Influence of high-resolution surface databases on the modeling of local atmospheric circulation systems

L. M. S. Paiva¹, G. C. R. Bodstein², and L. C. G. Pimentel²

¹Federal Center of Technological Education, Rio de Janeiro, Brazil
²Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

Received: 16 October 2013 – Accepted: 30 October 2013 – Published: 16 December 2013

Correspondence to: G. C. R. Bodstein (gustavo@mecanica.coppe.ufrj.br) and L. C. G. Pimentel (pimentell65@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

Large-eddy simulations are performed using the Advanced Regional Prediction System (ARPS) code at horizontal grid resolutions as fine as 300 m to assess the influence of detailed and updated surface databases on the modeling of local atmospheric circulation systems of urban areas with complex terrain. Applications to air pollution and wind energy are sought. These databases are comprised of 3 arc-sec topographic data from the Shuttle Radar Topography Mission, 10 arc-sec vegetation type data from the European Space Agency (ESA) GlobCover Project, and 30 arc-sec Leaf Area Index and Fraction of Absorbed Photosynthetically Active Radiation data from the ESA GlobCarbon Project. Simulations are carried out for the Metropolitan Area of Rio de Janeiro using six one-way nested-grid domains that allow the choice of distinct parametric models and vertical resolutions associated to each grid. ARPS is initialized using the Global Forecasting System with 0.5°-resolution data from the National Center of Environmental Prediction, which is also used every 3 h as lateral boundary condition. Topographic shading is turned on and two soil layers with depths of 0.01 and 1.0 m are used to compute the soil temperature and moisture budgets in all runs. Results for two simulated runs covering the period from 6 to 7 September 2007 are compared to surface and upper-air observational data to explore the dependence of the simulations on initial and boundary conditions, topographic and land-use databases and grid resolution. Our comparisons show overall good agreement between simulated and observed data and also indicate that the low resolution of the 30 arc-sec soil database from United States Geological Survey, the soil moisture and skin temperature initial conditions assimilated from the GFS analyses and the synoptic forcing on the lateral boundaries of the finer grids may affect an adequate spatial description of the meteorological variables.
1 Introduction

Numerical modeling of unsteady three-dimensional turbulent atmospheric flow is a natural approach to describe mean properties of various physical processes that are not often captured by field measurements collected at a few scattered points in space. One of the most robust, versatile and modular mesoscale models designed to resolve atmospheric flows in many scales is the Advanced Regional Prediction System (ARPS; Xue et al., 1995, 2000, 2001). Its numerical code has been developed and disseminated freely in the scientific community to allow changes and developments in its subroutines. The ARPS code solves a set of Partial Differential Equations (PDE’s) written for the three-dimensional compressible time-dependent atmospheric flow, under dry, moist and non-hydrostatic conditions. ARPS is a typical mesoscale model that can be run at fine space and time resolutions in order to be able to represent highly complex terrain based on the high-resolution surface databases that are available nowadays. ARPS incorporates heterogeneous land-surface conditions and time-dependent synoptic boundary forcing, but they are typically limited by outdated coarse resolution. As other current mesoscale-models, such as Weather Research and Forecasting – WRF (Skamarock et al., 2001, 2005), Fifth-Generation of National Center for Atmospheric Research – NCAR/Penn State Mesoscale Model – MM5 (Grell et al., 1993; Dudhia et al., 2005), Regional Atmospheric Modeling System – RAMS (Walko et al., 1995), Meso-Eta Model (Black, 1994), Mesoscale Non-Hydrostatic Model – Meso-NH (Lafare et al., 1998) and Aire Limitee Adaptation Dynamique Developement International – ALADIN (Bubnová et al., 1993; Radnóti et al., 1995; Horányi et al., 1996), ARPS allow significant refinement of the numerical grid to the point where LES can be used, since some turbulent parametric models developed for LES are available in the code. We chose the LES-ARPS model as our main tool because it is based on a 1.5-order TKE scheme and the Moeng and Wyngaard (1989) turbulence model, and because it has been thoroughly tested (Chow, 2004; Chow et al., 2006) and used as
Several numerical studies available in the literature have adopted significant refinement of the grid in mesoscale simulations. As an example, we may cite the simulations carried out by Grell et al. (2000), who used MM5 to compute the atmospheric flow in some regions of the Swiss Alps with horizontal resolutions of up to 1 km. It’s noteworthy that previous works, such as Lu and Turco’s (1995), point out that the increase in the spatial resolution of the grid can generate more detailed and reliable solutions. In fact, most studies show that high-resolution numerical grids tend to improve the quality of numerically simulated data when compared to observed data (Revell et al., 1996; Grønås and Sandvik, 1999; Grell et al., 2000; Chow et al., 2006). However, in regions of steep and extensive slopes, the topography may be poorly represented because the ramps that form the slopes on the surface may become irregular as the grid resolution is increased. In this case, the coarse spatial resolution of the topographic database may not add any additional information in simulations with fine numerical grids. Chow (2004) expresses the same concern in his simulations. Usually, high-resolution numerical grids are often employed for simulations in small areas, since the number of grid points grows excessively as the resolution increases, which implies in a high computation cost. Revell et al. (1996) and Grønås and Sandvik (1999) used numerical grids with resolutions of 250 m to perform LES simulations, but the wind field was not reproduced accurately in the regions studied. The authors considered that the main source of inaccuracy was the absence of high-resolution surface data. On the other hand, Zhong and Fast (2003) were successful in capturing the general characteristics of the surface fluxes present in the Salt Lake Valley, in the state of Utah, using three of the mesoscale models cited above: RAMS, MM5, and Meso-Eta Model. All models were initialized using synoptic data and used horizontal grid resolutions of 560 and 850 m, which are close to the topographic database resolution of the models. Even so, the simulations were not able to capture the local circulation and the surface fluxes. In order to improve the results, Zhong and Fast...
suggested changes in the vertical mixing terms, in the radiation model and in the parameterizations adopted for the surface fluxes. Chen et al. (2004) have used ARPS to also simulate the atmospheric flow in the Salt Lake Valley. The results were more satisfactory because they increased the numerical domain size and the horizontal resolution to 250 m. Sensitivity tests have been performed by Chow et al. (2006) running ARPS in LES mode for the Riviera Valley, situated in the Swiss Alps. Their numerical results were in good agreement with field data recorded during the 1999 campaign of the Mesoscale Alpine Programme (MAP Riviera Project; Rotach et al., 2004), although sensitive the soil temperature and moisture initialization. In simulations of the type discussed here, which are characterized by short spin-ups, the initialization of soil moisture and skin temperature may become one of the main issues of the modeling, since it may require off-line models to provide proper initial conditions. Chow et al. (2006) have tried many different ways to solve this problem, but the statistical indices were still lower than expected.

In the numerical studies performed by Hanna and Yang (2001), who have used four different mesoscale-models in their simulations, the discrepancies that appeared in the calculation of the wind direction and speed were attributed to the misrepresentation of the turbulent fluxes, mainly due to the land-use database and the subgrid-scale parameterization adopted. On the other hand, Zängl et al. (2004) and Gohm et al. (2004) simulated foehn winds using MM5 with two-way nested grids and found non-negligible differences between simulated and observed data even using a horizontal resolution of 267 m for the innermost grid. Based on that they concluded that the topographic data require high resolution and that the lateral BC’s are poorly satisfied if the grid points are too far apart from each other. In contrast, Zängl et al. (2004) found that the effect of the horizontal computational mixing model was larger than the effect of grid resolution. Their model performed better with an improved computational mixing scheme at coarse resolution (3 km) than with the traditional mixing scheme at fine resolution (1 km). As a matter of fact, most studies also point out that the soil and the vegetation databases are also important sources of error. De
Wekker et al. (2005), using RAMS, have shown good agreement between numerical and observed data, but their modeling did not capture accurately the wind structure in a region characterized by valleys and mountains, even though their grid resolution of 333 m was very fine. The probable cause was the bad representation of the topographic database provided by RAMS.

Many numerical weather- and climate-prediction models are sensitive to the heat and moist surface fluxes (Beljaars et al., 1996; Viterbo and Betts, 1999). Because these surface transport processes occur in the subgrid scales, they cannot be solved directly and, therefore, they need to be parameterized. In practice, the moist fluxes at the soil surface are estimated by soil and vegetation models (Pitman, 2003). The transfer of moisture is usually described by semi-empirical aerodynamic coefficients, which are based on the similarity functions presented by Businger et al. (1971) and Deardorff (1972). Recently, Weigel et al. (2007) showed that the moist flux from the soil surface to the atmosphere is not controlled by the turbulent eddies only. The authors note that other mechanisms are also important, such as the mass transport due to the geometry of the topography and the interactions that exist in the thermally induced circulations present in regions of valleys and mountains.

Recent studies have shown that LES has been adopted frequently, mainly due to increased computational power to solve high-resolution atmospheric flow. Wyngaard (2004) observed that LES is not restricted to applications where the flow occurs in the smallest resolvable turbulent scales. Chow et al. (2006) and Weigel et al. (2006, 2007) indicated that the complex thermal structure and dynamics of the atmospheric flow over the complex terrain present in the Swiss Alps may be reproduced in detail using ARPS with Moeng and Wyngaard’s (1989) LES model turned on. Michioka and Chow (2008) also showed that ARPS performs well when configured to run in the LES mode. These authors coupled ARPS to a code that calculates the dispersion of passive pollutants and ran simulations in regions of highly complex terrain using one-way nested grids, where the highest resolution was 25 m in the horizontal directions. Recently, Chow and Street (2009) implemented a new turbulent-flux parameterization
model in ARPS in the form of a Taylor-series expansion, which aims to reconstruct the resolved sub-filter-scale turbulent stresses. Variations of this series expansion are combined with dynamic eddy-viscosity models for the subgrid-scale stresses to create a Dynamic Reconstruction Model – DRM (Chow, 2004; Chow et al., 2005). The authors evaluated the performance of DRM computing the flow over the Askervein hill (Taylor and Teunissen, 1985, 1987) and found promising results. The Atmospheric Boundary Layer (ABL) flow over the Askervein Hill has also been studied by Chow and Street (2009), who added a new stress tensor in the ARPS code to compute LES. The authors concluded that the use of explicit filters and DRM avoids common problems in the ground-surface Boundary Conditions (BC’s) that appear in the LES model and that the results were quite satisfactory when compared to other field data and numerical models (Castro et al., 2003; Lopes et al., 2007).

The main objective of this paper is to evaluate the local atmospheric circulation system of the Metropolitan Area of Rio de Janeiro (MARJ), Brazil, by setting up a high-resolution numerical model using ARPS in the LES mode as the main tool, and having detailed and updated topography and land-use databases incorporated into the model. New pre-processors are developed and incorporated into ARPS to assimilate the information from the database files of the Shuttle Radar Topography Mission (SRTM; Farr and Kobrick, 2000) and the European Space Agency (ESA) GlobCover (Bicheron et al., 2008) and GlobCarbon (Eyndt et al., 2007; Arino et al., 2008) projects in order to generate appropriate non-homogeneous surface BC’s for the present model.

2 Site characterization, period synoptic analysis and surface stations data

In order to assess the capability of our high-resolution numerical model to compute the local atmospheric circulation over a complex-terrain region we perform numerical simulations for the Metropolitan Area of Rio de Janeiro (MARJ). The MARJ is characterized by complex topography, nonhomogeneous land-use and land-cover surfaces, and is surrounded by water bodies, such as the Sepetiba Bay, the Guanabara
Bay and the Atlantic Ocean (see Zeri et al., 2011). Also, the MARJ is influenced by the South Atlantic Subtropical Anticyclone (SASA), by low frontal pressure systems on the synoptic scale, and by breeze systems on meso and local scales (sea/land and valley/mountain). The influence of the several atmospheric scales over the local circulation in this region becomes a major challenge for assessment of the performance of the mesoscale model ARPS. The MARJ is a high population-density area that accounts for about 80% of the population of the State of Rio de Janeiro, Brazil, and contributes the most to the emission of pollutants into the atmosphere. The main reasons for that are the existence of many industrial and mobile pollution sources that have emerged after new investments in infrastructure have been made that are associated with the expansion of the Itaguaí Harbor, the installation of the Steel Company of the Atlantic Ocean and the existing petrochemical industry. The high emission rates of pollutants in conjunction with the characteristics of the local atmospheric flow lead to the degradation of the air quality of the MARJ’s air basins. There are a few surface weather-observation stations and just one upper-air station in the MARJ, indicating the need to use high-resolution simulated data of the atmospheric flow to provide support to the air quality modeling in the region.

Our simulations are carried out for a physical time of 48 h, between 00:00 UTC on 6 September 2007 and 00:00 UTC on 8 September 2007. In this period, the synoptic analyses have indicated the dominance of the SASA over the MARJ. This is a system of semi-permanent high pressure, characterized by the presence of horizontal synoptic winds that rotate counter-clockwise, vertical subsidence wind that generates divergence near the surface, clear sky, calm weather, stable conditions and such that the SASA’s location and intensity changes seasonally (Richter et al., 2008; Zeri et al., 2011). The influence of the SASA contributes to inhibit cloudiness and the advancement of high-latitude frontal-systems in the region of interest (Lucena et al., 2012). Meteorological mesoscale and microscale systems, as the sea-breezes that act in the MARJ, can be hidden by the SASA system, but they are not totally destroyed as is the case when fronts pass by the area. During this period, the SASA system
remained mostly over the Atlantic Ocean, between latitudes −50 and 0 arc-deg south and longitudes −50 and 0 arc-deg west, whereas the directions of prevailing synoptic-scale winds were northeast and east in the MARJ.

For comparison, we use hourly observational data for potential temperature, water–vapor mixing ratio, and wind direction and speed obtained from seven available surface weather-observation stations located in different parts of the MARJ, as seen at Fig. 1 and Table 1. The data from Marambaia and Ecologia Agricola surface stations were obtained from the Brazilian National Institute of Meteorology, and the METAR code data of the aerodromes Santa Cruz (SBSC), Campo dos Afonsos (SBAF), Jacarepagua (SBJR), Santos Dumont (SBRJ) and Galeão (SBGL) were obtained from the Meteorology Network of the Brazilian Air Force Command. It’s worth mentioning that SBGL has also sounding data available.

3 Numerical modeling setup

The procedures we employed to run accurate numerical simulations of the atmospheric flow in the MARJ are described here. These procedures can also be applied to other regions of the Earth planet as well, since we have developed new subroutines to process all the satellite data needed. The steps taken include the setup of the high-order numerical method of ARPS, the structure of a high-resolution one-way nested grid, the incorporation of a detailed and updated topography and land-use databases on ARPS, and the more sophisticated choice of radiation, turbulence closure, microphysics and cumulus parameterizations. A control run (CTL) set up with the (outdated) original ARPS surface databases serves as a reference for comparison between the high-resolution (HR) surface database final runs and the field observational data.
3.1 Numerical schemes

We setup ARPS to employ a fourth-order spatial differencing for the advection terms and a mode-splitting technique for the temporal discretization to accommodate high-frequency acoustic waves. Large time steps ($\Delta t$) are chosen based on the leapfrog method. For the small time steps ($\Delta \tau$) we use first-order forward-backward explicit time stepping, except for terms responsible for vertical acoustic propagation, which are treated semi-implicitly.

3.2 Grid nesting and topography

We adopt a one-way nested-grid structure that is set up based on the tutorials of Mesinger and Arakawa (1976) and Warner et al. (1997). An external grid (GEXT) is setup in order to pre-process the data from the 0.5°-Global Forecasting System analyses (0.5°-GFS; Kanamitsu, 1989), and to produce the first Initial Conditions (IC’s) and lateral BC’s at 3 h intervals for the outermost domain (G1), which has 27 km of horizontal resolution. Relaxation on the values of the BC’s is applied to a 5–10 grid-cell zone around the domain boundary, depending on the grid. All simulations were carried out continuously for 48 h of physical time to produce output data at hourly intervals, starting at 00:00 UTC on 6 September 2007. The output data computed by a coarse grid is also employed as IC’s and BC’s at 3 h intervals for a subsequent fine-grid simulation in the one-way nested-grid domain. In order to determine the horizontal resolutions for grids G1 to G4 of the one-way nested-grid setup, we employ a ratio of one grid size to the next equal to 3, all centered with respect to the coarsest grid. The finest grids (G5 and G6) have a horizontal resolution of 300 m (see Table 2) and are not nested to each other, as illustrated in Fig. 1. This setup has been chosen because we are interested in understanding how the atmospheric local circulations behave when they cross the boundaries of air basins I and II–III of the MARJ, which are covered by the G5 and G6 domains, respectively. This grid setup also helps to reduce the computational cost that we would obtain if a single domain were used with the finest-
grid resolution. It is important to remember that the higher-resolution grids process information coming from the 0.5°-GFS analyses and results of the previous simulation performed on the coarser-resolution grids, which embed physical effects of the various scales being modeled. When one-way nested grids are used, the energy transfer occurs from large to small scales. Nevertheless, this procedure used by ARPS allows that each vertical resolution be treated separately for each numerical grid, ensuring that the flow computation is performed in the LES mode, and which we found to be necessary to accommodate the steep terrains of the MARJ. Currently two-way nesting schemes are available in other codes, such as RAMS, that do not allow for vertical resolution changes, but the effect of two-way nesting will be explored in future work. The grid domains from GEXT to G4 are centered at a point located at −22.82 arc-deg south-latitude and −43.50 arc-deg west-longitude, whereas the center of grid G5 is slightly shifted to −22.88 arc-deg south-latitude and −43.72 arc-deg west-longitude and G6 to −22.80 arc-deg south-latitude and −43.27 arc-deg west-longitude, as can be seen in Fig. 1. A Mercator projection is used with the true latitude and longitude located at the center of each domain to minimize distortion in the main area of interest, particularly in smaller areas that have the highest resolutions. For each nested subdomain, the terrain is smoothed next to its boundary to match the local neighboring grids that have lower resolution.

The ARPS original files of the 30 arc-sec USGS (≈ 900 m) topography database are pre-processed for grids G1 to G3 in our simulations. Depending of the run configuration, the ARPS files for grids G4 and G6 are either pre-processed from the 30 arc-sec USGS (≈ 900 m) topography database or the 3s-SRTM (≈ 90 m) detailed and updated high-resolution topography database, which we have recently incorporated into ARPS. This incorporation procedure required substantial modifications to the arpstrn.f90 source file of the original ARPS code, which can be downloaded directly from http://meteoro.cefet-rj.br/leanderson/arps/ and run with the 3s-SRTM data on the ARPS model version 5.2.8. Details on how the arpstrn.f90 routine works can be seen in the file’s comment lines and also Xue et al. (1995). In Fig. 2a and b it is possible to compare
these two distinct databases processed for G5, where we can easily notice that the topographic details are much better reproduced by the 3s-SRTM data, especially in regions where the topography exceeds 800 m a.g.l. The high-resolution topography database allows a much more appropriate definition of the surface boundary condition. Although a comparison between these databases for grid G6 is not shown the east side of the G5 domain provides a good idea of what the topography maps look like for G6.

3.3 Vertical resolution and grid aspect ratio

ARPS incorporates a $\sigma$-coordinate system that follows the ground surface. The grids are stretched using a hyperbolic tangent function (Xue et al., 1995) that produces an average spacing $\Delta z_{med}$ and a domain height equal to $(n_z - 3)\Delta z_{med}$. The smallest vertical spacing $\Delta z_{min}$ used in each grid can be found in Table 2. To resolve the smallest structures of the atmosphere it is necessary to adopt high vertical resolutions, but the grid aspect ratio on the surface ($\Delta / \Delta z_{min}$, where $\Delta = \Delta x = \Delta y$) should not be extremely large to avoid numerical errors, especially in the horizontal gradients (Mahrer, 1984), and also to avoid distortion of the resolvable turbulent structures when runs are carried out in LES mode (Kravchenko et al., 1996). Poulos (1999) and De Wekker (2002) have found that the aspect ratio of the grid should be small, especially for terrains with steep topography. Following the tutorials of Mahrer (1984), Kravchenko et al. (1996), Poulos (1999) and De Wekker (2002) and similar procedures adopted by Chow et al. (2006) we set up grids G1 and G2 with grid aspect ratio of 540 and 180 near the surface, respectively, to represent the scales in the atmosphere. These choices are adequate because the characteristic scales of the topography and the resolvable flow are large enough for the G1 and G2 domains, in addition to the fact that the 1.5-TKE parameterization scheme for the closure of the turbulent fluxes are used in both grids. Tests with high values of $\Delta z_{min}$ for grids G1 and G2 degraded the representation of the synoptic structures. The same proportion was used for the G3 and G4 aspect ratios, since their $\Delta z_{min}$ values are equal to the coarser grids. Particularly, for the G5 and
G6 grids, we avoided to increase the aspect ratio more than necessary and impose a substantial decrease in the value of $\Delta z_{\text{min}}$. Thus, we adopted an aspect ratio equal to 15, which results in a first level at 10 m.a.g.l. and $\Delta z_{\text{min}}$ equal to 20 m for the finest grids.

### 3.4 Land-use databases

In addition to the atmospheric model component, our simulations work together with the ARPS soil-vegetation model that has been constructed based on the parametric scheme developed by Noilhan and Planton (1989) and modified by Pleim and Xiu (1995). This scheme is in function of the soil and vegetation types, vegetation cover fraction and the Normalized Difference Vegetation Index (NDVI) and/or Leaf Area Index (LAI) processed by the numerical grids. Thus, it is important that the databases contain detailed updated information about urban and vegetated areas and water bodies with relatively high spatial resolution to allow more suitable non-homogeneous BC’s to simulate accurately the atmospheric flow, mainly in LES mode. ARPS models physically 13 soil-types (including water and ice) and 14 vegetation-types, according to the classification of the United States Department of Agriculture (USDA).

For the soil-type representation, the original 30s-USGS database files are processed and mapped into the grid categories set up in ARPS by selecting values from the nearest data points. For the vegetation-type representation, original files of either the 30s-USGS database or the 10s-ESA ($\approx 300$ m) GlobCover-project database that we incorporated into ARPS are employed, depending on the run. The incorporation of the vegetation data into ARPS is carried out through the modifications that we have introduced to the original arpsssfc.f90 and arpsssflcib.f90 source files. Particularly, we have developed and added two new subroutines to the arpsssfclib.f90 source file, referred to as GET_10S and MAPTY10S, which are similar to the GET_30S and MAPTY30S original subroutines. The GET_10S subroutine reads 10 arc-sec or 300 m vegetation type data resolution files from the ESA GlobCover project. The MAPTY10S subroutine transforms the 22 categories from 10 arc-sec vegetation-type...
data into the simpler 14 original vegetation-type USDA/ARPS categories and feed them into the model domain by choosing the data values at the nearest grid points. Beyond that, the setting of the surface-roughness \( z_0 \) map is processed by choosing values associated to the vegetation-type classes, according to the same conversions shown in Table 3. Both the arpsssf.f90 and arpsssfclib.f90 source files have been commented to explain the modifications and they can been downloaded directly from http://meteoro.cefet-rj.br/leanderson/arps/.

When the 30s-USGS vegetation-type database files are adopted in our runs, LAI is calculated from 30s-USGS monthly NDVI database for herbaceous vegetation and trees, respectively (Xue et al., 1995). The relation between the NDVI and LAI for herbaceous vegetation can be consulted in Asrar et al. (1984), and for trees in Nemani and Running (1989). Also, vegetation fraction data from the 30 arc-sec National Environmental Satellite Data and Information Service (30s-NESDIS), supported by the National Oceanic and Atmospheric Administration (NOAA), are derived from the same NDVI data using the methodology suggested by Gutman and Ignotov (1998). On the other hand, whenever the 10s-ESA vegetation-type database files are employed in ours runs, LAI and vegetation fractions are directly obtained from the 30s-ESA GlobCarbon project database, and little corrections on the mapped 30s-USGS soil-type data are needed near the coastlines of the water bodies as illustrated by comparison in Fig. 3a and b for the G5 domain. We point out that we have developed the GETLAIGLOBCARBON and the GETFAPARGLOBCARBON subroutines and included them into the arpsssfclib.f90 source file of the ARPS code. These routines are able to read the LAI and calculate the vegetation-fraction values, respectively, from the GlobCarbon project database and interpolate them into the model domain. In addition, we have also developed the MAPTYLAIGLOBCARBON and MAPTYFAPARGLOBCARBON subroutines to transform the 30s-ESA LAI and the FAPAR data into the simpler LAI and vegetation-fraction data and feed them into the model domain by choosing the data values at the nearest grid points. For an adequate reproduction of our results we also provide the namelist files, i.e., the arps.input files.
and the SRTM and ESA databases at http://meteoro.cefet-rj.br/leanderson/arps/, in order to allow any set up we have used to run on all domain grids employed in this work.

Figure 4a and b highlight a comparison between the vegetation-type maps processed by G5 with the 30s-USGS and the 10s-ESA databases, respectively. It can be noticed that 10s-ESA vegetation-type mosaic presents more detailed and smoothed areas than the 30s-USGS database. The Figs. 5a and b, 6a and b and 7a and b illustrate the comparison of the surface roughness, LAI and vegetation fraction maps for both cases at G5. The ESA land-use database seems better than the USGS.

Additionally, we use two soil layers with depths of 0.01 and 1 m for the computation of the temperature and moisture balances according to the ARPS soil-vegetation model. The Sea Surface Temperature (SST) and the soil skin temperature and moisture initial databases to G1 grid are obtained from the 0.5°-GFS analyses. For the subsequent grids, initial values of these surface characteristics are obtained by numerical interpolation performed in each preceding grid. We also point out that, for all grids, we adopted the Colette et al. (2003) topography shading scheme, the Chou (1990) and Chou and Suarez (1994) short and long-wave radiation schemes, the Kain and Fritsch (1990, 1993) microphysics scheme and the 1.5-Turbulent Kinetic Energy (TKE) Moeng and Wyngaard (1989) turbulence model. The Kessler (1969) and Lin et al. (1983) cumulus scheme is turned on only for the G1 and G2 synoptic-grids. The Table 4 summarizes the differences adopted in our one-way nested-grid runs.

4 Results and discussion

The simulations are performed in parallel on a cluster comprised of three identical machines with Intel Xeon E5450 processor of 3.0 GHz of RAM and cache size of 6144 KB. Our runs cover the period from 6 to 7 September 2007, which is long enough to infer the dependence of the simulations on initial and boundary conditions, topographic and land-use databases and grid resolution, and to compare the results to
surface and upper-air observational data. We anticipate that, in the absence of frontal systems and depending on the positioning of SASA in the southeastern region of Brazil, the prevailing winds that blow in the region of interest are the result of mesoscale and microscale mechanisms that occur as a function of the land–sea contrast, mountain–valley and land-use.

4.1 Meteograms, upper-air profiles and statistics indexes

Figures 8–14 show the time cross-section data, or meteograms, of potential temperature, water–vapor mixing ratio \( (q_v) \) and wind direction and speed that compare the simulated data from the CTL and the HR runs with the observational data from the surface weather stations. We clearly notice that there is no significant discrepancy between the CTL and the HR short spin-up runs for all surface weather stations, except in relation to the wind direction for some periods of time at a few surface stations. In general, the simulated results underestimate the potential temperature and the water–vapor mixing ratio in comparison with the observational data and overestimate the wind speed. Among these variables the largest differences occur for the water–vapor mixing ratio, because the comparison is carried out between the values computed from the first level of the G5 and G6 grids (10 m a.g.l.) and the data observed from the surface level of the weather stations, which is normally at 2 m a.g.l. This may also be related to the influence of the initialization procedure for soil temperature and moisture from 0.5°-GFS. On the other hand the largest discrepancies between simulated and observed data occur for the wind direction time series.

Specifically at the Marambaia station, the potential-temperature daily cycle is well simulated for both days when compared with the observational data (Fig. 8a and b). However, for 7 September 2007 the difference between the simulated and observed data for the water–vapor mixing ratio is relatively high (about 2 g kg\(^{-1}\)), as Fig. 8d shows. Good agreement between the simulated and the observational data for the wind direction and speed can be noticed at almost all times, mainly for the HR run (Fig. 8e–h). We can also see a predominance of observed southeast (SE) winds
at 14:00–16:00 UTC that turns clockwise to form south (S) winds at 17:00 UTC on 7 September 2007 (Fig. 8f). This change may be more associated to the sea-breeze winds coming from the Atlantic Ocean than to the sea-breeze mechanism generated in the interior of the Sepetiba Bay. The simulation results show just a slight discrepancy with respect to this behavior, since ARPS computes SE winds starting at 14:00 UTC that turns counterclockwise, i.e., turns to east (E), after 21:00 UTC (Fig. 8f), showing a possible influence of the synoptic forcing in the modeling process of the ABL and hiding the real effect of the sea-breeze from the Atlantic Ocean.

The observed daily cycles for the potential temperature and water–vapor mixing ratio also represented qualitatively well by the simulation for both runs at the Ecologia Agricola station (Fig. 9a–d). The time cross-section of the wind direction computed from the CTL and HR runs agree qualitatively well with the observational data, except between 00:00–06:00 and 16:00–21:00 UTC on 7 September 2007 (Fig. 9e and f). Moreover, the computed results for the HR run present better agreement with the observational data than the CTL run for wind direction and speed, although the results show a systematic trend to overestimate the wind speed (Fig. 9g and h). The Ecologia Agricola station is positioned in a direction transverse to the coastline of the Sepetiba Bay, but it is relatively far from it. Thus, the effect of the sea-breeze sometimes drives the wind direction, as probably occurred at 16:00–21:00 UTC on 7 September 2007 (Fig. 9f), although this may not be always true, as the results show for the same time period on 6 September 2007 (Fig. 9e). Apparently, the ARPS results are not representing adequately the first situation. We believe that the poor representation of the 30s-USGS soil-type can greatly influence the results of the simulation, in spite of the best representation of the topography and vegetation-type provided in the HR run.

The numerical results for SBSC station present good agreement with the observational data in terms of the potential temperature, mainly on the second day of the simulation (Fig. 10a and b). Despite the fact that the daily cycle of the water–vapor mixing ratio is qualitatively well simulated, non-negligible differences are founded when simulated values are compared with the observational data (Fig. 10c and d).
This may indicate that the soil-moisture initialization does not possess an adequate resolution, even considering that there is a difference of about 8 m between the level of observation and the first grid level of the simulations. With respect to the wind direction, the northeast (NE) winds observed are well computed from the beginning of the simulation until 04:00 UTC of the first day, but after that the numerical data present differences of 45°, approximately, until 18:00 UTC (Fig. 10e). The simulations seem to detect when the sea-breeze wind start to turn coming from the Sepetiba Bay direction (19:00–20:00 UTC), but it fails to detect when the wind turns completely from the southwest (SW) direction at 21:00 UTC (Fig. 10e). On the second day there is a significant discrepancy between the simulations and the observed wind direction, mainly starting 15:00 UTC when SW sea-breeze winds from Sepetiba Bay dominate (Fig. 10f). These results indicate that the fine settings proposed for the databases were not appropriate to represent the wind direction behavior at the SBSC station. In this case, there are combined effects due the synoptic scale and the sea-breeze coming from the Atlantic Ocean and the Sepetiba Bay, despite the fact that the computed wind-speed daily cycle have performed well for both runs (Fig. 10g and h). The low resolution of the soil-type database associated to soil temperature and moisture initialization data may be affecting the modeled results in this region too.

The daily cycles for the potential temperature and water–vapor mixing ratio computed at the SBAF station also presented successful results, as observed for other stations (Fig. 11a–d). In general, the calculated wind speeds are lower than the measured data between 00:00 and 15:00 UTC on 6 and 7 September, probably due to the SBAF location, where the wind direction can be driven by a sum of two factors: the catabatic winds from Pedra Branca and Tijuca Massifs; and the land-breeze (Fig. 11e–h). At this time interval we note that the wind blows from E (Fig. 11e) on the first day, whereas it blows from N (Fig. 11f) on the second day. Between 18:00 and 21:00 UTC on 6 September, the flow accelerates slightly when the wind blows from S and SE, especially in the valley formed between the Pedra Branca and Tijuca Massifs (Fig. 11g). The same is observed between 15:00 and 21:00 UTC on 7 September (Fig. 11h),
suggesting the occurrence of a canalized jet due the sea-breeze effect. The HR run appears to overcome the CTL run when we compare the wind direction and simulated speed data to observational data of 7 September.

At the SBJR station no observational data were recorded between 22:00 and 09:00 UTC in the days analyzed (Fig. 12a–h). However, good agreement is observed between the observational and the simulated data and there are few discrepancies between the CTL and the HR runs, showing that the land-use database did not change too much the characteristics of the SBJR region and vicinity. It is worth mentioning that the maximum values of wind speed occurs around 18:00 UTC, when the wind blew from S and SE under the influence of sea-breeze from Atlantic Ocean as observed at SBAF (Fig. 12e–h).

Satisfactory results for daily cycles of the potential temperature can be seen on Fig. 13a and b for SBRJ station. Small discrepancy between the CTL and the HR runs are detected when compared to the observed values of the water–vapor mixing ratio, as also noticed at the others stations, probably due the difference between the observed level and the first grid point (Fig. 13c and d). The occurrence of calm winds, which reach 6 m s$^{-1}$ for most of the afternoon, is the main cause of the discrepancies found in the wind direction (Fig. 13e–h). At SBRJ, the true winds blows from N in the land-breeze period (00:00–14:00 UTC) and turns from S, in the Guanabara Bay entrance, in the sea-breeze period (15:00 UTC onwards) for both days. On the second day, the simulations present good results for the wind direction and speed during the sea-breeze period.

Good agreement is noted between the simulated and observational data for all variables plotted on the time cross-section diagrams for the SBGL station (Fig. 14a–h). The results present a few differences between the CTL and the HR runs. We highlight the occurrence of calm winds between 00:00 and 12:00 UTC and moderate S and SE winds in the afternoon and night (between 17:00 and 22:00 UTC), just when the sea-breeze flow is completely developed (Fig. 14e–h). The wind direction values computed by the HR run are better than the CTL run, mainly between 17:00 and
22:00 UTC for both days. At other times the discrepancy is larger due to the calm winds simulated, since the turbulence parameterization schemes typically show considerable difficulties in modeling the flow near the ground under conditions of calm, intermittent flow and normally when mesoscale models are downscaling to LES domains. However, it is important to point out that the results obtained from the simulations reproduce qualitatively the observed maximum wind speed behavior in the afternoon and the minimum wind speed in the morning.

The upper-air or vertical profiles of potential temperature, water–vapor mixing ratio and wind direction and speed computed for the SBGL upper-air sounding station show encouraging results when compared to the observed data collected at the Galeão Airport (see Figs. 15a–d and 16a–d) in the first 6 km a.g.l. at 12:00 UTC for both days. Despite the trend of the model to underestimate the observational data, analysis based on the potential temperature profile at 12:00 UTC on 6 September (Fig. 15a) shows good agreement between the observational data and the CTL and HR results, since a stable layer predominates in the whole atmosphere. The water–vapor ratio distribution calculated by ARPS underestimates the observed data up to 1600 m a.g.l., maybe due to the initialization used, but there is good agreement above this height (Fig. 15b). In the ABL region the variation of the wind direction with height is not well represented by the ARPS results (Fig. 15c). The ARPS results do not reproduce adequately the height AGL where the wind speed maximum occurs and the CTL run overestimates more than the HR run the wind speed value up to 100 m a.g.l. (Fig. 15d). From 100 m up to 800 m a.g.l. the CTL run perform better than the HR run. The comparison of simulated meteorological variables with observational data on the second day present a few differences between the CTL and the HR runs and overall better agreement between the runs and the observational data (Fig. 16a–d). In general, the ARPS results reproduce correctly the physical behavior of the atmosphere when compared to the observed data at the SBGL station, mainly with respect to the atmospheric stability, the wind shear, and the occurrence of low wind speed near the ABL and higher wind speed above 2000 m.
Analogously to Chow et al. (2006), Table 5 illustrates the mean errors (bias) and the Root-Mean-Square Errors (RMSE) computed for the potential temperature $\theta$, water–vapor mixing ratio $q_v$ and wind direction and speed for both runs of the second day. The results completely exclude the fast spin-up time of the indices and maintain only the computationally-stable time results of a one-day cycle. The bias and RMSE are computed as follows

$$\text{bias} = \frac{1}{N} \sum_{i=1}^{N} \phi'_i,$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\phi'_i)^2},$$

where $\phi' = \phi_f - \phi_o$ represents the difference, or deviation, between any forecasted and observed variable, and $N$ is the total number of verifications. For wind direction, a positive deviation means that the simulated wind vector deviates clockwise in relation to observed wind vector. Because the largest possible error in wind direction is 180 arc-deg, the definition of the deviation $\phi'$ needs to be changed according to

$$\phi' = (\phi_f - \phi_o) \left(1 - \frac{360}{|\phi_f - \phi_o|}\right), \text{ if } |\phi_f - \phi_o| > 180 \text{ arc-deg.}$$

In terms of the potential temperature, the HR run presents incontestable better results than the CTL run for the SBJR, SBRJ and SBGL surface stations. However, even for the other stations where the bias for $\theta$ for the CTL run is less than for the HR run, only half of these stations, i.e., Marambaia and SBAF station, presents RMSE greater for the HR run. Overall the ARPS results for $\theta$ tend to underestimate the observational data. It is important to highlight that, except for SBAF, we see slight discrepancies between forecasted and observed potential temperature for both runs. This behavior may be associated mainly with the detailed updated information on the vegetation-type.
used for the HR run, as discussed by Chow et al. (2006) for their Riviera Valley results, where a new set of 30 vegetation and 14 soil types were incorporated into ARPS. In relation to the water–vapor mixing ratio $q_v$ we observe similar differences between the CTL and the HR runs. The relatively high values of the bias confirm the discrepancies observed in Figs. 8–14, indicating that the ARPS results for $q_v$ tend to overestimate the observational data. However, we recall that, in this case, the comparison is made between the values computed from the first level of the G5 and the G6 grids with the data observed from the surface level of the weather stations.

The wind direction is probably the most complicated variable to obtain an accurate forecast. Also, we can often interpret erroneously the statistics indexes, for example, when bias is calculated for curves which remain lagged every moment. For the SBSC station, the bias of wind direction indicates very low values, but RMSE confirms that errors are as large as we have seen on the curves of Fig. 10f. At the SBJR the differences indicated by the RMSE values are small too, although the statistical indices were not calculated between 00:00 and 09:00 UTC, i.e., night to dawn and early morning. In general the HR run results appear to behave better than the CTL run results at most stations analyzed. Specifically, good performance is obtained for both runs on Marambaia station, where the sea-breeze effects seem more evident due to its proximity to the ocean. At SBRJ and SBGL the CTL run results outperform slightly the HR ones for the wind speed only. At the SBAF and Marambaia stations the simulations present very satisfactory results if we consider the low RMSE values. On the other hand, the worst bias values for the wind speed obtained with both runs occurred at the Ecologia Agricola station, where ARPS overestimates the wind speed excessively at almost all times, as shown in Fig. 9h. We must emphasize that the data at the observation level closest to the ground may be influenced by surface effects due to the plant canopy, which is not totally represented in mesoscale models due to difficulties with the turbulence parameterization schemes, incipient grid resolution, poor urban soil-type database, and low resolution of the topography database outside of the fine-resolution domains. These effects imply that the model is unable to provide accurate
lateral BC forcing for the fine grids, as discussed by Gohm et al. (2004). Despite these issues, our results indicate that the scale modeled in the most resolute grids of this work, based on “Terra Incognita” (Wyngaard, 2004), is still accurate enough to provide encouraging results.

4.2 Potential temperature and wind vector fields

Overall, our simulations indicate that the meteorological variables have similar behavior on both days analyzed, presenting local circulation patterns well defined by the differential heating created from sea–land and valley–mountain contrasts. The results for the potential temperature gradient (contoured lines and shaded areas) and the wind field (vectors) at 10 m a.g.l. at 14:00 UTC on 7 September 2007 of the G5 domain for the CTL and the HR runs are illustrated in Fig. 17a and b, respectively. Around this time (11:00 Local Time – LT) the onset of the sea-breeze front from the Atlantic Ocean occurs, just when the wind turns from SE near the coastline between −43.60 and −43.40 arc-deg west-longitude, in the vicinity of the Marambaia station. This effect occurs due to the pronounced sea–land temperature gradient computed for both runs in this region. Both runs indicate that there is a remarkable competition between the southeast winds from the sea-breeze and the northern winds from the land-breeze into the continent, but apparently with a major propagation of the sea-breeze front into the continent for the HR run (Fig. 17a and b). This behavior shows the importance of increasing the density of surface-weather observation stations at MARJ in order to evaluate whether the physical trend captured with the high-resolution ARPS modeling is in agreement with the sea breeze pattern observed. It is also important to highlight the occurrence of a high gradient of potential temperature in the northwest side of the G5 domain obtained from the CTL run (Fig. 17b) next to a water reservoir. Such high gradient may be influenced by the low resolution of the land-use database used in the CTL run (Figs. 3a and 4a), which maximizes erroneously the water body, unlike the HR run (Figs. 3b and 4b).
4.3 Vertical-latitudinal cross-section analysis

Figures 18a–d and 19a–d illustrate the vertical-latitudinal cross-section distribution of TKE, potential temperature and meridional-vertical wind vector components, computed at −43.60 arc-deg west-longitude (Marambaia longitude location) in the period 14:00–17:00 UTC on 7 September 2007 for the CTL and the HR runs. The calculations were performed in this period because we are interested in highlighting the TKE production close to the sea-breeze front indicated by the meteogram analysis from the Marambaia station (Fig. 8). One can see that the TKE production in the north part of the continental region (between −22.90 and −22.65 arc-deg south-latitude) may be associated, at all the times, to the near-surface vertical shear and buoyancy effects caused mainly by the mountain waves (see Fig. 1 for location), as also suggested by the numerical results for the wind velocity vector and vertical potential temperature gradient. In this region of the cross-section, the TKE intensity presents higher values at 15:00 and 16:00 UTC, mainly for the HR run (see Figs. 18b and c and 19b and c) due to the higher temperatures and wind shear computed. At 14:00 UTC (Figs. 18a and 19a), a stable stratified ABL occurs over the ocean region and near the coastline (between −23.11 and −23.00 arc-deg south-latitude) for both runs, and one can see a northerly wind blowing along the vertical-latitudinal cross-section beyond 200 ma.g.l. Below this level, on the surface layer, there is evidence that a southerly wind component starts the sea-breeze flow for both runs, which is associated with a weak horizontal gradient and a light TKE production at the sea-breeze front near the coastline for the CTL run and a more intense production for the HR run.

At 15:00 UTC (Figs. 18b and 19b) the horizontal gradient of temperature increases and the sea-breeze front reaches −22.96 arc-deg south-latitude, approximately, for both runs, characterized by a very near-surface southerly wind component which turns to the north direction about 150 ma.g.l. However, the wind is still weak and some vertical motion is generated, even near the horizontal limit of the sea-breeze flow. Into a shallow layer, the gradient of TKE increase more in the HR run than in the CTL run,
leaving a TKE trail in the onshore region. From 16:00 to 17:00 UTC (Figs. 18c and d and 19c and d) we can see that air rises over the warm land near the shoreline and into the continent (near −22.95 arc-deg south-latitude), and cooler air from the water is advected by the southerly wind component to replace it. The noticeable temperature drop observed in Figs. 18c and 19c is typical of a mesoscale cold front, as observed by Bastin and Drobinski (2006). The CTL run predicts that the low-level convergence occurs further into the continent (around 15 km from the shore) than the HR run (around 10 km from the shore) at 16:00 UTC (Figs. 18c and 19c). Also, the CTL run presents a wind speed more intense when compared to the HR run and the observed data at the Marambaia station (see also Fig. 8f). In both runs a return circulation (the anti-sea-breeze) about 5 m s$^{-1}$ brings warmer air back to the sea, which descends towards the sea surface and closes the circulation circuit. We can also note that, at 14:00 and 15:00 UTC (Figs. 18a and b and 19a and b), the speed of the following sea breeze is a little faster than the speed of the sea-breeze front, and a light wave propagates upstream of the sea-breeze front as the southerly flow collides with the adverse northerly flow. When the following sea-breeze reaches the front, convergence results in a “head-shaped” updraft, as also seen by Reible et al. (1993) and Bastin and Drobinski (2006). This “head” is a zone of intense mixing, which is supported by the significant values of the TKE shown in Figs. 18b and 19a and b. We observe that, in both runs, the highest TKE values occur in the mesoscale cold-front region around 450 m a.g.l. and near the ground-surface for the HR run. The vertical motion of the air in the CTL run is much weaker than in the HR run at 14:00–15:00 UTC (see Figs. 18a and b and 19a and b). The maximum depth of the sea-breeze is observed to be around 300 m a.g.l. for both runs at all times. After 17:00 UTC the magnitude of the updraft decreases quickly, as well as its vertical extent, and the TKE intensity decreases slowly with time near the surface. Also, the northerly component of the upper-level wind strengthens, whereas the TKE decreases and remains confined near the surface, where some shear appears.
5 Summary and conclusions

Numerical simulations of the ABL flow are strongly influenced by several factors, namely the parametric models adopted in the boundary value problem that represents the physical situation, the numerical methods applied to solve the conservation equations, the numerical-grid scheme, and the boundary conditions related to the synoptic forcing and surface databases. In order to reduce the influence of these factors, we incorporate into ARPS the 3s-SRTM topographic database, the 10s-ESA vegetation-type database and the 30s-ESA LAI and FAPAR databases, which are preprocessed by subroutines we developed for the ARPS architecture. The numerical simulations are carried out in the LES mode for a period of 48 physical hours, running on six one-way nested-grids that are setup separately, such that different vertical resolutions and parameterizations are chosen for each scale being modeled.

We observed satisfactory agreement between numerical and field data for some periods of the days we investigated, mainly with the high-resolution run proposed in this work. Analogously to Chen et al. (2004), our simulations also showed that the increased resolution by itself leads to better numerical results. Overall, the slight discrepancies between the observed data and the simulated potential temperature results obtained for the HR run presents significantly lower errors than for the CTL run. On the other hand, the discrepancies are higher between the CTL and the HR runs for the water–vapor mixing ratio results. The calculated bias and RMSE values in the wind speed results for the CTL run are better than for the HR run at the SBRJ and SBGL stations. The HR run appears to overcome the CTL run when we compare the simulated data for the wind direction and speed to observational data on 7 September. At the SBAF station, for example, both runs forecasted well the canalized jet triggered by the sea-breeze effect. On the other hand, at the SBGL station, the wind direction sometimes presents more discrepancies due the calm winds that appear in the simulation. As Hanna and Yang (2001) pointed out, the turbulence parameterization schemes typically have difficulties in modeling the flow
near the ground under conditions of calm, intermittent flow and when mesoscale models are downscaling to LES domains. Equivalent statistics results and conclusions were obtained also by Chow et al. (2006), indicating the difficulties in modeling the wind field. Some clear discrepancies were observed, mainly at the moments when a flow transition occurs to form the sea-breeze. However, the remarkable TKE production on the sea-breeze front shows a pattern very similar to the one found by Bastin and Drobinski (2006), which is evidence of a consistent physical behavior. In accordance with the literature, our results indicate that an improved representation of the properties and characteristics of the land-use can dramatically influence the calculation of the momentum, heat and moisture fluxes between the surface and the atmosphere, and they may significantly affect the calculation of the meteorological field quantities. This suggests a need for improving our soil-type databases, soil moisture and temperature initialization.

Acknowledgements. The authors acknowledge the collaboration of: Fotini Katapodes Chow, from the University of California at Berkeley for your precious suggestions; Vasileios Kalogirou, from ESA/ESRIN, for his invaluable support with ESA database; Kevin W. Thomas, Ming Xue and Yunheng Wang, from the Center for Analysis Prediction of Storms, for their technical support with ARPS. The authors would also like to acknowledge the Brazilian National Institute of Meteorology and Meteorology Network of the Brazilian Air Force Command for leaving available the observational surface data, in addition to CNPq, CAPES and FAPERJ for their financial support.

References

Influence of high-resolution surface databases

L. M. S. Paiva et al.


Influence of high-resolution surface databases

L. M. S. Paiva et al.


Radnóti, G., Ajjaji, R., Bubnová, R., Caian, M., Cordoneanu, E., Von Der Emde, K., Gril, J. D., Hoffman, J., Horányi, A., Issara, S., Ivanovici, V., Janousek, M., Joly, A., Le Moigne, P., and


Table 1. Location of the surface and \textsuperscript{a}upper-air weather observation stations.

<table>
<thead>
<tr>
<th>Weather stations</th>
<th>Synop code</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (AMSL\textsuperscript{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marambaia</td>
<td>–</td>
<td>23.05° S</td>
<td>43.60° W</td>
<td>09.70 m</td>
</tr>
<tr>
<td>Ecologia Agricola</td>
<td>–</td>
<td>22.80° S</td>
<td>43.68° W</td>
<td>34.00 m</td>
</tr>
<tr>
<td>SBSC</td>
<td>83741</td>
<td>22.93° S</td>
<td>43.72° W</td>
<td>03.00 m</td>
</tr>
<tr>
<td>SBAF</td>
<td>83748</td>
<td>22.87° S</td>
<td>3.38° W</td>
<td>34.00 m</td>
</tr>
<tr>
<td>SBJR</td>
<td>83111</td>
<td>22.97° S</td>
<td>43.37° W</td>
<td>03.00 m</td>
</tr>
<tr>
<td>SBRJ</td>
<td>83755</td>
<td>22.88° S</td>
<td>43.17° W</td>
<td>03.00 m</td>
</tr>
<tr>
<td>SBGL\textsuperscript{a}</td>
<td>83746</td>
<td>22.82° S</td>
<td>43.25° W</td>
<td>06.00 m</td>
</tr>
</tbody>
</table>

\textsuperscript{a} AMSL: above mean sea level.

Source: Brazilian National Institute of Meteorology and Meteorology Network of the Brazilian Air Force Command.
Table 2. Structure of the one-way nested numerical grids. Given the number of grid points $n_x$, $n_y$ and $n_z$, the physical domain size can be calculated as: $L_x = (n_x - 3) \Delta x \times L_y = (n_y - 3) \Delta y \times H_z = (n_z - 3) \Delta z_{\text{med}}$.

<table>
<thead>
<tr>
<th>Grids</th>
<th>$n_{zg}$</th>
<th>$n_x \times n_y \times n_z$</th>
<th>$\Delta x$, $\Delta y$</th>
<th>$\Delta z_{\text{min}}$, $\Delta z_{\text{med}}$</th>
<th>$L_x \times L_y \times H_z$</th>
<th>$\Delta t / \Delta \tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEXT</td>
<td>2</td>
<td>$39 \times 39 \times 53$</td>
<td>55 km, 50 m</td>
<td>500 m</td>
<td>$1980 \text{km} \times 1980 \text{km} \times 25 \text{km}$</td>
<td>–</td>
</tr>
<tr>
<td>G1</td>
<td>2</td>
<td>$51 \times 51 \times 53$</td>
<td>27 km, 50 m</td>
<td>500 m</td>
<td>$1296 \text{km} \times 1296 \text{km} \times 25 \text{km}$</td>
<td>15 s/15 s</td>
</tr>
<tr>
<td>G2</td>
<td>2</td>
<td>$95 \times 78 \times 53$</td>
<td>9 km, 50 m</td>
<td>500 m</td>
<td>$828 \text{km} \times 675 \text{km} \times 25 \text{km}$</td>
<td>8 s/8 s</td>
</tr>
<tr>
<td>G3</td>
<td>2</td>
<td>$99 \times 78 \times 53$</td>
<td>3 km, 50 m</td>
<td>500 m</td>
<td>$288 \text{km} \times 225 \text{km} \times 25 \text{km}$</td>
<td>2.5 s/2.5 s</td>
</tr>
<tr>
<td>G4</td>
<td>2</td>
<td>$171 \times 111 \times 63$</td>
<td>1 km, 50 m</td>
<td>400 m</td>
<td>$168 \text{km} \times 108 \text{km} \times 24 \text{km}$</td>
<td>2 s/1.5 s</td>
</tr>
<tr>
<td>G5</td>
<td>2</td>
<td>$227 \times 177 \times 63$</td>
<td>$300 \text{m}$, 20 m</td>
<td>$350 \text{m}$</td>
<td>$67.2 \text{km} \times 52.2 \text{km} \times 21 \text{km}$</td>
<td>1 s/0.2 s</td>
</tr>
<tr>
<td>G6</td>
<td>2</td>
<td>$187 \times 187 \times 63$</td>
<td>$300 \text{m}$, 20 m</td>
<td>$350 \text{m}$</td>
<td>$55.2 \text{km} \times 55.2 \text{km} \times 21 \text{km}$</td>
<td>1 s/0.2 s</td>
</tr>
</tbody>
</table>
Table 3. Conversion of the original 30s-USGS and updated 10s-ESA categories of vegetation-type to the USDA classification that is adopted by ARPS. See tables in Xue et al. (1995) and Bicheron et al. (2008) user guide to know what each value corresponds specifically in their own vegetation-type categories.

<table>
<thead>
<tr>
<th>ARPS</th>
<th>Definition</th>
<th>Categories</th>
<th>$z_0$ (m)</th>
<th>USGS</th>
<th>ESA</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Desert</td>
<td>Bare Grnd. Tundra</td>
<td>0.011</td>
<td>23</td>
<td>200, 201, 202, 203</td>
</tr>
<tr>
<td>02</td>
<td>Tundra</td>
<td>Herb. Tundra, Wooden Tundra Mixed Tundra</td>
<td>0.076</td>
<td>20, 21, 22</td>
<td>145</td>
</tr>
<tr>
<td>03</td>
<td>Grassland</td>
<td>Crop./Grs. Mosaic, Grassland</td>
<td>0.030</td>
<td>5, 7</td>
<td>140, 141, 143</td>
</tr>
<tr>
<td>04</td>
<td>Grassland with shrub cover</td>
<td>Mix Shrb./Grs. Savana</td>
<td>0.238</td>
<td>9, 10</td>
<td>142, 130, 131, 132, 134, 135, 136</td>
</tr>
<tr>
<td>05</td>
<td>Grassland with tree cover</td>
<td>Crop./Wood Mosc.</td>
<td>0.563</td>
<td>6</td>
<td>110, 120, 144, 30, 32</td>
</tr>
<tr>
<td>06</td>
<td>Deciduous forest</td>
<td>Decids. Broadlf., Decids. Needlf, Mixed Forest</td>
<td>0.826</td>
<td>11, 12, 15</td>
<td>50, 60, 80, 90, 91, 100, 101, 102, 16</td>
</tr>
<tr>
<td>07</td>
<td>Evergreen forest</td>
<td>Evergrn. Braodlf, Evergrn. Needlf.</td>
<td>1.089</td>
<td>13, 14</td>
<td>41, 42, 70, 92</td>
</tr>
<tr>
<td>08</td>
<td>Rain forest</td>
<td></td>
<td>2.653</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>09</td>
<td>Ice</td>
<td>Snow or Ice</td>
<td>0.011</td>
<td>24</td>
<td>220</td>
</tr>
<tr>
<td>10</td>
<td>Cultivation</td>
<td>DryInd Crop. Past., Irrg. Crop. Past., Mix. Dry/Irrg. C.P.</td>
<td>0.075</td>
<td>2, 3, 4</td>
<td>10, 14, 15, 20, 21, 22, 31</td>
</tr>
<tr>
<td>11</td>
<td>Bog or marsh</td>
<td>Herb. Wetland, Wooded wetland</td>
<td>0.100</td>
<td>17, 18</td>
<td>11, 12, 13, 160, 161, 162, 170, 180, 181, 182, 183, 184, 185, 186, 187, 188</td>
</tr>
<tr>
<td>12</td>
<td>Dwarf shrub</td>
<td>Shrubland</td>
<td>0.853</td>
<td>8</td>
<td>133</td>
</tr>
<tr>
<td>13</td>
<td>Semidesert</td>
<td>Urban Bar. Sparse Veg.</td>
<td>1.500</td>
<td>1, 19</td>
<td>150, 151, 152, 153, 190</td>
</tr>
<tr>
<td>14</td>
<td>Water</td>
<td>Water Bodies</td>
<td>0.002</td>
<td>16</td>
<td>210</td>
</tr>
</tbody>
</table>
Table 4. Summarized surface database configuration of the one-way nested numerical grids.

<table>
<thead>
<tr>
<th>Grids</th>
<th>Topography</th>
<th>Soil-type</th>
<th>Vegetation-type</th>
<th>LAI</th>
<th>Vegetation Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1–G3</td>
<td>30s-USGS</td>
<td>30s-USGS</td>
<td>30s-USGS</td>
<td>30s-USGS</td>
<td>30s-NESDIS</td>
</tr>
<tr>
<td>G4–G6</td>
<td>CTL 30s-USGS</td>
<td>HR 3s-SRTM</td>
<td>30s-USGS 10s-ESA</td>
<td>30s-USGS</td>
<td>30s-NESDIS</td>
</tr>
</tbody>
</table>

L. M. S. Paiva et al.
Table 5. Mean errors (bias) and Root-Mean-Square Errors (RMSE) for potential temperature $\theta$, water–vapor mixing ratio $q_v$, wind direction and speed.

<table>
<thead>
<tr>
<th>Weather stations and runs</th>
<th>Variables and statistics</th>
<th>$\theta$ (K)</th>
<th>$q_v$ (gkg$^{-1}$)</th>
<th>Direction (arc-deg)</th>
<th>Speed (ms$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>bias RMSE</td>
<td>bias RMSE</td>
<td>bias RMSE</td>
<td>bias RMSE</td>
<td>bias RMSE</td>
</tr>
<tr>
<td>Marambaia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-0.13 1.55</td>
<td>-2.81 2.82</td>
<td>-16.84 61.44</td>
<td>2.18 2.72</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-0.20 1.56</td>
<td>-2.80 2.81</td>
<td>-15.65 53.96</td>
<td>1.15 1.93</td>
<td></td>
</tr>
<tr>
<td>Ecologia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricola</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>0.80 3.19</td>
<td>-3.62 3.66</td>
<td>-44.49 81.04</td>
<td>3.31 4.40</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>1.09 2.84</td>
<td>-3.63 3.68</td>
<td>-46.52 83.79</td>
<td>2.83 3.54</td>
<td></td>
</tr>
<tr>
<td>SBSC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-0.01 2.39</td>
<td>-3.87 3.89</td>
<td>7.52 110.14</td>
<td>1.39 2.80</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>0.26 2.10</td>
<td>-3.89 3.91</td>
<td>0.13 113.98</td>
<td>1.34 2.44</td>
<td></td>
</tr>
<tr>
<td>SBAF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-3.61 4.14</td>
<td>-4.01 4.03</td>
<td>-16.60 99.55</td>
<td>1.19 2.11</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-4.27 5.02</td>
<td>-3.90 3.93</td>
<td>-31.38 71.86</td>
<td>0.26 1.72</td>
<td></td>
</tr>
<tr>
<td>SBJR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-1.88 2.66</td>
<td>-4.02 4.09</td>
<td>-11.32 62.09</td>
<td>1.90 2.34</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-1.66 2.37</td>
<td>-3.99 4.06</td>
<td>-7.73 37.45</td>
<td>1.68 2.26</td>
<td></td>
</tr>
<tr>
<td>SBRJ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-2.23 2.59</td>
<td>-2.96 2.98</td>
<td>48.75 85.59</td>
<td>-0.07 1.16</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-1.94 2.27</td>
<td>-2.97 2.99</td>
<td>-20.87 88.73</td>
<td>0.22 1.56</td>
<td></td>
</tr>
<tr>
<td>SBGL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTL</td>
<td>-2.03 2.87</td>
<td>-2.91 2.93</td>
<td>41.92 104.32</td>
<td>0.50 1.54</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-1.80 2.69</td>
<td>-2.99 3.01</td>
<td>24.50 101.17</td>
<td>0.55 1.64</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1. The G5 and G6 innermost grids nesting to G4 and the terrain map of the Metropolitan Area of Rio de Janeiro (MARJ) showing the location of the surface weather observation stations.
Fig. 2. Shaded-topographic elevation maps for G5, processed with (a) 30s-USGS and (b) 3s-STRM databases.
Fig. 3. Soil-type shaded maps for G5, processed with 30s-USGS database (a) not-adjusted and (b) adjusted.
Fig. 4. Vegetation-type shaded maps for G5, processed with (a) 30s-USGS and (b) 10s-ESA databases.
Fig. 5. Surface-roughness shaded maps for G5, processed with (a) 30s-USGS and (b) 10s-ESA databases.
Fig. 6. Leaf-Area-Index (LAI) shaded maps for G5, processed with (a) 30s-USGS and (b) 30s-ESA databases.
Fig. 7. Vegetation cover fraction shaded maps for G5, processed with (a) 30s-USGS and (b) 30s-ESA databases.
Fig. 8. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G5 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at Marambaia surface weather station.
Fig. 9. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G5 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at Ecologia Agricola surface weather station.
Fig. 10. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G5 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at SBSC surface weather station.
Fig. 11. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at SBAF surface weather station.
Fig. 12. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at SBJR surface weather station.
Fig. 13. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at SBRJ surface weather station.
Fig. 14. Time cross-section data of (a, b) potential temperature $\theta$ (K), (c, d) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (e, f) wind direction (arc-deg) and (g, h) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 00:00 UTC on 6 September to 00:00 UTC on 8 September 2007 at SBGL surface weather station.
Fig. 15. Vertical data profiles of (a) potential temperature $\theta$ (K), (b) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (c) wind direction (arc-deg) and (d) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 12:00 UTC on 6 September 2007 at SBGL upper-air weather station.
Fig. 16. Vertical data profiles of (a) potential temperature $\theta$ (K), (b) water–vapor mixing ratio $q_v$ (g kg$^{-1}$), (c) wind direction (arc-deg) and (d) speed (m s$^{-1}$) observed and simulated by G6 of CTL and HR runs for 12:00 UTC on 7 September 2007 at SBGL upper-air weather station.
Fig. 17. Horizontal cross-section of potential temperature (K – contoured lines and shaded areas) and wind field (m s$^{-1}$ – vectors) at 10 m a.g.l. simulated by G5 of (a) CTL and (b) HR runs for 14:00 UTC on 7 September 2007.
Fig. 18. Vertical-latitudinal cross-section of Turbulent Kinetic Energy (TKE – m$^2$s$^{-2}$ – shaded), potential temperature (K – solid line) and meridional-vertical wind vector components (m$^{-1}$ – vectors) up to 1.5 km a.g.l. simulated by G5 of CTL run for (a) 14:00, (b) 15:00, (c) 16:00 and (d) 17:00 UTC on 7 September 2007, at −43.60 arc-deg west-longitude (Marambaia longitude location).
Fig. 19. Vertical-latitudinal cross-section of Turbulent Kinetic Energy (TKE – m$^2$s$^{-2}$ – shaded), potential temperature (K – solid line) and meridional-vertical wind vector components (m$^{-1}$ – vectors) up to 1.5 km a.g.l. simulated by G5 of HR run for (a) 14:00, (b) 15:00, (c) 16:00 and (d) 17:00 UTC on 7 September 2007, at −43.60 arc-deg west-longitude (Marambaia longitude location).