USING THE WRF MODEL TO REFINE NCEP CFSV2 REANALYSIS ATMOSPHERIC PROFILE: A SOUTHERN BRAZIL TEST CASE

Lucas Ribeiro Diaz¹, Silvia Beatriz Alves Rolim¹, Daniel Caetano Santos², Pâmela Suélen Käfer¹, Nájila Souza da Rocha¹ and Rita de Cássia Marques Alves¹

ABSTRACT. The vertical distribution of atmospheric parameters is pivotal to understand atmospheric dynamic processes. This study analyzes the use of the Weather Research and Forecasting (WRF) model to generate atmospheric profiles as an alternative for a local radiosonde. We performed simulations in the WRF model for two dates with distinct weather conditions (01/09/2018 – cloudy and 03/14/2018 – clear-sky). The simulated data were compared with radiosondes launched in the study area, both (i) along the entire profile and (ii) with focus on the Planetary Boundary Layer (PBL) – for the variables water vapor mixing ratio ($q$), potential temperature ($\theta$) and wind speed ($\text{Speed}$). The results showed a high correlation between the simulated and observed profiles, with most R values higher than 0.9. Bias and RMSE, respectively, ranged between -1.29 – 0.66 g/kg and 0.48 – 2.01 g/kg for $q$; -0.52 – 0.25 K and 0.44 – 3.00 K for $\theta$; and -0.20 – 1.31 m/s and 1.61 – 2.77 m/s for $\text{Speed}$. The difference between the profiles from different WRF nested grids was minimal, suggesting that a resolution of approximately 12 km is enough to a good trade-off between detailing and computational cost. WRF model seems suitable to simulate fine atmospheric profiles.

Keywords: vertical profiles; numerical weather prediction; meteorology; remote sensing; planetary boundary layer.

RESUMO. A distribuição vertical dos parâmetros atmosféricos é fundamental para entender os processos dinâmicos da atmosfera. Este estudo analisa o uso do modelo Weather Research and Forecasting (WRF) na geração de perfis atmosféricos, como uma alternativa à radiosondagens locais. Foram realizadas simulações com o WRF para duas datas em condições de tempo distintas (09/01/2018 – alta nebulosidade e 14/03/2018 – céu claro). Os dados simulados foram comparados com radiosondagens lançadas na área de estudo, (i) ao longo de todo o perfil e (ii) com foco na camada limite planetária (PBL) – para as variáveis razão de mistura do vapor de água ($q$), temperatura potencial ($\theta$) e velocidade do vento ($\text{Speed}$). Os resultados mostraram uma alta correlação entre os perfis simulados e observados, com a maioria dos valores de R superiores a 0,9. Viés e RMSE, respectivamente, variaram entre -1,29 – 0,66 g/kg e 0,48 – 2,01 g/kg para $q$; -0,52 – 0,25 K e 0,44 – 3,00 K para $\theta$; e -0,20 – 1,31 m/s e 1,61 – 2,77 m/s para $\text{Speed}$. Mínimas diferenças entre os perfis das diferentes grades aninhadas sugerem que uma resolução horizontal de aproximadamente 12 km é um bom balanço entre detalhamento e custo computacional. O WRF mostrou potencial na simulação de perfis atmosféricos refinados.

Palavras-chave: perfis verticais; previsão numérica do tempo; meteorologia; sensoriamento remoto; camada limite planetária.

Corresponding author: Lucas Ribeiro Diaz

¹Universidade Federal do Rio Grande do Sul – UFRGS, State Research Center for Remote Sensing and Meteorology, Av. Bento Gonçalves, 9500, Campus do Vale, 91501-970, Porto Alegre, RS, Brazil – E-mails: lucas.diaz@ufrgs.br, silvia.rolim@ufrgs.br, pamelaskafer@gmail.com, najila.rocha@ufrgs.br, rita.cma@terra.com.br
²Universidade Federal de Santa Maria – UFSM, Center for Natural and Exact Sciences, Department of Physics, Av. Roraima, 97105-900, Santa Maria, RS, Brazil – E-mail: danielcae@gmail.com
INTRODUCTION

Modeling climate and weather on regional and global scales requires accurate monitoring of atmospheric parameters such as air temperature, wind speed, and water vapor content in the air (Sobrino et al., 2015). Water vapor plays a critical role in governing the terrestrial atmospheric processes (Sherwood et al., 2010; Jiang et al., 2019; Chang et al., 2020). It can be considered the most important greenhouse gas, influencing the planetary radiative budget and therefore the climate balance (Held & Soden, 2000; Filioglou et al., 2017; Chang et al., 2020; De Rosa et al., 2020). Atmospheric temperature also plays an important role in Earth’s climate and in deciding the thermodynamic state of the atmosphere (Thorne et al., 2005; De Rosa et al., 2020; Rao et al., 2020). Thus, the analysis of variations and trends in such atmospheric parameters has received growing interest in recent years due to global climate changes (IPCC, 2014; Sobrino et al., 2015; Jiang et al., 2019; De Rosa et al., 2020).

Notably, the continuous and accurate monitoring of the vertical structure of atmospheric parameters on global and regional scales is a key issue for assessing climate change (Sobrino et al., 2015; Chang et al., 2020; Rao et al., 2020). Besides all of the above weather and climate related issues, the accurate description of the vertical water vapor and temperature variability is also required to Numerical Weather Prediction (NWP) and to improve the mesoscale models for any meteorological research (Turner et al., 2000; Chang et al., 2020; Zhao et al., 2020).

Historically, monitoring the atmosphere has been characterized by sparse in situ measurements (Sobrino et al., 2015). The characterization of the vertical structure of the atmosphere, based on the launching of radiosondes, is no different. Radiosondes are meteorological balloons that can measure atmospheric physical parameters such as profiles of air temperature, pressure, and humidity (Rahimzadeh & Mobasheri, 2011; Hassanli & Rahimzadeh, 2019). A radiosonde profile can characterize atmospheric situations at specific latitudes and longitudes with a high precision (Yang et al., 2020). Nevertheless, it is only available at specific sites (e.g., airports) and times, usually twice a day (00:00 and 12:00 UTC) (Wang et al., 2002; Hassanli & Rahimzadeh, 2019). Moreover, radiosondes are quite expensive and have poor coverage over oceans and in the southern hemisphere (Li et al., 2003; Xu et al., 2015; Chang et al., 2020; De Rosa et al., 2020). Particularly, radiosonde observational networks in South America are geographically sparse and mostly located in or near urban settlements. Hence, it recorded data set is used mainly as “ground truth” in atmospheric research.

Reanalysis data is a growingly popular alternative to overcome most of these limitations (Kalnay et al., 1996; NCEP et al., 2000; Kanamitsu et al., 2002; Dee et al., 2011; Rienecker et al., 2011; Saha et al., 2010, 2014; Kobayashi et al., 2015; Hersbach et al., 2020). Using numerical atmospheric model outputs and data assimilation techniques and observations from multiple data sources (e.g., radiosondes, satellite data, aircraft and ground observations) for multiple variables (e.g., temperature, humidity, solar radiation, geopotential height, etc.) are combined to represent the state of atmosphere at different temporal and spatial scales (Mooney et al., 2011; Alghamdi, 2020; Yang et al., 2020). The resulting data set is gridded and referred to as retrospective analysis – reanalysis data. These data have global coverage.
and an extended homogeneous time series with no temporal and spatial gaps (Alghamdi, 2020). They are provided at low spatial resolution of several degrees every, generally, one, three or six hours (Wang & Zeng, 2012; Yang et al., 2020).

The increase in spatial resolution of atmospheric profile data makes them useful not only for global studies, but also for studies at regional scales. An important application of these vertical profiles is related to the atmospheric correction of thermal infrared (TIR) remote sensing data, aiming at the retrieve of land surface parameters, such as temperature (LST) and emissivity (LSE) (Barsi et al., 2003; Jiménez-Muñoz et al., 2010; Coll et al., 2012; Pérez-Planells et al., 2015; Rosas et al., 2017; Meng & Cheng, 2018; Duan et al., 2019; Yang et al., 2020). These are key parameters closely connected to the Earth’s surface energy balance (Tardy et al., 2016). They are of great significance to the studies of climate, hydrology, geology, ecology, agronomy, and natural disasters (Vicente & Souza Filho, 2010; Anderson et al., 2012; Huo et al., 2015; Gallego-Elvira et al., 2016; Sobrino et al., 2016; Cheng & Kustas, 2019; Nill et al., 2019; Roy et al., 2020).

Although meteorological reanalysis data sets are widely used in scientific studies, it accuracy is generally lower for regions with inadequate spatial coverage of permanent observatories, such as the oceans and many countries in the Southern Hemisphere (Tonooka, 2001; Chen et al., 2014; Chen & Liu, 2016; Alghamdi, 2020). Because they are spaced at grid points, they may also have the accuracy of meteorological phenomena on a sub-grid scale affected. The same may occur for phenomena of variable time scale, since the time intervals are typically 6 h (Tonooka, 2001). Previous studies have showed that reanalysis data are not consistently reliable or equally effective due to differences in assimilation schemes and the observational data used (Chen et al., 2014; Bao & Zhang, 2019; Alghamdi, 2020).

Reanalysis data efficiency is local, time and weather dependent, and small differences between reanalysis and observations may lead to substantially different results (Mooney et al., 2011; Alghamdi, 2020). Studies evaluating the performance of reanalysis data do not have a wide application due to diverse regional climate conditions (Bao & Zhang, 2013, 2019; Alghamdi, 2020). Dong et al. (2017) and Bao & Zhang (2019) intercompared and evaluated reanalysis products for horizontal wind, temperature, and water vapor mixing ratio (and relative humidity) over the Tibetan Plateau. Graham et al. (2019) evaluated atmospheric reanalysis profiles over the Arctic, and Jones et al. (2016) in Antarctica. Chen et al. (2014) assessed four reanalysis products representing the warm-season diurnal cycle over East Asia and reported that the products performed differently at the sub-regional scale than at a large scale. Differences in local atmospheric processes and topography are among the most significant factors affecting the efficiency of reanalysis (Schafer et al., 2003; Alghamdi, 2020).

New weather prediction models enjoy better computing performance and parameterization of physical processes to improve the reanalysis forecast accuracy (Evans et al., 2012; Hassanli & Rahimzadegan, 2019; Prasad et al., 2020). As mentioned above, coarse resolution of global-scale data may not be suitable for local use. Thus, mesoscale atmospheric models are used in local areas with global models data as the boundary and initial conditions (Wee et al., 2012; Hassanli & Rahimzadegan, 2019). The Weather Research and Forecasting (WRF) model (Skamarock et al., 2008)
is among the most frequently used mesoscale models (Knievel et al., 2007; Powers et al., 2017; Hassanli & Rahimzadegan, 2019). As its name indicates, the WRF is designed for both research and NWP. Community-based, free and open source, the WRF model provides specialized resources for a variety of applications in terrestrial systems (Skamarock et al., 2008; Powers et al., 2017).

Prasad et al. (2020) simulated multiple cases (including vapor and temperature profiles) over northeastern Australia using the WRF model, initialized by and nudged to reanalysis data and tested with various schemes. Resulting in overall skillfully WRF simulations. Lin et al. (2018) analyzed the impact of WRF model grid resolution on water vapor simulations. Whereas Ruiz et al. (2010) and Santos & Nascimento (2016) assessed WRF parameterization schemes, against vertical profiles and surface variables, over South America.

In this context, this paper aims to analyze the feasibility of using the mesoscale WRF model (Skamarock et al., 2008), to generate vertical atmospheric profiles, increasing the spatial and temporal resolutions of the global reanalysis, thus being an alternative to the need for a radiosonde. We performed atmospheric simulations using, as initial and boundary conditions, the National Centers for Environmental Prediction (NCEP) Climate Forecast System Version 2 (CFSv2) (Saha et al., 2014) reanalysis data. The resulting atmospheric profiles were evaluated using in situ radiosonde observations.

MATERIALS AND METHODS

Study Area

The study was conducted in an area comprising a transgressive dune field, located in the North Coast of Rio Grande do Sul State, Brazil (Tomazelli et al., 2008), between the cities of Tramandaí and Cidreira (Fig. 1). The dune field covers an area of approximately 30 km² and is composed of 99.53% quartz (Käfer et al., 2019, 2020) – mineral that presents characteristic spectral features (Reststrahlen band) in the TIR region (Salisbury & D’Aria, 1992). We are particularly interested on future applications of refined atmospheric profiles for atmospheric correction of TIR remote sensing data purposes. Thus, this dune field was selected as study area, once it can be assumed as a pseudo-invariant site and homogeneous in terms of emissivity (Hulley et al., 2009; Hulley & Hook, 2009). Therefore, this area can be used for validations in approaches that are based on the TIR remote sensing (Hulley et al., 2009; Li et al., 2013) in future and associated studies.

WRF Model Configuration

The Weather Research and Forecasting (WRF) model is a numerical weather prediction and atmospheric simulation system developed for both research and operational applications (Skamarock et al., 2008). In this study, the WRF version 3.9.1.1 with the Advanced Research WRF (ARW) dynamic solver was used. Reanalysis data from the NCEP Climate Forecast System Version 2 (CFSv2) (Saha et al., 2014) were utilized as initial and boundary conditions to initialize the simulations, with the 6-hourly product (00:00, 06:00, 12:00 and 18:00 UTC). These data are arranged in grids with spatial (horizontal) resolution of 0.5° x 0.5° and in 37 vertical (pressure) levels (1000 to 1 mbar). The reanalysis CFSv2 data is
preprocessed with the WRF Preprocessing System (WPS). It ingests, reformats, and interpolates the reanalysis data to the selected domains and put these inputs on the model’s vertical levels (Powers et al., 2017).

To perform the simulations the model was set for two domains, i.e. two nested grids, with horizontal resolutions of 12 km (G12) and 3 km (G03) – in 1-way nesting mode. This nesting mode refers to how the coarse grid and the fine grid interact. The fine grid boundary conditions are interpolated from the coarse grid forecast. In a 1-way mode, this is the only information exchange between the domains (Skamarock et al., 2008). Figure 2 shows the nested grids used in the simulations, in which G12 and G03 are indicated by light and dark gray shading, respectively. The asterisk indicates the point from where the outputs used in the study were extracted. This point corresponds to the grid point closest to the launch coordinate of the radiosondes in the study area (Section Radiosonde launch experiment).

Moreover, with regard to physical parameterizations, the schemes used in the study followed those adopted in Santos & Nascimento (2016), and are summarized in the Table 1.

The simulations were performed for two different dates: January 9, 2018 (Day 1) and March 14, 2018 (Day 2), with different weather conditions. On the first day, the sky was
completely covered, while in the second, there was predominance of sun with clear conditions. For both dates the simulations were started at 00:00 UTC and ended at 00:00 UTC on the following days.

Thereafter, the outputs of the ARW-WRF simulations were extracted from the point in Figure 2, containing the distribution of meteorological variables in 33 vertical levels for both grids in both days. An output file was generated each 1-hour simulation for the G12, and each 30-minute simulation for the G03 – for both dates.

Table 1 – Physical parameterization schemes used in the study.

<table>
<thead>
<tr>
<th>Parameterization</th>
<th>Scheme</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Microphysics</td>
<td>Lin et al.</td>
<td>Lin et al. (1983)</td>
</tr>
<tr>
<td>Planetary Boundary Layer</td>
<td>Yonsei University (YSU)</td>
<td>Hong et al. (2006)</td>
</tr>
<tr>
<td>Shortwave Radiation</td>
<td>Dudhia</td>
<td>Dudhia (1989)</td>
</tr>
<tr>
<td>Longwave Radiation</td>
<td>RRTM</td>
<td>Mlawer et al. (1997)</td>
</tr>
<tr>
<td>Land-Surface Model</td>
<td>Unified NOAH</td>
<td>Tewari et al. (2004)</td>
</tr>
</tbody>
</table>
**Evaluation of the simulations**

**Radiosonde launch experiment**

To evaluate the performance of the atmospheric simulations, the data simulated by the WRF model were compared with field data from radiosondes launched in the study area. These radiosonde observations were treated as ground truth data and were obtained on Day 1 and Day 2. In this experiment, the vertical structure of the atmosphere was characterized by Vaisala RS41 radiosondes (Jauhiainen et al., 2014). It consists of small sensors (temperature, humidity, and pressure) integrated in a light structure (Fig. 3a), which are launched and lifted into the atmosphere by meteorological helium balloons (Fig. 3b). The meteorological variables (i.e., temperature, humidity, and pressure) are then measured every second, transmitted to the surface by radio signals with the aid of a ground station to receive the signal (Fig. 3c), and then stored at regular intervals of 30 seconds. The other variables are calculated from these, and the wind speed and direction are estimated from the GPS information.

For the Day 1, the radiosonde was launched at 12:15:24 UTC with cloudy conditions, and it reached a final altitude of 24.81 km at 13:38:54 UTC. For the Day 2, it was launched at 13:38:20 UTC with clear-sky conditions, and it reached a final altitude of 24.64 km at 14:59:20 UTC. The launches occurred from the point of coordinates 30°04'04.1"S; 50°09'53.1"W (Fig. 1). Due to the need for infrastructure, this point was located in an adjacent position, as close as possible to the dune field.

**Evaluation and statistical methods**

In order to compare the data simulated by the ARW-WRF and those observed with the radiosondes, the model output data were interpolated from the simulated vertical levels to radiosondes coincident levels. This interpolation was performed through a weighted linear interpolation (Santos & Nascimento, 2016), Eq. (1):

\[
D(h) = \left[ d[h_1] \times \left[ 1 - \frac{(h-h_1)}{(h_2-h_1)} \right] \right] + \left[ d[h_2] \times \left[ 1 - \frac{(h_2-h)}{(h_2-h_1)} \right] \right]
\]

where \(D(h)\) is the interpolated value at the observed level \(h\), and \([h]\) is the simulated value at the two neighboring model levels \(h_1\) and \(h_2\) (Santos & Nascimento, 2016).

The interpolation was performed for each of the domains (G12 and G03) and for both days. The simulated vertical profiles were acquired at the times closest to the times when the radiosondes reached the pressure level of approximately 500 hPa, i.e., 12:33:54 UTC for Day 1, and 13:54:50 UTC for Day 2. Therefore, the interpolation was processed for the simulated profiles from 13:00 UTC and 12:30 UTC for grades G12 and G03, respectively, on Day 1. For Day 2, the interpolated WRF profiles were the 14:00 UTC for both grids.

The atmospheric profiles were compared by computing the standard statistical metrics: correlation coefficient (R), bias and the Root-Mean Squared Error (RMSE). Bias measures the model’s tendency to overestimate or underestimate a variable and is defined by Eq. (2):

\[
\text{bias} = \frac{1}{n} \sum_{i=1}^{n} y_i - o_i
\]

where \(y_i\) is the values simulated by the WRF (G12 and G03) while \(o_i\) is the value observed through radiosonde. \(n\) is the number of simulations/observations, which in this study refers to the number of vertical levels in the...
If bias is positive, the simulated values tend to be an overestimation of the observations, while if it is negative, the simulated values tend to be an underestimation of the observed ones. The bias does not provide information about the typical magnitude of the errors of individual simulations, but only an average trend. Hence, it is not a measure of precision (Wilks, 2006).

The RMSE is the square root of the mean of the individual quadratic differences between simulations and observations:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - o_i)^2} \tag{3}$$

it is considered a typical parameter for forecast/simulation errors, but is sensitive to outliers, once the errors are squared in this equation before they are summed.

The Pearson correlation coefficient (R) is a measure in a single value of the association between two variables. The R is defined as follows:

$$R = \frac{\sum_{i=0}^{n}(y_i - \bar{y})(o_i - \bar{o})}{\sqrt{\sum_{i=0}^{n}(y_i - \bar{y})^2 \sqrt{\sum_{i=0}^{n}(o_i - \bar{o})^2}} \tag{4}$$

where $\bar{y}$ and $\bar{o}$ are the arithmetic means of the simulated and observed values, respectively. R is sensitive to outliers and not sensitive to bias (Wilks, 2006).

This statistical analysis was performed for the profiles of the selected meteorological variables water vapor mixing ratio ($q$), potential temperature ($\theta$) and wind speed (Speed) (Wilks, 2006; Santos & Nascimento, 2016).

The assessment was performed throughout the entire atmospheric profile and also an analysis focusing only on the Planetary Boundary Layer (PBL). The PBL is the portion of the troposphere closest to the ground level (Stull, 2017) and where the highest concentrations of water vapor in the atmosphere are located. This analysis is valuable because the water vapor is the principal factor for atmospheric effects in optical satellite images (Sobrino et al., 1991; Jiménez-Muñoz et al., 2010). In this study, the PBL height was estimated.
using the Gradient Method (Sullivan et al., 1998) and the evaluation of $q$ and $\theta$ vertical profiles. In the Gradient Method, the PBL height is defined as the vertical position of the largest increase in $\theta$, i.e., the vertical location of the $\theta$ maximum gradient (Eq. (5), where $z_{iq}$ is the PBL height).

$$z_{iq} = z, \quad \text{where } \frac{\partial \theta(x,y,z)}{\partial z} \text{ is maximum}$$

Lastly, to manipulate the results, WRF simulations were post-processed with the WRF-ARW post-processing system (ARWpost), the data were extracted with the Grid Analysis and Display System (GrADS), and the statistical analysis was performed with R and Python languages.

**RESULTS AND DISCUSSION**

The results are presented through diagrams, comparing the vertical profiles according to the height, and tables with the computed standard statistical metrics. Their presentation will be divided between the analysis along the entire atmospheric profile and the analysis focusing on the PBL.

**Analysis along the entire atmospheric profile**

The data simulated by the WRF model were interpolated, resulting in atmospheric profiles with 124 and 121 vertical levels for Day 1 and Day 2, respectively. The visual comparison of the profiles of wind speed, potential temperature and mixing ratio plotted as a function of height is shown in Figures 4 (Day 1) and 5 (Day 2).

Analyzing the profiles, it is observed that although there are some variations, the data simulated by the WRF follow the behavior of what was observed in situ. However, a more detailed and quantitative comparison is made by plotting the errors vertically and calculating the static metrics for the entire profile. Thereby, the errors for each vertical level, and the correlation coefficient (R), bias and RMSE were computed between simulated and observed data – the results are shown in Figures 6 (Day 1) and 7 (Day 2).

The simulations presented a high correlation with the field observed data. With R values higher than 0.9, for both the parent grid (G12) and the finer grid (G03). The exception is in the wind speed profiles of Day 1, where the value of the correlation coefficient drops to 0.81 for both grids. Nevertheless, these values still indicate a strong correlation between observed and simulated data. The tendency of the model to overestimate the values of wind speed in the atmosphere condition of Day 1 is also noted, whereas for $\theta$ and $q$ the tendency is that the simulated values are lower than those observed.

On the other hand, for the weather conditions of Day 2 this average trend is inverted, being to underestimate the wind speed and overestimate $\theta$ and $q$, for the two domains. It is interesting to observe the existence of an inverse relationship between wind speed and humidity concentration and temperature, with respect to the tendency of the model to overestimate or underestimate these variables. Relative to the RMSE, the values found of wind speed were lower for Day 1, while for $\theta$ and $q$, Day 2 presented lower RMSE values.

The analysis of the vertical distribution of errors shows that for the day with clear-sky conditions the difference between the two WRF simulated grids are irrelevant. While on a cloudy day, the differences between the errors of the two grids are more noticeable. Mainly for $q$ at the levels closest
Figure 4 — Vertical profiles of wind speed (a), potential temperature (b), and water vapor mixing ratio (c) as a function of height: field observed/radiosonde and WRF simulated with horizontal resolutions of 12 km (G12) and 3 km (G03). For January 9, 2018 (Day 1).

Figure 5 — Vertical profiles of wind speed (a), potential temperature (b), and water vapor mixing ratio (c) as a function of height: field observed/radiosonde and WRF simulated with horizontal resolutions of 12 km (G12) and 3 km (G03). For March 14, 2018 (Day 2), of 12 km (G12) and 3 km (G03). For January 9, 2018 (Day 1).
Figure 6 — Errors at each vertical level between local radiosonde and WRF (G12 and G03): wind speed (a), potential temperature (b), and water vapor mixing ratio (c). For Day 1 (cloudy).

Figure 7 — Errors at each vertical level between local radiosonde and WRF (G12 and G03): wind speed (a), potential temperature (b), and water vapor mixing ratio (c). For Day 2 (predominance of sun and clear-sky).
to the surface. Despite this, even in the case of Day 1, the overall differences between G12 and G03 are not considered as significant. In addition, for the two days, errors in Speed are distributed along the profile. While in $\theta$, errors are relatively more concentrated at higher altitudes. In contrast, $q$ biggest ones are concentrated in the first levels of the atmosphere. Clearly when the presence of water vapor in the atmosphere is lower, the $q$ errors also tend to be. This contributes to the low overall values of biases and RMSE along the entire $q$ profile. The region where the biggest $q$ errors are located comprises the PBL. That is why it is essential to analyze the statistical metrics considering exclusively this layer.

**Analysis focusing on the PBL**

The PBL height estimate for Day 1 was close to 1.9 km (for both observed and simulated profiles), thus including 12 vertical levels. While for Day 2 the estimated heights were approximately 4.8 km, counting on 27 vertical levels. Figures 8 and 9 show the same vertical profiles of the previous session, except that only the first kilometers of the atmosphere are represented here, aiming to focus on the PBL.

Likewise, Table 2 presents the recalculated statistical metrics encompassing only the initial levels of the atmosphere – up to the estimated height of the top of the PBL.

When the analysis becomes only for the PBL, most of the high correlations remain. However, for Day 1 the R values for the wind speed were changed to 0.73 in G12 and 0.74 in G03. This reduction in the correlation is in agreement with the great presence and influence of the turbulence in the boundary layer, with the friction tension decreasing with the height (Foken, 2008; Stull, 2017). Nonetheless, it is for the water vapor mixing ratio of Day 2 that the largest variations occur between the data observed and those simulated by the WRF. Within the PBL, the correlation coefficient assumes a value of 0.68 in both grids. These results are consistent with the greater sensitivity of the vertical flow of humidity to the planetary boundary layer, surface layer and microphysical parametrization schemes (Hu et al., 2010; Ruiz et al., 2010; Santos & Nascimento, 2016).

In addition, for the cloudy day, the average tendency of the model was to overestimate the wind speed values and to underestimate the potential temperature and the mixing ratio in the two grids. As for the condition of predominance of sun and little cloud cover, the simulation tended to overestimate the wind speed and $q$, and to underestimate the observed values of $\theta$, also for the two nested grids.

The resulting RMSE values were slightly lower for the simulations of Day 1, and generally lower when compared to those resulting from the whole profile analysis, for wind speed. For the potential temperature within the PBL, the RMSE was lower on Day 2 in G12 and, in G03, on Day 1. These RMSE values were lower than those calculated along the entire profile. While for the water vapor mixing ratio, the RMSE values increase when the calculation includes only the boundary layer levels, being within the PBL lower on Day 1. Ratifying what was discussed in Figures 6 and 7, the biggest errors between the simulated and observed values are located in the highest altitudes of the profile for $\theta$, and in the first kilometers for $q$. 
Figure 8 – Vertical profiles of wind speed (a), potential temperature (b) and water vapor mixing ratio (c) as a function of height – for the first levels of the atmosphere: field observed/radiosonde and WRF simulated with horizontal resolutions of 12 km (G12) and 3 km (G03). For January 9, 2018 (Day 1).

Figure 9 – Vertical profiles of wind speed (a), potential temperature (b) and water vapor mixing ratio (c) as a function of height – for the first levels of the atmosphere: field observed/radiosonde and WRF simulated with horizontal resolutions of 12 km (G12) and 3 km (G03). For March 14, 2018 (Day 2).
Table 2 - Statistical metrics comparing the observed (through radiosondes) profiles with those simulated by the WRF (G12 and G03) – for the Planetary Boundary Layer (PBL).

<table>
<thead>
<tr>
<th></th>
<th>Day 1 – Cloudy – PBL</th>
<th>Day 2 – Clear-sky – PBL</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRF G12 - Observed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed θ q</td>
<td>Speed θ q</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.73 0.99 0.99</td>
<td>0.96 0.99 0.68</td>
</tr>
<tr>
<td>Bias</td>
<td>1.24 -0.06 -0.90</td>
<td>0.38 -0.28 0.66</td>
</tr>
<tr>
<td>RMSE</td>
<td>1.61 0.71 0.96</td>
<td>1.85 0.55 2.01</td>
</tr>
</tbody>
</table>

Furthermore, an important point to be reiterated is that in the simulations carried out for Day 1, there is a greater differentiation of the results of the two nested grids between them, when compared to Day 2. Particularly, q bias and RMSE values are higher in the finer resolution grid. Nevertheless, especially when considering the entire atmospheric profiles and the weather conditions of Day 2, the differences between nested grid results are minimal. The processes that influence the humidity content in the atmosphere are relative to the synoptic scale (e.g., low level jets (Vera et al., 2006; Do Nascimento et al., 2016)) and are not as sensitive to small differences in horizontal resolution (Santos & Nascimento, 2016; Lin et al., 2018), on the other hand local factors do. Local circulation, land use, surface models, and PBL parameterization have a major influence on the simulation of water vapor content and vertical transport in the atmosphere (Ruiz et al., 2010; Hari Prasad et al., 2017). Therefore, horizontal resolution is a way of determine the local characteristics of the study area. In a coarse resolution simulation, the grid point chosen to extract the vertical atmospheric profile can be in a distant location, with different local characteristics. Thus the availability of local humidity can be completely different and this influence will be more evident in the first levels of the atmosphere (within the PBL). So the importance of using a mesoscale model to include a more detailed local description and circulation. The minimal differences between the simulated profiles extracted from the nested grids G12 and G03 demonstrates a lower influence of the horizontal resolution in the simulation of the vertical profile of the atmosphere. It can suggest that a resolution of approximately 12 km represents a good compromise between a more spatially detailed simulation and computational cost for a domain covering this study area (Lin et al., 2018).

CONCLUSION

In this paper we investigated the potential of the WRF model in the generation of vertical atmospheric profiles as an alternative to the need for a local radiosonde. Atmospheric simulations were performed with the ARW-WRF model using CFSv2 reanalysis data as initial and boundary conditions. In order to evaluate the performance
of these simulations, the results were compared with field data obtained by launching radiosondes in the study area for different weather conditions.

The results indicated that there is an overall positive indication of this potential, once high correlations were seen between simulated and observed data, with correlation coefficients mostly higher than 0.9. With exceptions for wind speed in the weather conditions of Day 1. And the water vapor mixing ratio, when analyzed exclusively within the boundary layer on Day 2 – reaching an R of 0.68. However, even in these cases of lower correlation, it can be considered moderate and occur in already expected and coherent situations.

Relatively low values of bias and RMSE were found for the meteorological variables analyzed. For the bias, the lowest value occurred at the potential temperature (-0.06 K) for Day 1 at G12 grid, for PBL analysis. Whereas, the highest value was the positive bias of 1.31 m/s for G03, under the same conditions. Regarding the RMSE, the lowest value was observed in $\theta$: 0.44 K, in G03 and Day 1, in the boundary layer. On the other hand, the highest value of RMSE was 3.00 K and it was given in the analysis along the whole profile in G03, for Day 2 and also in $\theta$.

The results also showed a minimum difference between the profiles of the two different grids. This is an indication of the lower influence of the horizontal resolution on the simulated vertical atmospheric profiles in this case. Suggesting that a resolution of approximately 12 km it is enough to represents a good balance between a local detailed simulation and computational cost.

Since radiosondes are rarely available for specific sites and time, this study has shown that using the WRF model in the generation of atmospheric profiles is a good alternative for radiosondes absence. Despite prior model testing can be used to determine optimum model configurations; the best configuration is a function of location and meteorological conditions. Our results point to the WRF model as a useful tool in the simulation of the vertical profile, but local studies are always necessary.

Future work will focus on testing different physical parameterization schemes. In particular, those related to PBL and surface layer parametrization, trying to adjust the issues related to the humidity vertical flow and the water vapor mixing ratio. Additionally, another paper is in process, in which we apply and evaluate the results in the atmospheric correction of images. Finally, another requiring attention is to invest in the increase of the vertical resolution in the simulations, due to the finding of the lower impact of the horizontal resolution on the performance of the model.

ACKNOWLEDGMENTS

The CFSv2 reanalysis data are a courtesy of the National Centers for Environmental Prediction (NCEP), available at the Research Data Archive (RDA) – National Center for Atmospheric Research (NCAR). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES), finance code 001. The launch of local radiosondes was carried out in partnership with the Meteorology and Air Quality Laboratory (LMQA) – CEPSRM/UFRGS. The authors are also grateful to the anonymous reviewers for the valuable comments.

REFERENCES


HONG S-Y, NOH Y & DUDHIA J. 2006. A New


Brazilian Journal of Geophysics, 38(2), 2020


Received on July 23, 2020 / Accepted on September 2, 2020