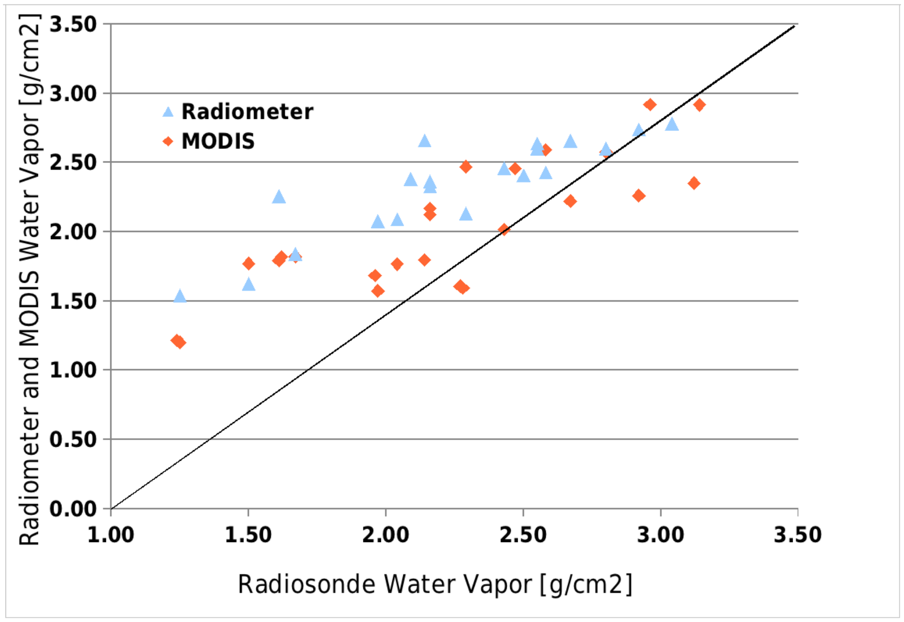

Figure 5: Correlation between Radiometer, MODIS and radiosondes, for 2007 radiometer measurement campaign in Mendoza.

MODIS shows in most cases smaller values than the radiometer and the radiosonde, probably due to the frequent lack of retrievals in the lower levels of the troposphere where water vapor is more abundant. Figure 6 shows a time series for January 2007 with MODIS and radiosonde values. The correlation is good, with maximum differences of 1g/cm^2. Mean differences are 0.2g/cm^2, in most cases being values reported by MODIS smaller than those reported by the radiosonde.

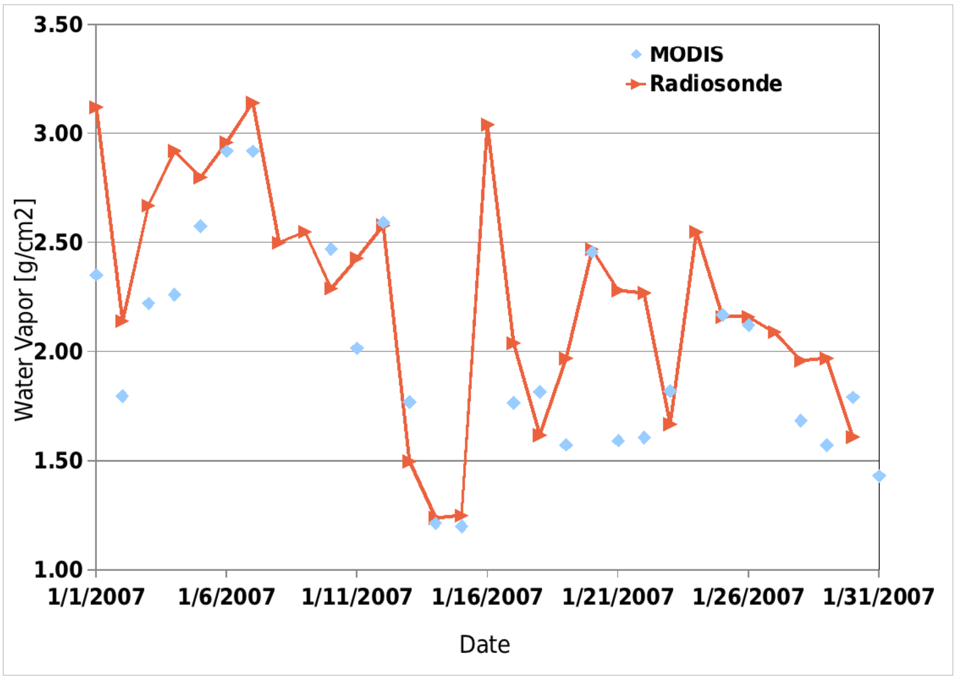

Figure 6: Time series for January 2007. Orange line represents water vapor total column reported by radiosonde. Light blue dots represent MODIS reported values.

3 WRF MODELING AND RESULTS VALIDATION

This section describes the local configuration of WRF used in this work to obtain high resolution meteorological fields. A complete description of the model setup and its characteristics can be found in Peckham et al. (2010).

3.1 WRF setup description

For WRF simulation we setup three nested domains with 27 vertical levels and a spatial resolution of 36km, 12km and 4km respectively, being the smaller domain centered in Mendoza city (32º 53' S, 68º 50' W, height: 750 m above sea level), as shown in Figure 7.

WRF requires static data related to the description of domains that includes: topography, land use and land cover. The package includes databases that WRF uses by default to include this information (Peckham et al., 2010). However, those sources have data with very coarse resolution, which result inadequate for most real cases simulation. Instead, we replaced that information with more accurate and fine resolution datasets. We used terrain elevation provided by Shuttle Radar Topography Mission SRTM3 (Rodríguez et al., 2005), with a resolution of 90m x 90m, more than 10 times finer than topographic data used by default.

Land use and land cover information was also replaced and enhanced with datasets generated by local institutions (Dirección de Ordenamiento Territorial de Mendoza (DOADU), Instituto Nacional de Tecnología Agropecuaria (INTA), Universidades Nacionales) (Cruzate et al., 2007; Dirección de Ordenamiento Ambiental y Desarrollo Urbano, 2009; Instituto de Desarrollo Rural, 2009)

Figure 7: Nested domains setup in WRF. Domain 3 centered in Mendoza.

The physical parameterizations we chose were the ones recommended for mid latitude locations. The microphysics scheme used is WRF Single Moment (WSM). We used no cumulus parameterization. We used the Noah Land Surface Model (LSM) with 4 surface levels for description of surface temperature and soil moisture, and sensible and latent heat transfer to the planetary boundary layer. The planetary boundary layer scheme used in this simulation is the Yonsei University (YSU PBL).

The model was setup to run from January to March, 2007, with outputs every 15 minutes.

3.2 WRF results evaluation

We compared WRF results with all the data sources discussed previously in this article. First we evaluate temperature profile using radiosonde information, following a procedure similar to the one described in section 2.2, in order to asses the long term behavior of WRF model. Then we compare water vapor content simulated with all measurements available in the period. Finally, we use the results of section 2 to show how MODIS atmospheric profiles can be used to validate WRF over the entire domain.

3.2.1 Long term analysis of WRF temperature profile.

Using the same statistical long term analysis we used to validate MODIS, we compare the temperature profiles obtained by WRF with those reported by the National Weather service for all days in 2007 period.

For every radiosonding available, we extracted WRF temperature profile at the radiosonde launching point location, and at 11:00 local time. Those profiles were linearly interpolated in the vertical direction in order to find temperature values at pressure levels defined in the corresponding radiosonde dataset. Then we computed mean differences between the profiles and standard deviation, for the same pressure ranges we used for MODIS.

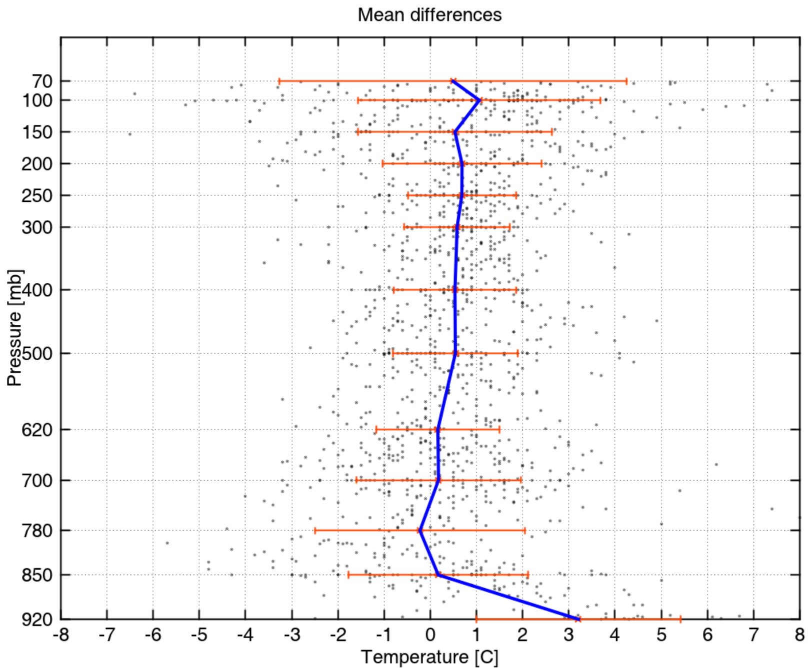

Figure 8: Mean differences and standard deviation between WRF and radiosonde temperature profiles, for 2007 summer days.

Figure 8 shows results of the statistical analysis. For levels above 850mbar mean differences are less than 0.9ºC with standard deviation between 1ºC and 2ºC, showing a very good degree of agreement. On the other hand, for lower pressure levels mean differences are around 3ºC with deviations of 2ºC. These discrepancies may be the result of a coarse representation of the lower atmosphere and/or a misrepresentation of land use or land cover, being these variables responsible for heat fluxes between soil and the atmosphere. Another possible source of error is the surface model chose for this simulation, as explained in Misenis and Zhang (2010).

3.2.2 WRF temperature profile validation with MODIS

Once MODIS atmospheric products had been validated, and the expected error determined, we compared WRF temperature profiles at every grid point of domain 3 with satellite retrievals, for February 20, 2007. Firstly, we remapped the MODIS image to match the WRF domain. Then we used an averaging kernel with a radius r=4 to process the image, accordingly to results obtained in section 2. Finally, we interpolated WRF profiles at every grid point in the vertical direction to match pressure levels in the MODIS profiles. We took differences and standard deviation at all levels. Results are depicted in Figure 9.

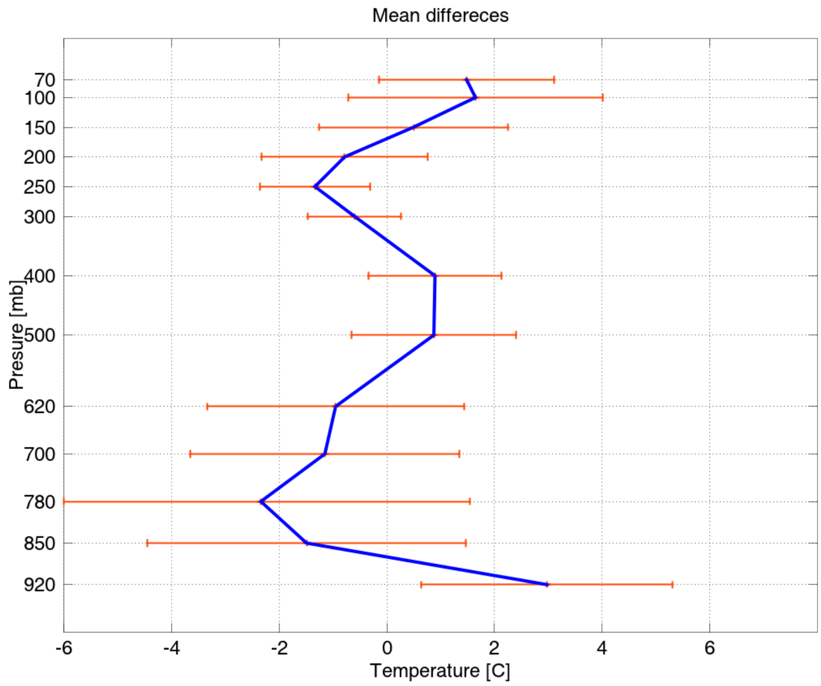

Figure 9: Mean differences between WRF and MODIS temperature profiles over the whole domain for February 20, 2007. Orange bars represent the standard deviation.

Although differences are big, in the order of 1,5ºC, they are under the values found in the long term analysis of MODIS, which means that WRF results may be within acceptable errors. Once again, differences are bigger in the lower levels. Unlike previous results, standard deviation values are in the order of 1ºC to 4ºC, maybe as a result of errors in the remapping process.

3.2.3 Tropospheric water vapor

In this section we analyze comparisons among WRF, radiometer and MODIS for water vapor total column. Figure 10 shows correlation for the three sources. Radiometer agrees with radiosonde within 0.1g/cm^2 for most days. High values reported by the radiometer may be due to presence of clouds in the area, which increase liquid water content in the atmosphere. On the other hand, MODIS seems to underestimate water vapor concentration. As previously explained, MODIS has low vertical resolution and difficulties to retrieve atmospheric variables in lower layers, were vapor is more abundant. Consequently, MODIS may not be an appropriate source of information to evaluate water vapor content.

Finally, WRF shows a good degree of agreement with radiosondings and radiometer measurement. Although WRF can simulate the presence of clouds, the location, characteristics and time of these cumuli may not be realistic. In such cases, WRF may differ greatly with the radiometer. In order to better capture these types of phenomenon, the model can be setup with special parameterizations, which were not used in this case, and are beyond the scope of this article (Santos et al., 2010).

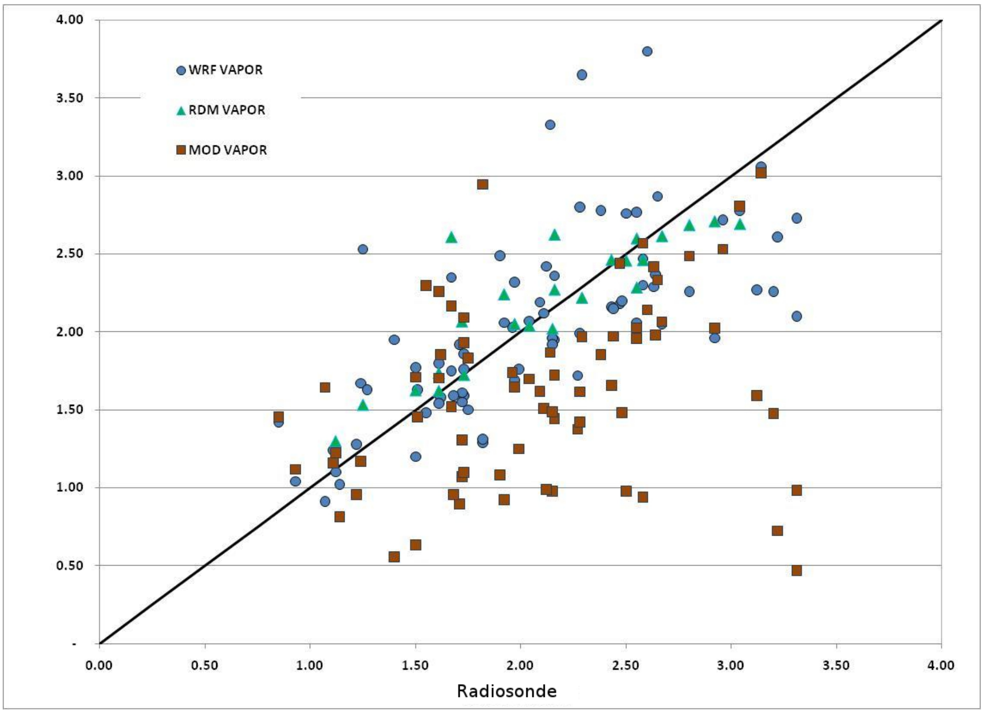

Figure 10: Correlation between WRF (blue circles), radiometer (green triangles), MODIS (brown squares) and radiosonde.

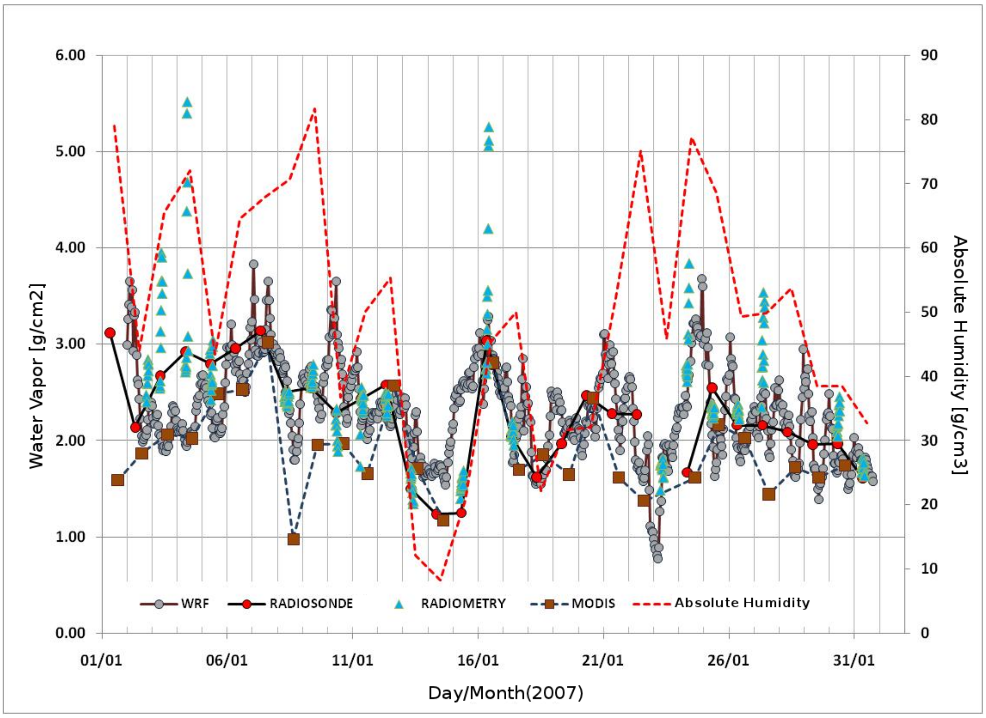

Figure 11: Water vapor time series for January, 2007. WRF (gray dots), radiosonde (red dots), radiometer (Light blue triangles), MODIS (brown squares) and absolute humidity (red dashed line).

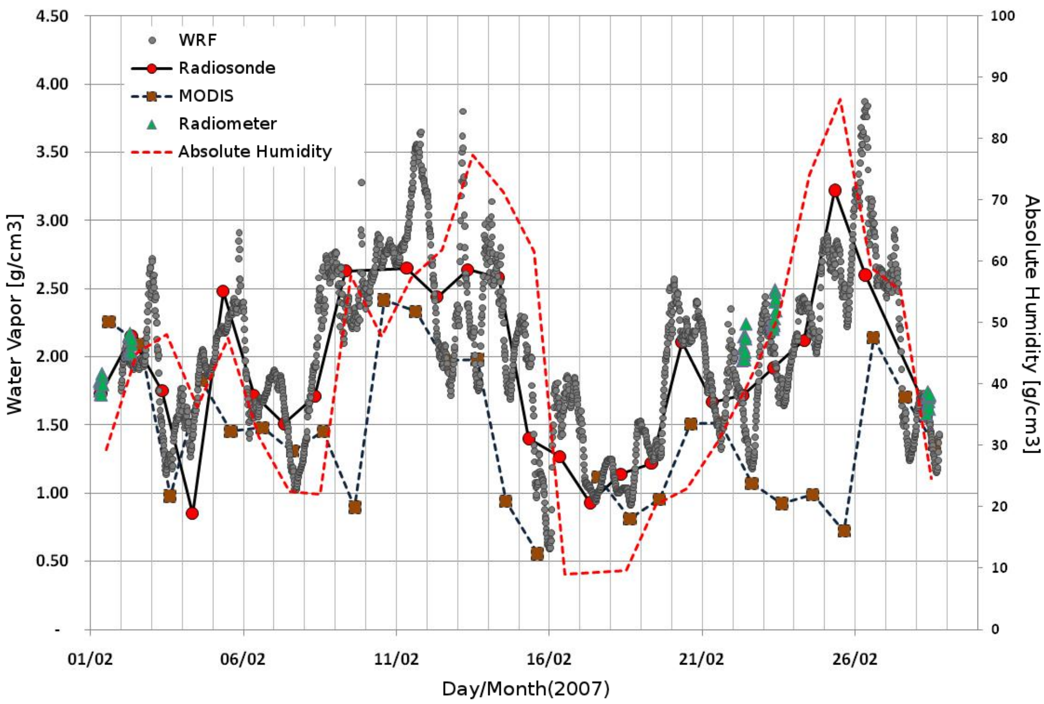

Figure 12: Water vapor time series for February, 2007. WRF (gray dots), radiosonde (red dots), radiometer (Light blue triangles), MODIS (brown squares) and absolute humidity (red dashed line).

Figures 11 and 12 show time series for January and February respectively. In general, WRF, radiosonde and radiometer agree with differences not grater than 0.5g/cm^2. Nonetheless, in January there are radiometer retrievals that in some cases are twice as high as the radiometer or WRF reported values. The reason for such large differences may well be the presence of clouds in the area where the radiometer was located. WRF could replicate the behavior of water vapor content, with a slight time retard by the end of February. In all cases, MODIS underestimates the water vapor total column, for reasons already discussed.

4 CONCLUSIONS

Weather forecasting models like WRF can be very useful tools in air quality studies, as they can provide a detailed description of meteorological variables at every grid point of the computational domain, with arbitrary spatial and temporal resolution. This approach may be mandatory when there are no other sources of meteorological data in the area of interest. However, model results should be verified prior to using them as input of an air quality model.

In this paper we discuss the use of radiosondes, radiometer and satellite retrievals for validating WRF results. First we study the behavior of MODIS profiles over a period of 4 years, over Mendoza city. We found that satellite information have to be preprocessed to avoid problems related with missing retrievals. We built a statistical index that allowed the definition of an averaging kernel for data processing. We found that with the proposed methodology, and a kernel radius r=4, the probability of finding a profile (temperature or dew point) for any point in a domain, and that meets the quality criteria described in section 2.2, increases from 40% to 72%. Mean differences remain within the same range, less than ±2ºC for mid altitude levels and between ±4ºC and ±6ºC. Such big discrepancies near the surface can be attributed to the differences in retrieval time.

For water vapor total column comparison, we used data from 2007 radiometer measurement campaign. In general, MODIS underestimate this variable, as a consequence of the frequent lack of retrievals in lower levels of the atmosphere. This result lead to the conclusion that using MODIS atmospheric products for determining water vapor content below 700mbar may not be appropriate.

With those results, we then evaluated WRF results for January, February and March, 2007, over Mendoza city. First, we follow the same procedure we used whit MODIS dataset to study the long term behavior of WRF, in terms of temperature profile. Differences with values reported by radiosondes were almost zero, with deviations that do not exceed 1.5ºC for pressure levels over 850mbar. Near surface differences reach 3ºC, probably due to bad soil representation in the model. Statistical analysis over the entire domain for one day in the period showed differences between MODIS and WRF that were within expected values according to results of section 2.

For tropospheric water vapor content comparison we used MODIS, radiometer and radiosonde. WRF showed good correlation with radiometer and radiosonde, except for particularly high values retrieved by the radiometer, which may be attributed to the presence of clouds in the area that WRF could not simulate. MODIS underestimates in most cases the water vapor content.

Addtitionaly, we have probed that the configuration of WRF seems very appropriate for the region under study, as variables studied were within expected errors.

Eventhough the validation presented in this paper was performed for MODIS data over the Province of Mendoza, the methodology presented can be extended to any other satellite product and ground-based measurement at any point of the world.

REFERENCES

Cruzate, G., Gomez, L., Pizarro, M. J., Mercury, P., Banchero, S., Mapas de suelos de la República Argentina. *p. Proyecto PNUD ARG/85/019,* 2007. http://geointa.inta.gov.ar/node/11/id=15.

D.O.A.D.U. Dirección de Ordenamiento Ambiental y Desarrollo Urbano, 2009. http://www.ambiente.mendoza.gov.ar/direcciones/doadu.php.

IDR. EcoAtlas. *Instituto de Desarrollo Rural*, 2009. http://www.ecoatlas.org.ar.

Misenis, C., Zhang, Y., An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options. *Atmospheric Research*, 97:315–334, 2010.

Ortiz, G., Puliafito, S., Determinación del contenido de vapor de agua troposférico mediante radiometría y radiosondeo. *Cuarto Encuentro de Investigadores y Docentes de Ingeniería*, 71–78, 2008.

Peckham, S., Grell, G., McKeen, S., Fast, J., Gustafson, W., Ghan, S., Zaveri, R., Easter, R.,Barnard, J., Chapman, E., Wiedeinmyer, C., Schmitz, R., Salzmann, M., Freitas, S.WRF/Chem Version 3.2 Users Guide. *National Center for Atmospheric Research (NCAR)*, *USA*, 2010.

Puliafito, E., Bevilacqua, R., Olivero, J., Degenhardt, W., Retrieval error comparison for several inversion techniques used in limb-scannig millimeter wave spectroscopy. *Journal of Geophysical Research*, 100:14,257–14,268, 1995.

Rodriguez, E., Morris, C. S., Belz, J. E., Chapin, E. C., Martin, J. M., Daffer, W., Anassessment of the SRTM topographic products. *Technical Report JPL D-31639*, 143, 2005.

Santos, R., Garino, C., Catania, C., Salinas, S., Martinez, P., Simulación numérica de eventos convectivos severos en la provincia de Mendoza. *Sexto Encuentro de Investigadores y Docentes de Ingeniería*, Energía y Medio Ambiente: 36.