CrossMark

RESEARCH ARTICLE

Evaluation of Nocturnal Temperature Forecasts Provided by the Weather Research and Forecast Model for Different Stability Regimes and Terrain Characteristics

Adriano Battisti[1] (iD) · **Otávio C. Acevedo[1]** · **Felipe D. Costa[2]** ·
Franciano S. Puhales[1] · **Vagner Anabor[1]** · **Gervásio A. Degrazia[1]**

Abstract The quality of nocturnal temperature forecasts made by the Weather Research and Forecast (WRF) numerical model is evaluated. The model was run for all July 2012 nights, and temperature fields compared to hourly observations made at 26 weather stations in southern Brazil. Four different planetary boundary-layer (PBL) schemes are considered: Bougeault–Lacarrere (BouLac), Quasi-Normal Scale Elimination (QNSE), Yonsei University (YSU) and Mellor–Yamada–Janjic (MYJ). Additional simulations to assess the role of higher horizontal and vertical resolutions were performed using the MYJ scheme. All schemes, except BouLac, underestimated the 2-m temperature, and in all cases the temperature bias is dependent on wind speed. At high wind speeds, all schemes exhibit a cold bias, which is greater for those that yield lower nocturnal surface-layer turbulent intensity. The elevation difference between each station and the model nearest grid point $H_{station} - H_{gridpoint}$ is highly correlated with the temperature simulation error. We found that a consistent cold bias is restricted to conditions with low-level clouds. We concluded that one possible means of improving nocturnal temperature forecast is to use parametrization schemes that allow for higher turbulent intensity in near-neutral conditions. The results indicate that this improvement would partially counteract the misrepresentation of the low-level cloud cover. In most stable cases, a post-processing algorithm based on terrain characteristics may improve the forecasts.

Keywords Nocturnal temperatures · Stable boundary layer · Turbulence formulations · Weather Research and Forecasting model

✉ Adriano Battisti
adrianotbattisti@gmail.com

[1] Universidade Federal de Santa Maria, Santa Maria, RS, Brazil

[2] Universidade Federal do Pampa, Alegrete, RS, Brazil

1 Introduction

Processes such as frost, fog formation, and ground freezing as well as thermal comfort indices are directly affected by nocturnal near-surface air temperature, itself strongly reflecting the magnitude of the surface-atmosphere interaction. As a consequence, local surface features, such as terrain, vegetation, land-use, and proximity to obstacles may exert a significant control by enhancing the horizontal variability of near-surface air temperature, complicating its forecast (Gustavsson et al. 1998; Acevedo and Fitzjarrald 2003; Mahrt 2006; Bodine et al. 2009; Acevedo et al. 2013; Medeiros and Fitzjarrald 2014, 2015).

The degree of nocturnal temperature horizontal variability is itself highly dependent on the stable boundary-layer (SBL) flow. On the nights with a sufficient horizontal boundary-layer pressure gradient, or in concert with reduced radiative cooling associated with cloudiness, the weakly stable case, air near the surface is connected through turbulence to the upper SBL. For such conditions, the horizontal variability of near-surface temperature is largely reduced over areas of linear dimension from hundreds of metres (Bodine et al. 2009; Acevedo et al. 2013) to tens of kilometres (Acevedo and Fitzjarrald 2001). Conversely, on clear-sky nights with reduced large-scale forcing, the surface often decouples from the upper SBL. These nights are classified as very stable and, for such conditions, cold-air pooling is initiated, so that nocturnal air temperature varies significantly over small differences in altitude. Bodine et al. (2009) found a 5 K temperature difference over 300 m of horizontal separation with an altitude difference of 25 m. Acevedo et al. (2013) found that such a difference reached 10 K over similar horizontal and vertical distances. In the latter case, it is likely that the temperature difference was increased by the fact that the lower station was obstructed, further limiting local mixing and subsequent turbulent interaction with other regions of the SBL. Thus, on calm nights, although much of the nocturnal temperature variability may be attributed to topography, other local features such as the proximity to obstacles must also be considered. On such nights, the lower atmosphere over higher locations usually remains turbulent throughout the night, therefore becoming substantially warmer relative to the lower, decoupled locations (Acevedo and Fitzjarrald 2003; Medeiros and Fitzjarrald 2014, 2015). According to an objective classification of the nocturnal boundary layer based on the heat-flux dependency on stability (Mahrt 1999; Acevedo and Fitzjarrald 2003), weakly stable conditions occur when the absolute flux increases with stability because of the enhanced thermal gradient, while in the very stable case the enhanced stability dampens turbulence such that the heat-flux magnitude decreases with stability.

Forecasting nocturnal temperature is, therefore, a difficult task. The large horizontal variability of temperature over relatively small areas implies that, unless windy conditions prevail, no single temperature is representative over the area of a common grid cell of a mesoscale numerical model, which is usually of a few kilometres square. Furthermore, each model is sensitive to the boundary-layer turbulence parametrization scheme (Cuxart et al. 2006; Steeneveld et al. 2006; Svensson et al. 2011).

The parametrization schemes usually relate turbulence to atmospheric stability through a stability function, which is obtained from theoretical considerations and observations (Louis 1979; Delage 1997). One particularly important issue regarding stability functions refers to the turbulence they prescribe for strong stability. Basic SBL theory indicates that there is a threshold (related to the critical Richardson number) above which turbulence is effectively suppressed. However, when such a characteristic is implemented in a numerical weather model, it often leads to a process known as "runaway cooling" (Louis 1979; Steeneveld et al. 2006; Holtslag et al. 2013), for which the surface temperature decreases unrealistically

due to radiative loss not being counteracted by the turbulent transport of warmer air from above. In fact, the stability function proposed by Louis (1979) retains some turbulence in very stable conditions to avoid such a problem. The same idea of retaining finite mixing even for intense stratification has also been justified as a means of accounting for localized turbulent activity within the area of a model grid cell, as is usually observed at locations with relatively higher altitudes in a region (Mahrt 1987; Delage 1997). Medeiros and Fitzjarrald (2014, 2015) found that the model use of the high Richardson number threshold could be understood observationally to be a result of spatially averaging surface temperatures in regions of moderately complex topography.

Cuxart et al. (2006) compared single-column models using different stability functions and concluded that schemes used by each operational weather forecasting model tended to overestimate mixing in the SBL. Svensson et al. (2011) made a similar comparison, finding that all compared schemes provided excessive turbulence at night, except for those considered in the Weather Research and Forecasting (WRF) model, viz., the Yonsei University (YSU) and Mellor–Yamada–Janjic (MYJ) schemes.

In the WRF model (Skamarock et al. 2008), the planetary boundary-layer (PBL) parametrization schemes have been subject to different adjustments, many of which focus on the level of nocturnal mixing for very stable conditions. Hu et al. (2010) found that the YSU PBL scheme provided higher nighttime temperatures based on an update that increased nocturnal mixing in the referred scheme (implemented in version 3.0), while the MYJ parametrization scheme produced lower temperatures when compared to observations. Jiménez et al. (2012) found that wind-speed simulations improved with the use of different stability functions and of a lower friction velocity minimum. The proposed changes have been implemented in subsequent versions of the WRF model (from 3.1.1). An additional update to the YSU PBL parametrization scheme that reduced turbulence (implemented in version 3.4.1) has been tested by Hu et al. (2013) and found to reduce the simulated nocturnal temperatures, which then approached values simulated using the MYJ parametrization scheme. Kleczek et al. (2014) found that the WRF model tends to underestimate nocturnal temperatures, regardless of the PBL scheme used.

In the present study, it is hypothesized that, in addition to the difficulty in properly quantifying turbulence in the stable environment, a large fraction of the nocturnal temperature forecast errors arise from the naturally occurring large horizontal variability of meteorological variables in the period. The WRF model nocturnal temperature simulations are compared to observations over a network of 26 stations in southern Brazil, for 31 different nights. The results using four PBL parametrization schemes, Bougeault–Lacarrere (BouLac), Quasi-Normal Scale Elimination (QNSE), YSU and MYJ are compared. The MYJ and YSU schemes have been chosen because they are popular choices of local and non-local schemes in similar studies, having been chosen by Svensson et al. (2011) to represent the WRF model in their extensive comparison of numerical models and respective PBL schemes. The BouLac scheme is tested because it is a scheme that provides higher turbulence levels than the other schemes, so that its inclusion better allows association of the quality of the temperature forecasts with turbulence-related quantities. The QNSE scheme is considered because it has been recently proposed specifically to address the problems that arise in the SBL (Sukoriansky et al. 2005). The influence of horizontal and vertical grid resolutions is also considered, but only for the MYJ scheme.

The main novelty regards the fact that the range of stations and nights used in the comparison allow us to investigate not only the influence of the PBL schemes on temperature forecasts, but also how the schemes perform for locations with different terrain characteristics and for nights with different large-scale forcing. Therefore, we use stations located at lower

and higher altitudes relative to their surroundings. Furthermore, there are nights when the boundary layer at most of the stations is in a weakly stable state, while on the other nights, very stable conditions prevail.

We focus on the influence of terrain characteristics and atmospheric stability on nocturnal temperature forecast errors, and the results of an analysis on the effects of cloud cover on the temperature forecast are also presented. Influences that are not considered include those from land use, soil temperature and humidity, proximity to the coast or obstacles, occurrence of breezes. These aspects are not addressed for three main reasons: simplicity of the analysis; lack of observational data on many of these quantities at all stations, and the fact that it is shown that a large fraction of the errors are, in fact, explained by the factors included in the analysis.

2 Model Setup

We use the ARW-WRF model, version 3.6.1, with initial conditions provided by the Global Forecast System (GFS) model analysis. Boundary condition adjustments are provided every 6 h, also by the GFS analysis; all 31 nights of July 2012 have been simulated. A 15-h long spin-up time has been applied, so that the model simulation commences at 1200 UTC (0900 Local Standard Time, LST) of the previous day, being integrated until 0900 UTC. Hourly fields are extracted from 0300 UTC to 0900 UTC, resulting in seven values per station and per night for each variable. Three two-way nested domains are used (Fig. 1), where the larger covers longitudes 75.5°W−31.9°W and latitudes 11.3°S−48.1°S, with a 48-km resolution. The second grid covers 66.5°W−40.9°W and 41.0°S−18.4°S, with a 12-km resolution. In four of the model configurations, only the solutions from the second grid are considered. To address specifically the improvements that may arise from the use of a finer inner grid, several simulations with a higher horizontal resolution were also done. In these cases, a third grid covers longitudes 60.4°W−47.0°W and latitudes 35.3°S−24.1°S with a 4-km horizontal grid.

The Lambert projection is used with 28 vertical terrain-following hydrostatic-pressure levels (Skamarock et al. 2008; Wang et al. 2015) in the three domains. The lowest level is located at 28 m and the thickness of the subsequent layers increases with height until the domain top, located at 100 hPa, such that there are five levels in the lowest 500 m. Simulations with higher vertical resolution were done, and in such cases, there were 56 vertical levels, with the lowest near 10 m and a total of 18 levels in the lowest 500 m.

In the simulations with 28 vertical levels, the timestep for the coarser domain is 288 s, decreasing accordingly to the grid resolution at the inner domains, to 72 s at the second domain, and to 24 s at the finer grid. When the 56 vertical levels are used, all timesteps are correspondingly reduced by half. The topographic data used in the two coarser grids are from the United States Geological Survey (USGS) and have a horizontal resolution of 2 min, adjusted to the model grid points. For the finest grid, the terrain information is obtained from a topographic dataset with a 30-s horizontal resolution. Land-use, roughness and vegetation type have also been obtained from a 2-min resolution USGS database.

For the two coarser grids, the Lin microphysics scheme is used (Lin et al. 1983), while for the finer grid the WRF-model single-moment-3-class microphysics scheme is applied (Hong et al. 2004). The Noah land-surface scheme (Tewari et al. 2004) and the Kain–Fritsch cumulus physics scheme (Kain 2004) are used; for the longwave radiation, the Rapid Radiative
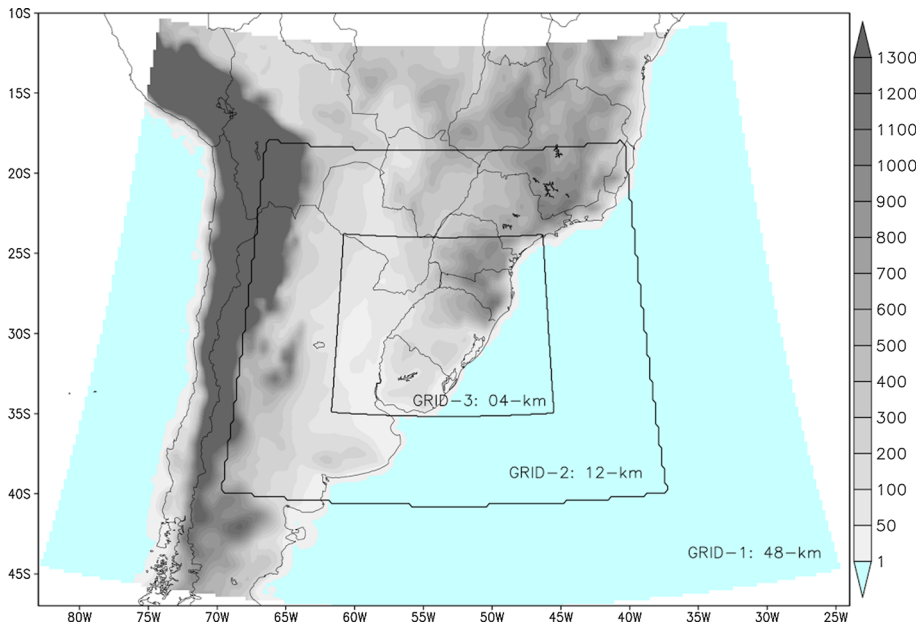
Fig. 1 Positions of the three domains used in the simulations with respect to the South American map

Table 1 Summary of all simulations performed

Simulation name	Boundary-layer parametrization scheme	Number of vertical points	Horizontal resolution (km^2)
BouLac	Bougeault–Lacarrere	28	12 × 12
QNSE	Quasi-Normal Scale Elimination	28	12 × 12
YSU	Yonsei University	28	12 × 12
MYJ	Mellor–Yamada–Janjic	28	12 × 12
MYJ-h	Mellor–Yamada–Janjic	56	12 × 12
MYJ-4	Mellor–Yamada–Janjic	28	4 × 4
MYJ-4h	Mellor–Yamada–Janjic	56	4 × 4

Transfer Model (RRTM) scheme is applied (Mlawer et al. 1997), and the Dudhia shortwave scheme is also used (Dudhia 1989).

Four PBL parametrization schemes are used for the two coarser grids: Bougeault–Lacarrere (BouLac; Bougeault and Lacarrere 1989), Quasi-Normal Scale Elimination (QNSE; Sukoriansky et al. 2005), Yonsei-University (YSU; Hong and Pan 1996; Hong et al. 2006) and Mellor–Yamada–Janjic (MYJ; Mellor and Yamada 1982; Janjic 1994, 2002). The simulations with higher horizontal and vertical grid resolutions are performed only using the MYJ scheme, chosen because this is the parametrization scheme that provides the lowest turbulence levels of all, so that it tends to enhance local effects. For this reason, it is also expected that the use of a finer grid in this scheme will have a greater influence than in the others. A summary of all simulations is presented in Table 1, while the main characteristics of the PBL parametrization schemes are described in the Appendix.
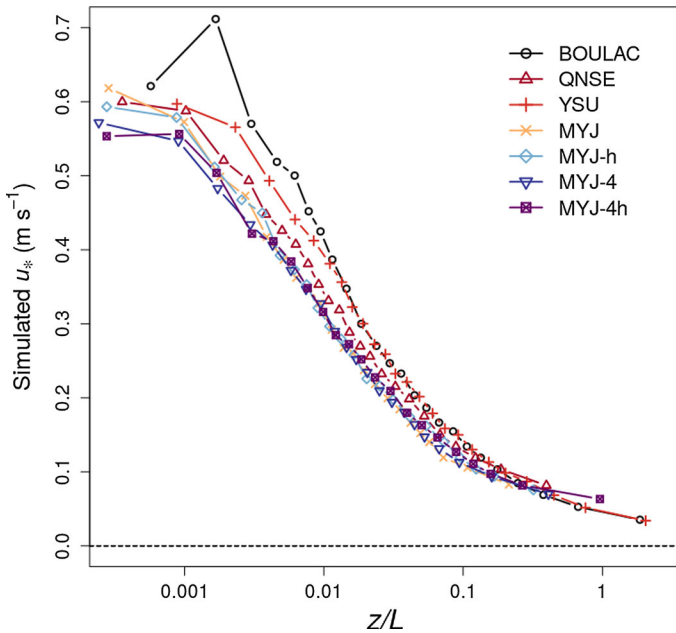
Fig. 2 Friction velocity u_* as a function of the stability parameter z/L. All values are obtained from the model fields. The *lines* represent the different PBL scheme, horizontal and vertical grids resolutions used (see Table 1). Results are block-averaged for all stations and nights being analyzed. Each *point* represents the average over 200 values

For similar values of the Obukhov stability parameter z/L (z is height, L is the Obukhov length), the MYJ scheme (regardless of the horizontal and vertical grid resolutions used) usually provides lower turbulence than the BouLac, QNSE and YSU schemes (Fig. 2). The exception is in the very stable limit (large z/L), when all schemes tend, on average, to provide similar values of friction velocity u_*. In most extreme cases of stability, the minimum values of u_* provided by the MYJ scheme are actually slightly greater than those provided by the other schemes (figure not shown), indicating that for the very stable limit the MYJ scheme retains greater turbulence levels than in the others. For $z/L > 0.02$, the BouLac and YSU schemes provide similar average values of u_*, a consequence of the use of the same surface-layer scheme in both parametrization schemes. In the near-neutral limit (small z/L), however, the BouLac scheme tends to provide higher turbulence levels than does the YSU scheme, while the QNSE scheme generally provides higher turbulence levels than the MYJ scheme, but lower levels than the BouLac and YSU schemes (for similar values of z/L).

3 Observations

The model fields are compared to observations collected at 26 stations from the Brazilian National Meteorology Service (Instituto Nacional de Meteorologia—INMET) during all 31 simulated nights. These are all stations in operation during that period in the southernmost Brazilian state, Rio Grande do Sul (Fig. 3). The stations represent a range of surface conditions, and include coastal stations, with those furthest inland about 550 km from the coast. Station altitudes vary from sea level to 1230 m (Online Resource 1), with the highest alti-
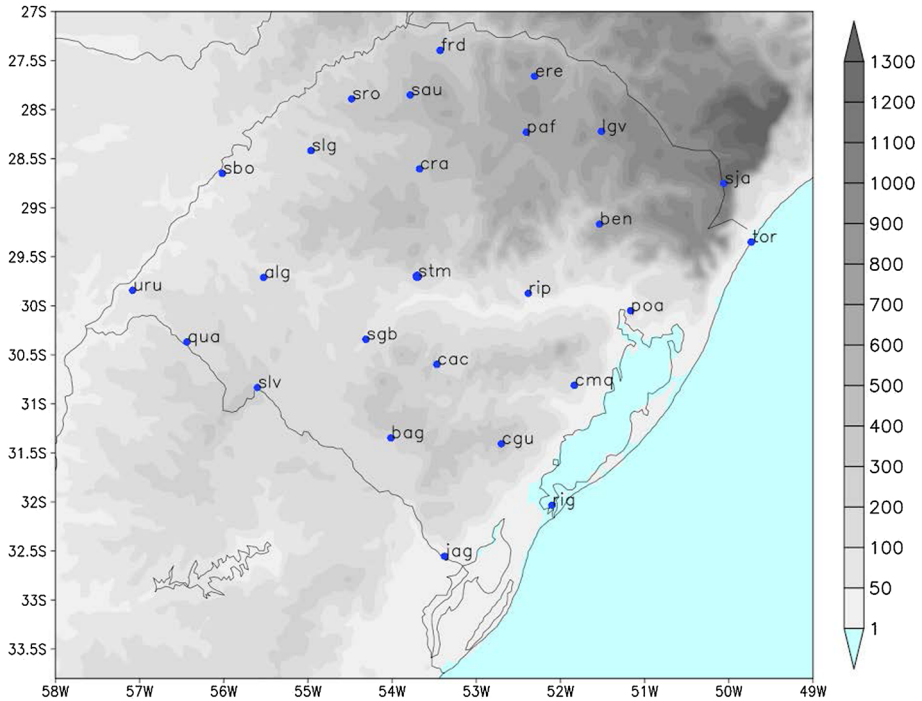
Fig. 3 Rio Grande do Sul state map, showing the locations of the stations used in the study. *Grey scale* shows topography (m)

tudes being located in the north-eastern portion of the State. At such higher regions, forests are still present. The western, central and south-western regions of the state are part of the Pampa plains, whose dominant vegetation is pasture, used for cattle grazing. Crop cultivation dominates the northern half of the state. Air temperature (at 2-m height) and wind speed (at 10 m) are observed as hourly grabbed samples at these stations, and those from 0000 to 0600 LST are used in the present study. The period of observations is during winter, with July climatologically the coldest month of the year.

The temperature and wind observations are compared to model observations at the nearest model grid point, following common practice (Buckley et al. 2004; Jiménez et al. 2010; Müller 2011; Soares et al. 2012; Xie et al. 2012). This procedure introduces an error because usually the nearest grid point and the station are located at different altitudes and in some cases this difference between the altitudes is large. In the present case, using the 12-km horizontal grid resolution, the lowest station with respect to its closest grid point is Santa Maria (grid point is 93 m higher relative to that of the station), while the opposite occurs at Caçapava do Sul, where the station is located at an altitude 200 m higher relative to the nearest grid point. When the 4-km horizontal resolution is used, these differences tend to diminish, but not substantially (Online Resource 1).

Because there is an altitude difference between station and the nearest grid point, it is important that the model temperatures are compensated for the adiabatic varia-tion before they are compared to the observations. This is presently done by subtracting $g \left(H_{station} - H_{gridpoint} \right) C_p^{-1}$ from the model temperature fields, where g is the acceleration due to gravity, taken as 9.8 m s^{-2}, H is the altitude of the station or nearest grid point with

respect to the mean sea level, and C_p is the specific heat of the air at constant pressure, taken as $1005 \, \mathrm{J \, kg^{-1} \, K^{-1}}$. At the station with the maximum height difference, Caçapava do Sul, this correction reduces the model temperature values by $2.0 \, ^{\circ}\mathrm{C}$, while at Santa Maria, the station lowest relative to the nearest grid point, the model temperature values are increased by $0.9 \, ^{\circ}\mathrm{C}$. In less than 10 % of the cases, the atmosphere was near saturation implying the need of a pseudo-adiabatic lapse rate for correction. Different corrections, dependent on the proximity to saturation, have been tested, minimally affecting the results. For this reason, to retain simplicity, the dry-adiabatic correction was used in all cases.

Cloud conditions have been analyzed in order to address whether there is any systematic temperature error associated with a given cloud cover. The cloud conditions were inferred for each hour and station using an algorithm based on brightness temperature (T_{sat}) from Geostationary Operational Environmental Satellite (GOES) 13 images. This was performed for 22 out of the 31 nights presently analyzed because such information was unavailable from 1 to 9 July. For each hour of analysis, an image produced within 20 min of the exact hour was used, and if no image was available in that period, no image was considered for that hour. The absence of images occurred in 16 % of the hours in the 22 nights considered. The cloud condition was classified as "clear sky" if $T_{sat} > (T_{station} - 5)^{\circ} \mathrm{C}$ where $T_{station}$ is the temperature observation from the surface network in $^{\circ}\mathrm{C}$. When $-20 \, ^{\circ}\mathrm{C} < T_{sat} < (T_{station} - 5)^{\circ} \mathrm{C}$, it was classified as "low cloud". The $5 \, ^{\circ}\mathrm{C}$ threshold is used because T_{sat} has been found to overestimate $T_{station}$ in clear-sky conditions (Prihodko and Goward 1997), and guarantees the correct classification of the cases classified as "low-clouds", although it is possible that some cases classified as "clear sky" actually presented very low clouds with warm cloud tops. It was classified as "mid-level clouds or shallow moist convection" when $-30 \, ^{\circ}\mathrm{C} < T_{sat} < -20 \, ^{\circ}\mathrm{C}$ and as "high-level cloud or deep moist convection" when $T_{sat} < -30 \, ^{\circ}\mathrm{C}$. This discrimination criterion has been based on the mean temperature of the cloud layer as determined by the observed temperature profile. Despite the efforts to ensure the best classification, it is a simple method and other errors may occur.

4 Results

4.1 Observed and Simulated Temperatures

The direct comparison between observed and simulated temperatures indicates that PBL parametrization schemes underestimate the nocturnal temperature (Fig. 4), in agreement with Xie et al. (2012) utilizing version 3 of the WRF model, Hu et al. (2013) utilizing the MYJ scheme and an updated version of the YSU scheme implemented in version 3.4.1, and Kleczek et al. (2014), using version 3.4.1. Using earlier WRF-model versions, previous studies that found temperature overestimation include Hu et al. (2010) for the YSU scheme, version 3.0.1, García-Díez et al. (2013) in winter with the MYJ and YSU schemes, version 3.1.1, and Bosveld et al. (2014), version 3.0. The aforementioned studies used different sets of PBL schemes. In general, the different choices of parametrization schemes give similar scatter with respect to the observations, with nearly identical root-mean-square (r.m.s.) error between parametrization schemes and observations (Table 2). Even using larger vertical or horizontal resolution has little effect on the mean r.m.s. error. The mean temperature bias is negative for all parametrization schemes, except for the BouLac scheme, confirming the tendency of the WRF model to underestimate surface temperatures. It should be noted that the higher temperatures are obtained with parametrization schemes that give higher turbulence
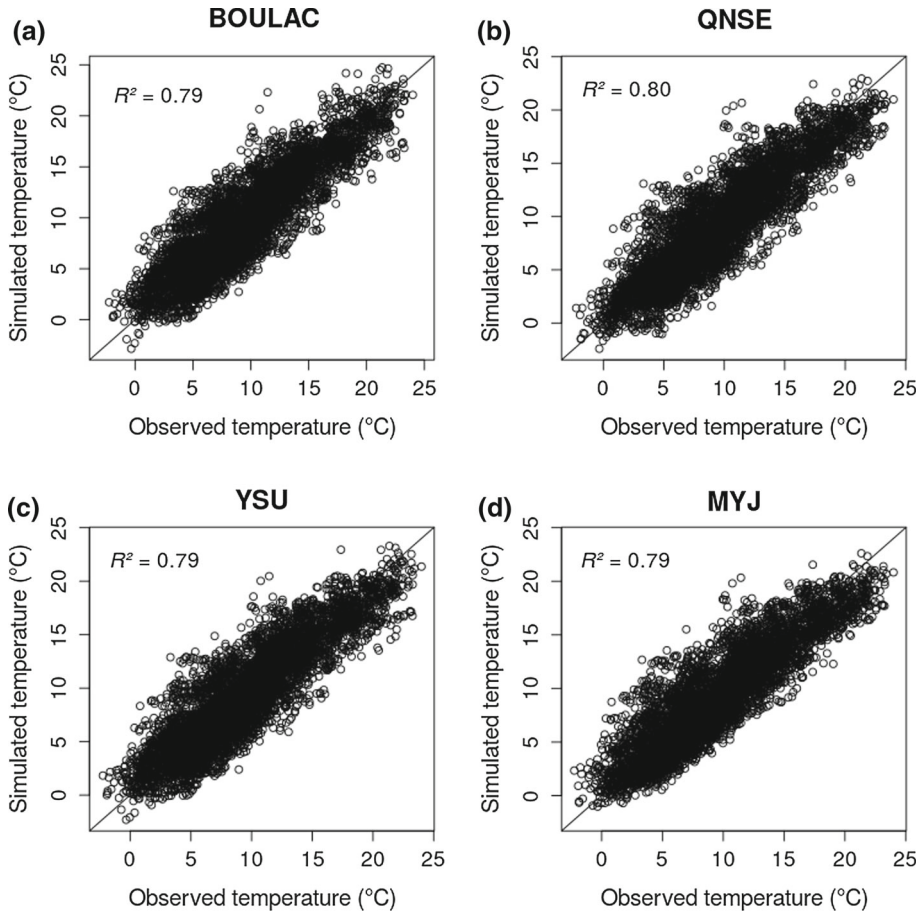
Fig. 4 Comparison between observed and simulated 2-m temperatures (height-corrected, as described in Sect. 3) for all nights and stations. Each *panel* shows the results using a different PBL scheme, as identified above the *panels*. All cases refer to the 12-km horizontal and lower vertical grid resolutions

levels for the same stability condition. For this reason, the BouLac scheme provides a warm bias (0.2 °C), while amongst the simulations with low horizontal and vertical resolutions, the QNSE scheme gives the coldest bias (-0.4 °C), despite yielding higher average turbulent intensity than does the MYJ scheme. This improved performance of the QNSE scheme is mostly influenced by the results from weak-wind cases, see below. This general result contrasts with Xie et al. (2012) and Kleczek et al. (2014), who found non-local PBL schemes, such as the YSU scheme, give higher temperatures than local schemes, such as the BouLac scheme. The use of a finer horizontal resolution reduces only slightly the cold bias of the MYJ scheme (to -0.2 °C, in run MYJ-4), and this improvement is mostly caused because the finer horizontal grid allows representing a given station by a closer grid point. Along this line, Müller (2011) found that enhanced horizontal resolution only improved the temperature forecast in some cases, mainly in complex terrain, while Zhang et al. (2013) stated that "… simulations at finer resolutions do not outperform those at coarser resolutions in most cases". The finer vertical resolution does not affect the mean r.m.s. error by much, but enhances the

Table 2 Root-mean square (r.m.s.) and mean bias errors of simulated 2-m temperatures with respect to the observed temperatures, for each simulation performed

	BouLac	QNSE	YSU	MYJ	MYJ-h	MYJ-4	MYJ-4h
r.m.s. error (°C)	2.4	2.4	2.4	2.3	2.4	2.4	2.4
Mean bias (°C)	0.2	−0.4	−0.1	−0.3	−0.4	−0.2	−0.4
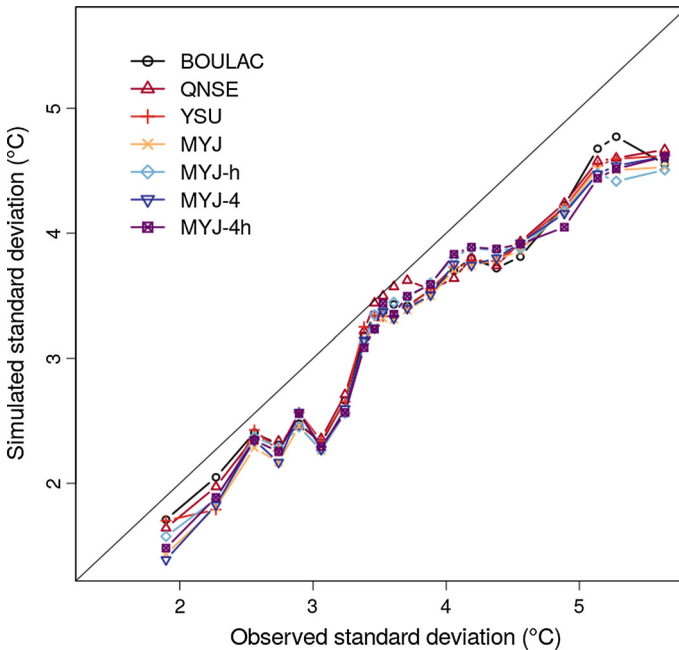
Fig. 5 Block-averaged simulated spatial standard deviation of 2-m temperatures inferred from the 26 values from all stations at each hour as a function of the same quantity inferred from the observations. Each *line* represents a different PBL scheme, horizontal and vertical grid resolutions used

cold bias of the corresponding MYJ simulations. This result is similar to that of Kleczek et al. (2014), who found slightly lower temperatures at night when a higher vertical resolution is used, and to that of Zhang et al. (2013), who found that the mean absolute error of nocturnal forecasts is not affected by the increase in the vertical resolution.

The model reduces the observed spatial variability, by underestimating the highest temperatures and slightly overestimating the smallest values. This tendency is confirmed when the observed spatial variability of temperatures over the network is compared with the variability of the simulated values. Regardless of the parametrization scheme, the spatial variability of temperatures in the model is always reduced with respect to what is observed across the surface stations (Fig. 5).

4.2 Dependence on Wind Speed

Larger deviations amongst the different parametrization schemes appear when the temperature forecast error is compared to the wind speed (Fig. 6a). In all cases, there is mean
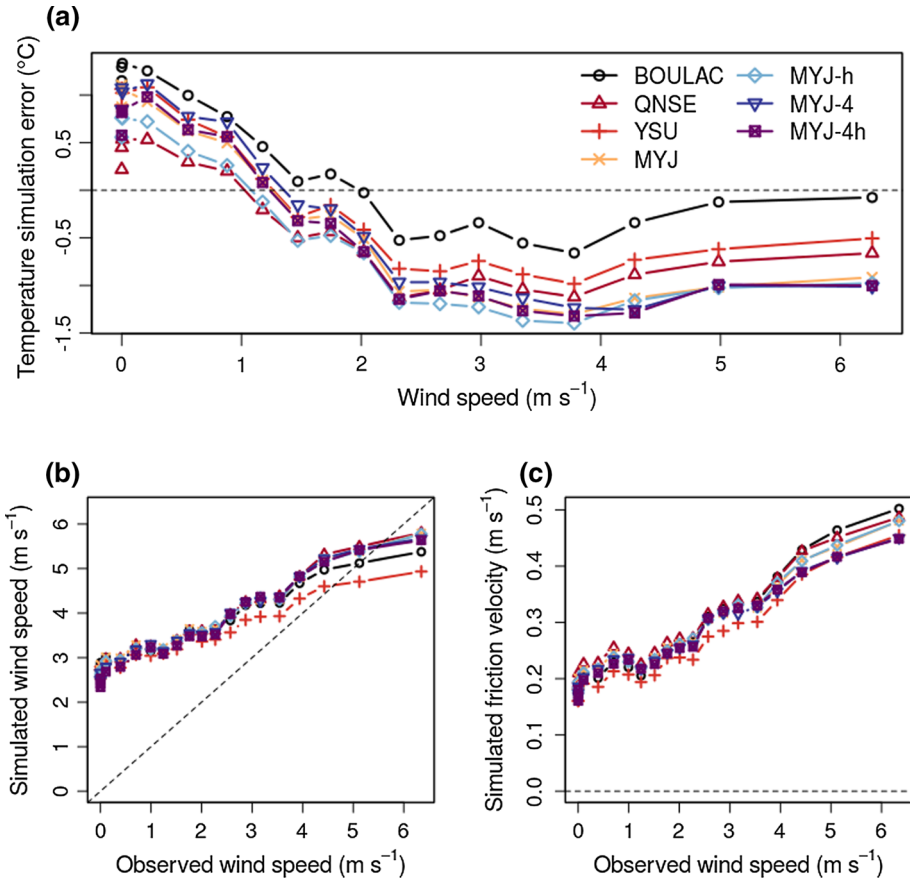
Fig. 6 a Block-averaged 2-m temperature simulation errors (height-corrected) as a function of the observed 10-m wind speeds, for each PBL scheme, horizontal and vertical grid resolutions used. **b** Same as in (**a**), but for block-averaged simulated 10-m wind speeds. **c** Same as in (**a**), but for block-averaged simulated friction velocities

temperature overestimation with observed small wind speeds, and underestimation with larger wind speeds. With larger wind speeds, the degree of temperature underestimation is directly dependent on the intensity of turbulence that the parametrization scheme gives for similar stabilities (show in Fig. 2). For wind speeds $>2\,\mathrm{m\,s^{-1}}$, the smallest cold bias is obtained with the scheme that provides highest turbulence levels (BouLac) and the cold bias progressively increases for the YSU, QNSE and MYJ schemes, in the same order as the turbulent intensity calculated by each scheme decreases. This fact indicates that, for wind speeds $>2\,\mathrm{m\,s^{-1}}$, parametrization schemes that provide higher turbulence levels for similar values of stability also produce higher temperatures, as expected. This result, along with the fact that in these conditions (observed wind speeds $>2\,\mathrm{m\,s^{-1}}$) all parametrization schemes underestimate the observed temperatures, leads to the conclusion that the temperature forecasts provided by all schemes tend to improve if the schemes provide higher turbulence levels in weakly stable conditions.

For weak winds, on the other hand, there is not a perfect correspondence between the temperature bias and the turbulence provided by each scheme. In such conditions, the QNSE

scheme performs better than all other schemes, having the smallest warm bias for observed wind speeds $<1\,\mathrm{m\,s}^{-1}$, showing that this scheme is suited for simulating very stable conditions. The use of higher vertical resolution affects mostly weak-wind conditions, as could be expected, given that this is the situation when the turbulence vertical scales are reduced. For small wind speeds, temperatures originated from a high vertical resolution model (labeled "MYJ-h" and "MYJ-4h" in Fig. 6a) are usually lower than the corresponding simulations with lower vertical resolution ("MYJ" and "MYJ-4", respectively). For the most stable cases, when there is a warm bias, this cooling effect causes the simulated temperatures to approach the observed values, but for wind speeds $>1.5\,\mathrm{m\,s}^{-1}$, the opposite occurs. For 10-m wind speeds $>5\,\mathrm{m\,s}^{-1}$, the vertical resolution has no effect on the average final temperatures.

The model tends to largely overestimate wind speed in weak-wind conditions, regardless of the chosen parametrization scheme (Fig.6b). For larger wind speeds, the wind-speed forecast approaches the observations on average, and these are the conditions when the temperature becomes underestimated. In the average for all stations, the friction velocity seems to approach a constant value at the low wind limit (Fig. 6c), a possible consequence of intentionally enhancing mixing for very stable conditions in all parameterization schemes. The cooling effect of the higher vertical resolution is associated with slightly lower turbulence levels (Fig. 6c). This fact suggests that the higher vertical resolution allows a better representation of surface decoupling in very stable conditions. However, Fig. 6c shows that the decrease in u_* with higher vertical resolution is very subtle indicating that, in most cases, doubling the vertical resolution is not enough to simulate properly this phenomenon.

A better understanding on how wind speed, turbulence and temperature interplay in the model with respect to the observations can be inferred if the results from specific stations are analyzed separately. In order to do that, the stations Alegrete and Caçapava are chosen, because these sites represent contrasting extremes of terrain. Alegrete station (121-m altitude) is located in relatively low terrain with respect to its surroundings, topography that favours nocturnal cold-air accumulation around the station. The consequence is that small wind speeds are common at Alegrete, where 60 % of the nocturnal observations show wind speeds $\leq 1\,\mathrm{m\,s}^{-1}$ (Fig. 7a). The problem in forecasting temperature at this station is enhanced because the nearest grid point is situated at a height considerably higher relative to the actual station (198 m in the 12-km grid, 181 m in the 4-km grid). The second station considered is at Caçapava do Sul (450-m altitude), where the actual altitude of the station is considerably higher relative to the nearest grid point (250 m in the 12-km grid, 268 m in the 4-km grid). At Caçapava, 88 % of the nocturnal observations used in the present study were between 2 and 6 m s^{-1} (Fig. 7a).

An obvious consequence of the fairly large differences in altitudes $H_{station} - H_{gridpoint}$ is that the WRF model wind speeds are consistently overestimated at Alegrete and often underestimated at Caçapava do Sul, for all parametrization schemes and regardless of the horizontal and vertical grids (Fig. 7b). The largest deviations between simulated wind speeds and observations occur for very small wind speeds that are common at Alegrete, which are not at all reproduced by the WRF model, suggesting that the model does not reproduce the decoupling of near-surface flow and that at higher levels in the SBL.

The wind-speed model bias explains the general trends of the temperature bias. In general, temperatures are always overestimated at Alegrete (Fig. 7c) and this fact can be directly attributed to the corresponding overestimation of wind speed. The largest warm bias occurs at Alegrete with observed small wind speeds, when all parametrization schemes give a reasonably large friction velocity (Fig. 7d). It is hypothesized here that most of the warm bias at Alegrete is associated with excessive turbulence in the model results. At Caçapava do Sul, on the other hand, temperatures are, on average, underestimated for all wind speeds, except for
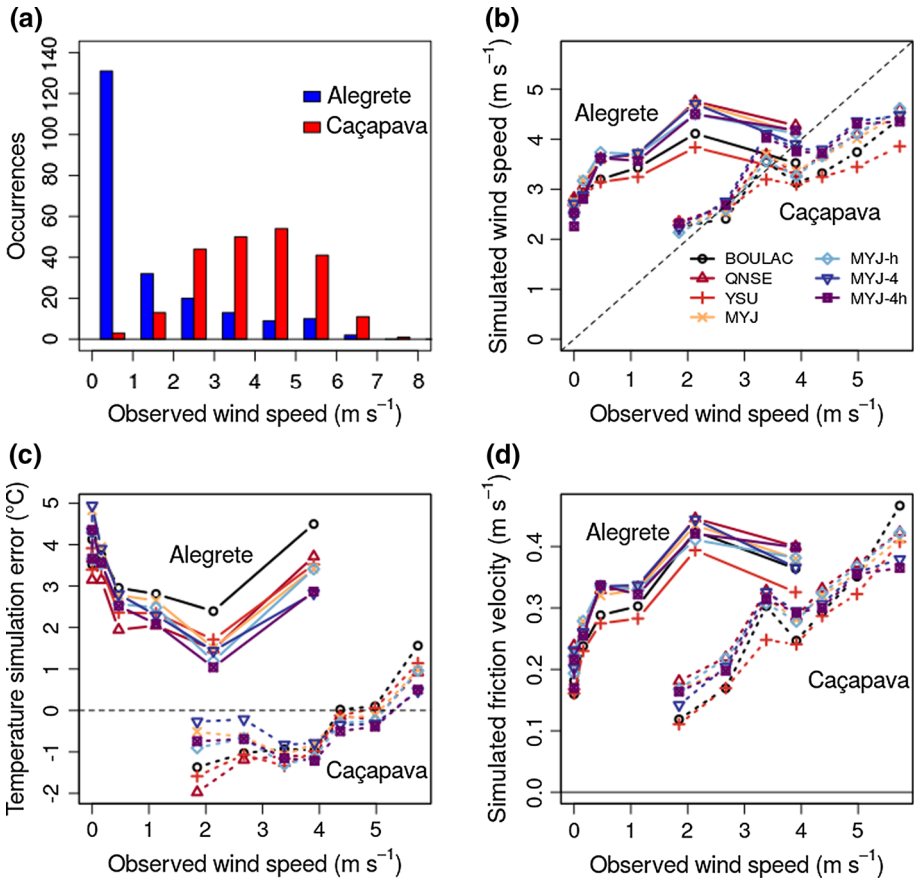
Fig. 7 **a** Frequency distributions of observed 10-m mean wind speeds in Alegrete (*blue*) and Caçapava do Sul (*red*). **b** Block-averaged simulated 10-m wind speeds as a function of the observed wind speeds for Alegrete (*solid lines*) and Caçapava do Sul (*dotted*), for each PBL scheme, horizontal and vertical grid resolutions used. **c** The same as in (**b**), but for the block-averaged 2-m temperature simulation errors (height-corrected). **d** The same as in (**b**), but for the block-averaged simulated friction velocities

the largest wind speeds (Fig. 7c). The largest cold bias at Caçapava do Sul occurs with wind speeds between 2 and 4 m s^{-1}, the same range for which the wind speed is well reproduced by the model (Fig. 7b). On the other hand, the good temperature forecasts at that station are obtained with wind speeds between 4 and 5 ms^{-1}(Fig. 7c); such wind speeds tend to be underestimated by all parametrization schemes (Fig. 7b).

4.3 Dependence on Terrain

When all stations are considered, it is clear that the altitude difference ($H_{station} - H_{gridpoint}$) explains a large fraction of the temperature bias, although other influences not being currently analyzed, such as the land use, may also be important. In all parametrization schemes, there is a clear tendency of stations that are lower relative to the grid point used for the comparison, having a positive temperature bias, while stations higher relative to the nearest grid point tend to show lower temperatures than the observations (Fig. 8). This is mainly caused by
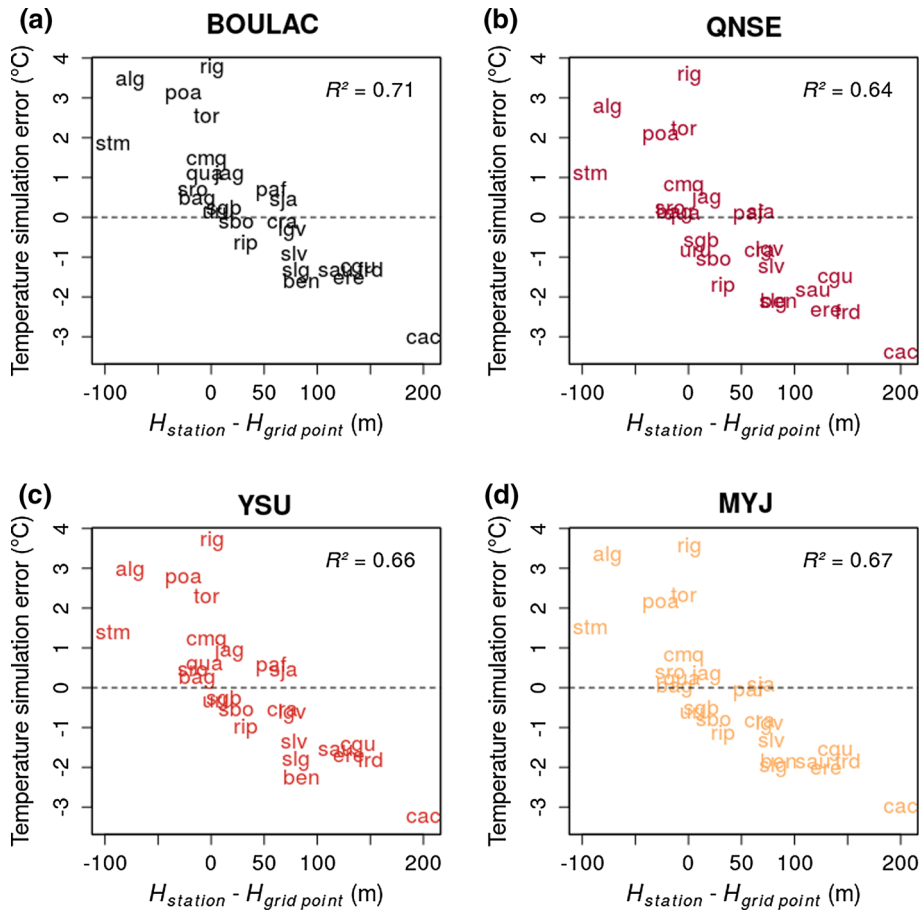
Fig. 8 2-m temperature simulation errors (height-corrected) for each station, averaged for all nights, as a function of altitude difference between station and model nearest grid point ($H_{station} - H_{gridpoint}$). Each *panel* refers to a different PBL scheme, as identified above the *panels*. All cases refer to the 12-km horizontal and lower vertical grid resolutions

the fact that wind speeds and, consequently, turbulent intensity, tend to be greater at higher altitudes. This fact alone causes the model to overestimate wind speeds in locations that are lower relative to the model grid point being compared. Furthermore, the model topography is a smoothed representation of a region, such that "the simulation tends to overestimate the wind speed over the valleys and to underestimate it at the mountain tops" (Jiménez et al. 2012). The excessive mixing at stations lower relative to the model grid point being compared warms the surface accordingly and the opposite occurs at stations that are higher relative to the closest grid point. If a station is at the bottom of a hill, for example, local terrain favours cold-air pooling favoring surface decoupling from the upper boundary layer. When it occurs, the observed nighttime temperature becomes proportional to the difference between the altitude of the station and that of its surroundings (Acevedo and Fitzjarrald 2003; Acevedo et al. 2013). Medeiros and Fitzjarrald (2014, 2015) showed that the convex locations served at 'hot spots' that fostered surface-atmosphere exchange on strongly stable nights, and this points toward seeking better horizontal resolution to define such locations in a landscape.

To compare the results for the different nights, it is important to classify them in terms of stability. To do so, based solely on the observations provided by the station network, a "spatial Richardson number" is defined as

$$Ri_{spat} \equiv \frac{g}{\bar{\Theta}} \Delta z \left(\frac{\theta_{\max} - \theta_{\min}}{V_{mean}^2} \right), \tag{1}$$

where $\theta_{\max}$ and $\theta_{\min}$ are respectively the maximum and minimum potential temperatures observed across the entire network for a given night, and V_{mean} and $\bar{\Theta}$ are respectively the mean wind speed and potential temperature for the entire network during that night. The height of the wind-speed observations, 10 m, is used for Δz.

Equation 1 provides a value of Ri_{spat} for each night and is, therefore, used as a stability classifier for the entire network. It is based on the idea that the horizontal temperature variability is enhanced in more stable conditions, while in less stable conditions the increased turbulent mixing provides horizontal temperature homogenization. Although the idea of a spatial Richardson number may seem to contradict the definition of the parameter, it is used here merely as a means of contrasting the background SBL state for the entire network over the different nights (Online Resource 2). Medeiros and Fitzjarrald (2014) applied the concept of a regional bulk Richardson number, which is applied to a spatially smaller network of stations than that used herein.

Among all nights, the largest wind speeds and corresponding lowest Ri_{spat} occurred on 31 July 2012. In this case, the error is not highly correlated with $H_{station} - H_{gridpoint}$, as given by the R^2 values between these two quantities, which never exceed 0.24 for any parametrization scheme (figure not shown). The most stable night was 26 July 2012, when much higher correlation coefficients between temperature simulation error and $H_{station} - H_{gridpoint}$, occurred, with R^2 values always exceeding 0.6, indicative that lower locations more likely experience decoupling. For example, the Alegrete station (Fig. 7) is much warmer than is indicated by the observations. On the very stable night of 26 July 2012, the QNSE scheme showed the smallest correlation between temperature error and $H_{station} - H_{gridpoint}$ and also the smallest r.m.s. error between forecasted and observed temperatures. This is a further evidence that this scheme is the most appropriate to simulate this type of night. The use of finer vertical grid (figure not shown) improves the temperature simulation for this very stable night (r.m.s. error of 2.7 °C for the MYJ scheme decreasing to 2.5 °C for the MYJ-h scheme), indicating that in such an environment, with reduced spatial scales of turbulence, the use of higher vertical resolution provides a better representation of the physical process.

The overall results are summarized in Fig. 9. In general, at stations higher relative to the nearest grid point the model provides lower temperatures than the observations, a fact that is independent of parametrization scheme and grid resolution. More interestingly, this result is also generally independent of nocturnal stability. This is because at those stations located at higher locations relative to the nearest grid point, there is a general tendency of windier and more turbulent conditions regardless of the night. In such locations, the surface tends to connect to the upper SBL both in the real world and in the model. At stations with positive $H_{station} - H_{gridpoint}$ the wind speed is generally overestimated, with the exception of some extreme cases with largest $H_{station} - H_{gridpoint}$ (Fig. 10). Therefore, although representing a lower location, the model still overestimates wind speed, a result similar to that found earlier (e.g., Zhang and Zheng 2004; Svensson et al. 2011; Jiménez and Dudhia 2012; Xie et al. 2012; Hu et al. 2013). At stations lower relative to the nearest grid point (negative $H_{station} - H_{gridpoint}$), on the other hand, the average temperature simulation error is highly dependent on Ri_{spat}. In these cases large warm bias occurs on the most stable nights, when such lower stations experience surface decoupling, not reproduced by any parametrization
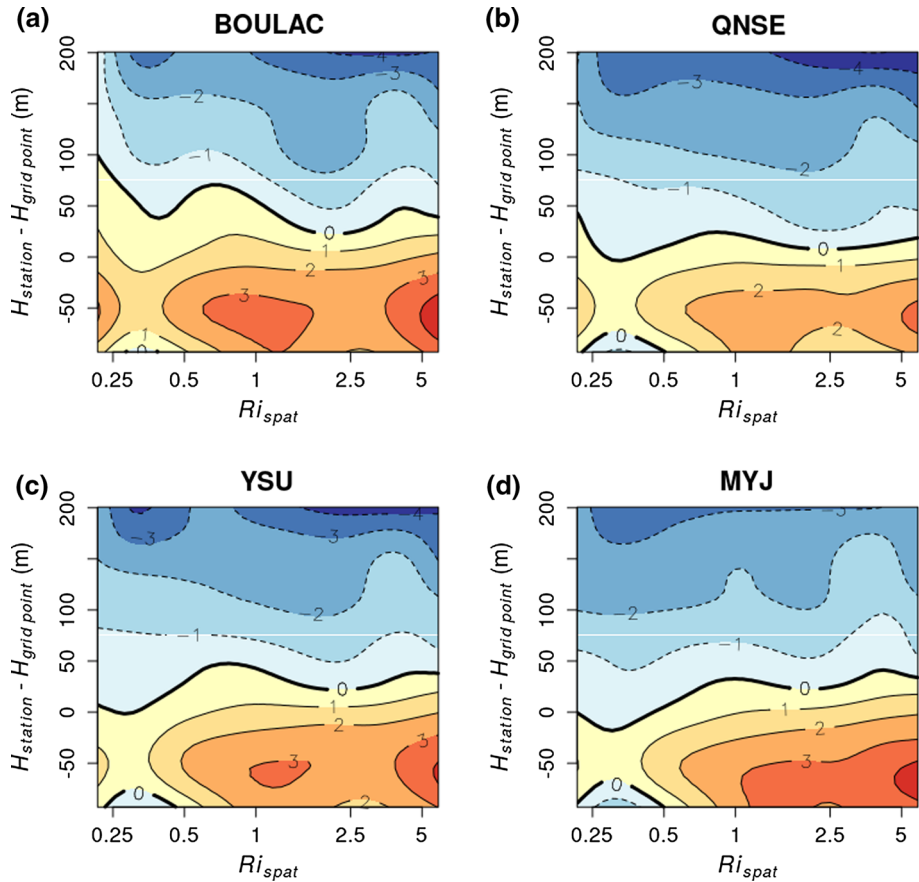
Fig. 9 Average 2-m temperature simulation errors (height-corrected) as a function of the spatial Richardson number (Ri_{spat}, see text) and the altitude difference between station and model nearest grid point ($H_{station} - H_{gridpoint}$). Each *panel* refers to a different PBL scheme, as identified above the *panels*. All cases refer to the 12-km horizontal lower vertical grid resolutions

scheme presently being compared. However, the average error at these stations approaches zero or becomes slightly negative on weakly stable nights, when windier conditions prevail over the entire area and decoupling is unlikely, even at lower locations.

4.4 Dependence on Cloudiness

When the temperature error is classified in terms of the cloud condition, as described in Sect. 3, some trends become clear (Table 3). The largest r.m.s. errors occur in clear-sky conditions, while the smallest r.m.s. errors occur for high-level clouds or deep convection. For cloudy conditions, the largest errors occur for low-level clouds.

Weather situations with observed low-level clouds are the only conditions in which the WRF model produces a mean cold bias regardless of the parametrization scheme, although such a bias varies substantially among them (from -0.1 for BouLac to -0.7 for MYJ-4h). This result may indicate a difficulty of the model in properly resolving this cloud type, therefore, affecting the temperature forecasts. A similar result has been found for simulations
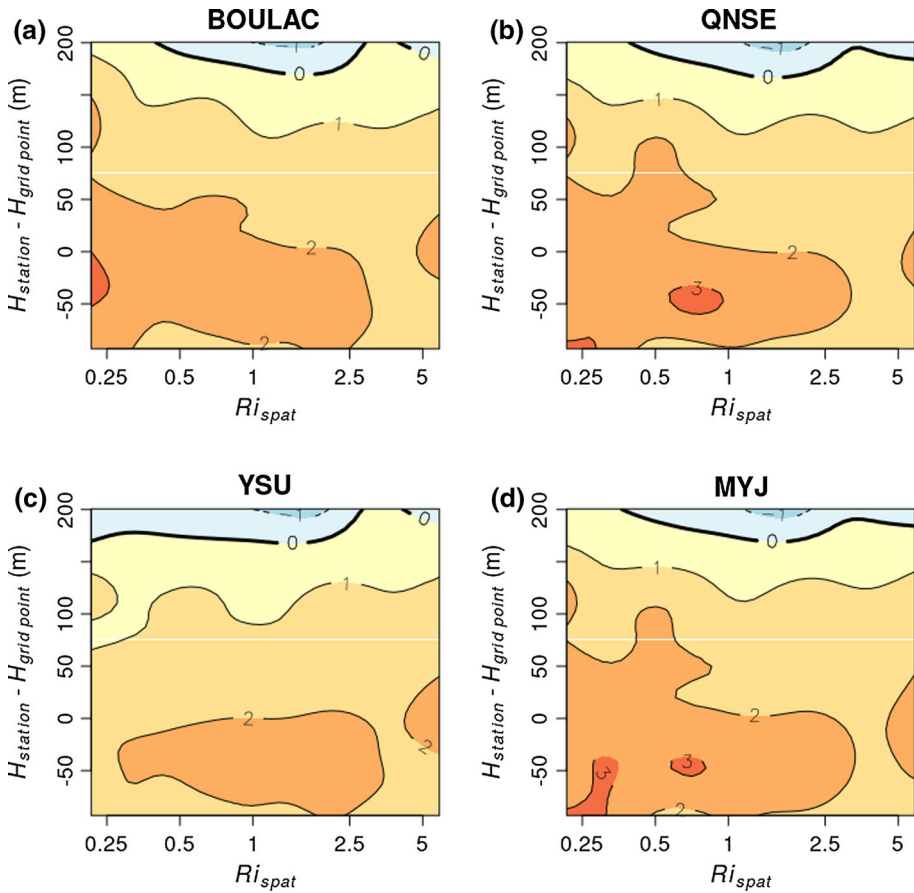
Fig. 10 The same as in Fig. 9, but for the average 10-m mean wind speed errors

of high-latitude weather by Dong and Mace (2003), Hines et al. (2011) and Bromwich et al. (2013) who concluded that "the model has difficulty predicting the downward longwave radiation and hence the surface energy balance…" and that "inadequate cloud representation is inferred to be responsible for the summer cold bias". Furthermore, it is also interesting to notice that the mean bias for low-clouds is generally colder for parametrization schemes with lower turbulence levels, exactly as it has been found to occur in windy conditions (Sect. 4.2) and on the weakly stable nights (Sect. 4.3).

To isolate whether the cold bias is produced by problems in the turbulence or in the low-cloud representation, the mean temperature bias is compared for cases with low clouds or clear skies and those with small or large wind speeds (Table 4). For simplification, in this analysis, only the averaged bias among all parametrization schemes is compared, because these results are not largely dependent on the parametrization schemes. For mean wind speeds $>2\,\mathrm{m\,s^{-1}}$, the cold bias occurs regardless of cloud condition. For windy, clear-sky conditions, the averaged bias of $-0.8\,^{\circ}\mathrm{C}$ shows that not all cold bias observed is produced by cloud misrepresentation. On the other hand, for the same range of wind speeds, the cold bias is enhanced to $-1.3\,^{\circ}\mathrm{C}$ when low clouds are present, indicating that their misrepresentation increases the problem. We conclude that the nocturnal temperature forecast improves if each

Table 3 R.m.s. and mean bias errors of simulated 2-m temperatures with respect to the observed temperatures, for each simulation performed, classified according to cloud condition inferred from satellite image

(°C)	BouLac	QNSE	YSU	MYJ	MYJ-h	MYJ-4	MYJ-4h	Mean
R.m.s. error clear sky	2.6	2.5	2.5	2.4	2.4	2.5	2.5	2.5
R.m.s. error low-level clouds	2.5	2.5	2.4	2.4	2.4	2.3	2.4	2.4
R.m.s. error mid-level clouds or shallow moist convection	2.2	2.0	2.2	1.9	2.1	1.9	2.1	2.1
R.m.s. error high-level clouds or deep moist convection	2.2	2.0	2.1	2.0	2.0	1.9	1.8	2.0
Mean bias clear sky	0.4	−0.2	0.1	0.1	−0.2	0.2	−0.1	0.0
Mean bias low-level clouds	−0.1	−0.6	−0.3	−0.5	−0.6	−0.6	−0.7	−0.5
Mean bias mid-level clouds or shallow moist convection	0.4	0.1	0.4	0.2	0.1	0.1	−0.0	0.2
Mean bias high-level clouds or deep moist convection	0.7	0.0	0.4	0.0	0.0	−0.1	−0.1	0.1

Table 4 Mean bias errors of simulated 2-m temperatures with respect to the observed temperatures (°C), averaged for all simulations, classified in terms of the two most common cloud conditions found and 10-m mean wind speed

	Clear sky	Low-level clouds
Wind speeds $<2\,\mathrm{m\,s}^{-1}$	0.7 (1212 cases)	0.2 (449 cases)
Wind speeds $>2\,\mathrm{m\,s}^{-1}$	−0.8 (887 cases)	−1.3 (399 cases)

model artificially allows for higher turbulence in weakly stable conditions. This improvement arises partially as a counteraction to poor low-cloud representation or deficiencies in other parametrization schemes, which are beyond the scope of the study.

5 Conclusion

We analyzed the influence of turbulence parametrization schemes on the quality of nocturnal temperature forecasts using the WRF model. The study is incomplete, since errors may arise due to other factors, such as deficiencies in the surface-layer scheme, in the land-surface scheme, in the radiation scheme, or by errors in the GFS boundary and initial conditions. A thorough analysis that included surface radiation and energy balance, as well as vertical profiles of the mean and turbulent quantities was not possible because there were no observations available of these variables.

We examined 31 nights, each with distinct nocturnal stability conditions at 26 sites over a broad area with diverse surface characteristics. Furthermore, four PBL parametrization schemes have been compared: BouLac, QNSE, YSU and MYJ. For the MYJ scheme, different horizontal and vertical grid resolutions have also been employed. Based on the results, a proper adjustment of the PBL schemes may correct the following deficiencies found in the model:

(1) The model has a cold bias in windy conditions, regardless of the PBL scheme used. Such a bias is reduced for formulations that allow higher turbulence levels for a given value of the stability parameter z/L, such as the BouLac scheme. This result, therefore, indicates that enhancement of turbulence in weakly-stable conditions tends to improve the nocturnal temperature forecasts. On the other hand, such turbulence enhancement also tends to increase the model 10-m wind speed. Given that this variable is generally overestimated by the model, such an action tends to worsen the model wind-speed forecast. It has also been found that the model has a cold bias when low-level clouds are present. Given that low-level clouds tends to occur along with weakly-stable conditions, in many of these situations the cold bias may be partially attributed to the low-level cloud misrepresentation. However, a mean cold bias is also observed when large wind speeds occur with clear skies, indicating that the poor turbulence representation also plays a role.

(2) The PBL schemes compared have difficulties in reproducing the decoupling between the surface and the upper SBL. This conclusion, also obtained in Shin and Hong (2011), is inferred from the fact that on the more stable nights the model temperature errors are highly correlated with $H_{station} - H_{gridpoint}$. On those nights, stations at a lower altitude to their surroundings (where $H_{station} - H_{gridpoint}$ is likely to be negative) experience decoupling in the real world, but not in the model, where the maintenance of turbulence throughout the night produces a warm bias. Therefore, the inability of the model to reproduce correctly the decoupled state is responsible for a large diversity of temperature biases on the low-wind nights, with a large warm bias at stations situated at lower heights relative to the nearest grid point and large cold bias otherwise. It is important to stress that the QNSE scheme performs appreciably better relative do the others in such very stable conditions.

The use of finer horizontal and vertical grids has little effect on the temperature forecasts. Slight improvements have been found with higher vertical resolution on the most stable nights, but in windier conditions the results are insensitive to the vertical grid resolution.

There is an inherent difficulty forecasting nocturnal temperatures in very stable conditions because cold-air pooling and consequent SBL surface decoupling lead to large variability over relatively small horizontal distances. Observations have shown large temperature differences over distances as small as a few hundred metres (Bodine et al. 2009; Acevedo et al. 2013). No mesoscale operational numerical weather forecast model can reproduce such variability, although recently Boutle et al. (2016) have shown that this may be improved through the use of a very fine nested grid. Therefore, an interesting strategy to predict the spatial distribution of temperatures on such nights may be to perform a post-processing algorithm, which relates surface temperature to height difference between a given location and its surroundings. An advantage of such an approach is that it can be applies to a much finer horizontal grid than that of the model simulations The near-linear relationship between temperature simulation error and $H_{station} - H_{gridpoint}$ found in the present study during those nights suggests that a post-processing correction of the kind may be very effective and fairly simple to implement. However, the successful implementation of such a correction also demands a better solution of the coupling state of the SBL by the numerical model.

Appendix: Brief Description of PBL Parametrization Schemes Used

Bougeault–Lacarrere (BouLac)

The scheme proposed by Bougeault and Lacarrere (1989) can be classified as a local turbulent kinetic energy (TKE) closure parametrization, where the turbulent exchange coefficients (K) depend on TKE, which is locally determined from its budget equation. Turbulent fluxes are given by a simple K-theory approach, viz

$$\overline{w'\chi'} = -K_\chi \frac{\partial \chi'}{\partial z}, \tag{2}$$

where K_χ is a generic variable, e.g. the horizontal wind-speed components, potential temperature, specific humidity or TKE.

The exchange coefficients K_χ are related to TKE as

$$K_\chi = Cle^{1/2}, \tag{3}$$

where C is a constant equal to 0.4, e is the TKE determined at each grid point by its budget equation and l is a characteristic length, defined as the distance that a particle with initial energy equal to the average TKE at its vertical level travels upward or downward due to buoyancy. In the SBL, such a thickness becomes small, and it is assumed that the thermal stratification is constant, so that the aforementioned integral leads to (Bougeault and Lacarrere 1989)

$$l = e^{1/2} \left(\frac{g}{\theta'} \frac{\partial \theta'}{\partial z} \right)^{-1/2}. \tag{4}$$

The BouLac boundary-layer module has been coupled with the surface-layer scheme fifth-generation Pennsylvania State University—National Center for Atmospheric Research Mesoscale Model (MM5). The exchange coefficients used are determined from similarity relationships proposed by Dyer and Hicks (1970), Paulson (1970) and Zhang and Anthes (1982). Such expressions directly relate the eddy diffusivity to the atmospheric stability and can be, therefore, regarded as stability functions. Two nocturnal stability conditions are considered, classified in terms of the Richardson number (Ri). When Ri falls below a critical value of 0.2, the stability functions are applied. If the stability is such that the critical Richardson number is exceeded, turbulence is set to a minimum value. In the model 3.6.1 version used herein, the revision proposed by Jiménez et al. (2012) is in effect. This reduces the minimum friction velocity u_*, previously assumed as 0.1, to $0.001\,\mathrm{m\,s^{-1}}$, with the purpose of allowing conditions with reduced turbulence, but maintaining u_* greater than zero to avoid runaway cooling. A minimum TKE value of $e = 0.2\,\mathrm{m^2\,s^{-2}}$ is also imposed.

Quasi-Normal Scale Elimination (QNSE)

The QNSE scheme has been proposed by Sukoriansky et al. (2005), and uses Eq. 3 to determine the exchange coefficients for the neutral case. Here, $C = 0.55$, e is, again, determined at each grid point by solving the TKE budget and the characteristic length l is given by

$$\frac{1}{l} = \frac{1 + \frac{\kappa z}{\lambda}}{\kappa z} + 1.33 \frac{N}{e^{1/2}}. \tag{5}$$

In (5), the first term in the right is the Blackadar scale, where $\lambda = 0.0063 u_*/f$, where f is the Coriolis parameter. The second term on the right side of (5) represents the turbulence dampening effect by the stratification, where N is the Brunt–Väisälä frequency. For non-neutral conditions, the exchange coefficients must be multiplied by stability functions α_M and α_H for momentum and heat, respectively, which depend locally on the gradient Ri. The QNSE scheme uses a specific surface-layer parametrization scheme, described by Sukoriansky (2008).

Yonsei University (YSU)

The YSU boundary-layer scheme is a first-order non-local closure, with fluxes again related to local gradients through (2), but, being a first-order scheme, the turbulent exchange coefficients do not depend on higher order statistical moments, such as the TKE. The non-local nature of the scheme arises from the dependence of the exchange coefficients on the SBL height,

$$K_\chi = \kappa u_* \phi_\chi^{-1} \left(1 - \frac{z}{h}\right)^2, \tag{6}$$

where κ is the von Karman constant, u_* is provided by the surface-layer module, ϕ_χ is a dimensionless gradient that acts as a stability function, z is the height above the ground and h is the SBL height. The stability function used is $\phi_\chi = 1 + 5\,(z/L)$ (Hu et al. 2013), where L is the Obukhov length. The SBL thickness h is evaluated from the virtual potential temperature and wind-speed vertical profiles (Troen and Mahrt 1986; Hong and Pan 1996). Equation 6 is used for the momentum exchange coefficient and those for heat and moisture are corrected by the turbulent Prandtl number (Hong and Pan 1996).

The surface-layer module that is coupled to the YSU boundary-layer scheme is MM5, the same used with the BouLac scheme and briefly described above.

Mellor–Yamada–Janjic (MYJ)

The MYJ boundary-layer scheme is similar to the BouLac scheme in the sense that it is also classified as a local TKE closure. Turbulent fluxes are again determined from (2), but the important difference lies in the exchange coefficient K_χ, which is given by

$$K_\chi = l\,(2e)^{1/2}\,S_\chi. \tag{7}$$

In (7), l is a master length scale, whose determination is given by Janjic (2002) and to which a stability-dependent maximum value is imposed. Stability functions S_χ (which differ for momentum and heat) are used. These expressions impose the reduction of the exchange coefficients as a function of stability.

The ETA surface-layer module (also known as the MYJ surface-layer module) is coupled to the MYJ boundary-layer scheme, where the exchange coefficients are given by

$$K_\chi = l^2 \left[S\sqrt{B\left(1 - R_f\right)S_\chi}\right] S_\chi, \tag{8}$$

where S is the wind shear and B is a numerical constant. The flux Richardson number R_f is determined from the gradient Ri, and S_χ are, again, stability functions, different for momentum and heat and also different from those used in the boundary-layer module. According to Janjic (2002), "… non-zero equilibrium conditions are possible for $Ri \leq$

0.505", which effectively means that there is a critical Richardson number $Ri_c = 0.505$, above which mixing is suppressed. However, there are no such occurrences in the simulations presently described.

References

Acevedo OC, Fitzjarrald DR (2001) The early evening surface-layer transition: temporal and spatial variability. J Atmos Sci 58:2650–2667

Acevedo OC, Fitzjarrald DR (2003) In the core of the night-effects of intermittent mixing on a horizontally heterogeneous surface. Boundary-Layer Meteorol 106:1–33

Acevedo OC, Oliveira P, Santos DM, Silva GM, Medeiros LE, Quadro MFL, Fuentes MV, Bastos MS, Nedel AS (2013) Estudo da turbulência atmosférica noturna sobre coxilhas (estância). Parte 2: resultados preliminares. Ciencia & Natura-Special edition Dec. 2013:392–395

Bodine D, Klein PM, Arms SC, Shapiro A (2009) Variability of surface air temperature over gently sloped terrain. J Appl Meteorol Climatol 48(6):1117–1141

Bosvled FC, Baas P, Steeneveld GJ, Holtslag AAM, Angevine WM, Bazile E, de Bruijn EIF, Deacu D, Edwards JM, Ek M, Larson VE, Pleim JE, Raschendorfer M, Svensson G (2014) The third GABLS intercomparison case for evaluation studies of boundary-layer models. Part B: results and process understanding. Boundary-Layer Meteorol 152(2):157–187

Bougeault P, Lacarrere P (1989) Parameterization of orography-induced turbulence in a mesobeta-scale model. Mon Weather Rev 117(8):1872–1890

Boutle IA, Finnenkoetter A, Lock AP, Wells H (2016) The London Model: forecasting fog at 333 m resolution. Q J R Meteorol Soc 142(694):360–371

Bromwich DH, Otieno FO, Hines KM, Manning KW, Shilo E (2013) Comprehensive evaluation of polar weather research and forecasting model performance in the Antarctic. J Geophys Res Atmos 118(2):274–292

Buckley RL, Weber AH, Weber JH (2004) Statistical comparison of Regional Atmospheric Modelling System forecasts with observations. Meteorol Appl 11(1):67–82

Cuxart J, Holtslag AAM, Beare RJ, Bazile E, Beljaars A, Cheng A, Conangla L, Ek M, Freedman F, Hamdi R, Kerstein A, Kitagawa H, Lenderink G, Lewellen D, Mailhot J, Mauritsen T, Perov V, Schayes G, Steeneveld G-J, Svensson G, Taylor P, Weng W, Wunsch S, Xu K-M (2006) Single-column model intercomparison for a stably stratified atmospheric boundary-layer. Boundary-Layer Meteorol 118(2):273–303

Delage Y (1997) Parameterising sub-grid scale vertical transport in atmospheric models under statically stable conditions. Boundary-Layer Meteorol 82(1):23–48

Dong X, Mace GG (2003) Arctic stratus cloud properties and radiative forcing derived from ground-based data collected at Barrow, Alaska. J Clim 16(3):445–461

Dudhia J (1989) Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. J Atmos Sci 46(20):3077–3107

Dyer AJ, Hicks BB (1970) Flux–gradient relationships in the constant flux layer. Q J R Meteorol Soc 96(410):715–721

García-Díez M, Fernández J, Fita L, Yagüe C (2013) Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. Q J R Meteorol Soc 139(671):501–514

Gustavsson T, Karlsson M, Bogren J, Lindqvist S (1998) Development of temperature patterns during clear nights. J Appl Meteorol Climatol 37(6):559–571

Hines KM, Bromwich DH, Bai L-S, Barlage M, Slater AG (2011) Development and testing of polar WRF. Part III: Artic Land. J Clim 24(1):26–48

Holtslag AAM, Svensson G, Baas P, Basu S, Beare B, Beljaars ACM, Bosveld FC, Cuxart J, Lindvall J, Steeneveld G-J, Tjernström M (2013) Stable atmospheric boundary-layers and diurnal cycles—challenges for weather and climate models. Bull Am Meteorol Soc 94(11):1691–1706

Hong S-Y, Pan H-L (1996) Nonlocal boundary-layer vertical diffusion in a medium-range forecast model. Mon Weather Rev 124(10):2322–2339

Hong S-Y, Dudhia J, Chen S-H (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Mon Weather Rev 132(1):103–120

Hong S-Y, Noh Y, Dudhia J (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev 134(9):2318–2341

Hu X-M, Nielsen-Gammon JW, Zhang F (2010) Evaluation of three planetary boundary-layer schemes in the WRF model. J Appl Meteorol Climatol 49(9):1831–1844

Hu X-M, Klein PM, Xue M (2013) Evaluation of the updated YSU planetary boundary-layer scheme within WRF for wind resource and air quality assessments. J Geophys Res Atmos 118(18):10490–10505

Janjic ZI (1994) The step-mountain Eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes. Mon Weather Rev 122(5):927–945

Janjic ZI (2002) Nonsingular Implementation of the Mellor–Yamada Level 2.5 Scheme in the NCEP Meso Model. Technical Report, NCEP, p 61

Jiménez PA, Dudhia J (2012) Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF model. J Appl Meteorol Climatol 51(2):300–316

Jiménez PA, González-Rouco JF, García-Bustamante E, Navarro J, Montávez JP, de Arellano JV-G, Dudhia J, Muñoz-Roldan A (2010) Surface wind regionalization over complex terrain: evaluation and analysis of a high-resolution WRF simulation. J Appl Meteorol Climatol 49(2):268–287

Jiménez PA, Dudhia J, González-Rouco JF, Navarro J, Montávez JP, García-Bustamante E (2012) A revised scheme for the WRF surface layer formulation. Mon Weather Rev 140(3):898–918

Kain JS (2004) The Kain–Fritsch convective parameterization: an update. J Appl Meteorol Climatol 43(1):170–181

Kleczek MA, Steeneveld GJ, Holtslag AAM (2014) Evaluation of the weather research and forecasting mesoscale model for GABLS3: impact of boundary-layer schemes, boundary conditions and spin-up. Boundary-Layer Meteorol 152(2):213–243

Lin Y-L, Farley RD, Orville HD (1983) Bulk parameterization of the snow field in a cloud model. J Clim Appl Meteorol 22(6):1065–1092

Louis JF (1979) A parametric model of vertical fluxes in the atmosphere. Boundary-Layer Meteorol 17(2):187–202

Mahrt L (1987) Grid-averaged surface fluxes. Mon Weather Rev 115(8):1550–1560

Mahrt L (1999) Stratified atmospheric boundary layers. Boundary-Layer Meteorol 90(3):375–396

Mahrt L (2006) Variation of surface air temperature in complex terrain. J Appl Meteorol Climatol 45:1481–1493

Medeiros LE, Fitzjarrald DR (2014) Stable boundary-layer in complex terrain. Part I: linking fluxes and intermittency to an average stability index. J Appl Meteorol Climatol 53(9):2196–2215

Medeiros LE, Fitzjarrald DR (2015) Stable boundary-layer in complex terrain. Part II: geometrical and sheltering effects on mixing. J Appl Meteorol Climatol 54(1):170–188

Mellor GL, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. Rev Geophys 20(4):851–875

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA (1997) Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J Geophys Res Atmos 102(14):16663–16682

Müller MD (2011) Effects of model resolution and statistical postprocessing on shelter temperature and wind forecasts. J Appl Meteorol Climatol 50(8):1627–1636

Paulson CA (1970) The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. J Appl Meteorol Climatol 139:261–281

Prihodko L, Goward SN (1997) Estimation of air temperature from remotely sensed surface observations. Remote Sens Environ 60:335–346

Shin HH, Hong S-Y (2011) Intercomparison of Planetary Boundary-Layer Parameterizations in the WRF model for a single day from CASES-99. Boundary-Layer Meteorol 152(2):213–243

Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X-Y, Wang W, Powers JG (2008) A description of the advanced research WRF version 3. Technical Report, NCAR, 113 pp

Soares PMM, Cardoso RM, Miranda PMA, Medeiros J, Belo-Pereira M, Espirito-Santo F (2012) WRF high resolution dynamical downscaling of ERA-Interim for Portugal. Clim Dyn 39(9):2497–2522

Steeneveld G-J, van de Wiel BJH, Holtslag AAM (2006) Modeling the evolution of the atmospheric boundary-layer coupled to the land surface for three contrasting nights in CASES-99. J Atmos Sci 63(3):920–935

Sukoriansky S (2008) Implementation of the Quasi-Normal Scale Elimination (QNSE) model of stably stratified turbulence in WRF. Technical Report, Report on WRF-DTC Visit

Sukoriansky S, Galperin B, Perov V (2005) Application of a new spectral theory of stably stratified turbulence to the atmospheric boundary layer over sea ice. Boundary-Layer Meteorol 117(2):231–257

Svensson G, Holtslag AAM, Kumar V, Mauritsen T, Steeneveld G-J, Angevine WM, Bazile E, Beljaars A, de Bruijn EIF, Cheng A, Conangla L, Cuxart J, Ek M, Falk MJ, Freedman F, Kitagawa H, Larson VE, Lock A, Mailhot J, Masson V, Park S, Pleim J, Söderberg S, Weng W, Zampieri M (2011) Evaluation of the diurnal cycle in the atmospheric boundary-layer over land as represented by a variety of single-column models: the second GABLS experiment. Boundary-Layer Meteorol 140(2):177–206

Tewari M, Chen F, Wang W, Dudhia J, LeMone MA, Mitchell KE, Ek M, Gayno G, Wegiel JW, Cuenca RH (2004) Implementation and verification of the unified Noah land-surface model in the WRF model. In:

20th Conference on weather analysis and forecasting/16th conference on numerical weather prediction. American Meteorological Society, Seattle, WA, pp 11–15

Troen IB, Mahrt L (1986) A simple model of the atmospheric boundary-layer: sensitivity to surface evaporation. Boundary-Layer Meteorol 37(1):129–148

Wang W, Bruyère C, Duda M, Dudhia J, Gill D, Kavulich M, Keene K, Lin H-C, Michalakes J, Rizvi S, Zhang X, Berner J, Smith K, Beezley JD, Coen JL, Mandel J, Chuang H-Y, McKee N, Slovacek T, Wolff J (2015) ARW version 3 modeling system user's guide. Technical Report, NCAR, 428 pp

Xie B, Fung JCH, Chan A, Lau A (2012) Evaluation of nonlocal and local planetary boundary-layer schemes in the WRF model. J Geophys Res Atmos 117:D12103

Zhang D, Anthes RA (1982) A high-resolution model of the planetary boundary layer–sensitivity tests and comparisons with SESAME-79 data. J Appl Meteorol 21(11):1594–1609

Zhang D-L, Zheng W-Z (2004) Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parameterizations. J Appl Meteorol Climatol 43(1):157–169

Zhang H, Pu Z, Zhang X (2013) Examination of errors in near-surface temperature and wind from WRF numerical simulations in regions of complex terrain. Weather Forecast 28(3):893–914