Boundary-Layer Meteorol (2014) 152:213–243
DOI 10.1007/s10546-014-9925-3

ARTICLE

Evaluation of the Weather Research and Forecasting Mesoscale Model for GABLS3: Impact of Boundary-Layer Schemes, Boundary Conditions and Spin-Up

Michal A. Kleczek · Gert-Jan Steeneveld · Albert A. M. Holtslag

Received: 11 July 2013 / Accepted: 18 March 2014 / Published online: 16 April 2014

Abstract We evaluated the performance of the three-dimensional Weather Research and Forecasting (WRF) mesoscale model, specifically the performance of the planetary boundary-layer (PBL) parametrizations. For this purpose, Cabauw tower observations were used, with the study extending beyond the third GEWEX Atmospheric Boundary-Layer Study (GABLS3) one-dimensional model intercomparison. The WRF model (version 3.4.1) contains 12 different PBL parametrizations, most of which have been only partially evaluated. The GABLS3 case offers a clear opportunity to evaluate model performance, focusing on time series of near-surface weather variables, radiation and surface flux budgets, vertical structure and the nighttime inertial oscillation. The model results revealed substantial differences between the PBL schemes. Generally, non-local schemes tend to produce higher temperatures and higher wind speeds than local schemes, in particular, for nighttime. The WRF model underestimates the 2-m temperature during daytime (about 2 K) and substantially underestimates it at night (about 4 K), in contrast to the previous studies where modelled 2-m temperature was overestimated. Considering the 10-m wind speed, during the night turbulent kinetic energy based schemes tend to produce lower wind speeds than other schemes. In all simulations the sensible and latent heat fluxes were well reproduced. For the net radiation and the soil heat flux we found good agreement with daytime observations but underestimations at night. Concerning the vertical profiles, the selected non-local PBL schemes underestimate the PBL depth and the low-level jet altitude at night by about 50 m, although with the correct wind speed. The latter contradicts most previous studies and can be attributed to the revised stability function in the Yonsei University PBL scheme. The local, turbulent kinetic energy based PBL schemes estimated the low-level jet altitude and strength more accurately. Compared to the observations, all model simulations show a similar structure for the potential temperature, with a consistent cold bias ($\approx$2 K) in the upper PBL. In addition to the sensitivity to the PBL schemes, we studied the sensitivity to technical features such as horizontal resolution and domain size. We found a substantial difference in the model performance for a range of 12, 18 and 24 h spin-up times, longer spin-up time decreased the

M. A. Kleczek (✉) · G.-J. Steeneveld · A. A. M. Holtslag
Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands
e-mail: mikekleczek@gmail.com

modelled wind speed bias, but it strengthened the negative temperature bias. The sensitivity of the model to the vertical resolution of the input and boundary conditions on the model performance is confirmed, and its influence appeared most significant for the non-local PBL parametrizations.

Keywords GABLS3 · Non-local schemes · Parametrizations · Spin-up time · Stable boundary layer · WRF model

1 Introduction

Despite many research efforts, the stable stratified atmospheric boundary layer (SBL), remains still a challenge for numerical weather prediction (NWP) and climate models (Holtslag 2006; Teixeira et al. 2008; Atlaskin and Vihma 2012; Holtslag et al. 2013). Depending on the stability, we may distinguish different SBL regimes that vary from weak to very stable conditions, as discussed by Mahrt (1999), Grachev et al. (2005), Banta (2008) among others. A significant factor that makes the SBL a difficult subject of investigation is the variety of processes and scales that play an important role in boundary-layer development, in particular turbulent mixing (Nieuwstadt 1984), radiative cooling (Savijärvi 2006), the low-level jet (LLJ) (Van de Wiel et al. 2010), the interaction with the land surface (Sterk et al. 2013), (orographically induced) gravity waves (Chimonas and Nappo 1989) and fog (Van der Velde et al. 2010). In the case of weak winds, turbulence cannot be maintained and the role of the other previously mentioned small-scale process becomes more important for SBL development. Another difficulty in NWP and climate models is the correct representation of the evening transition (Lothon and Lenschow 2010).

NWP and climate models are continuously developed, however, the multiplicity of processes during the diurnal cycle hampers parametrization development (LeMone et al. 2012). Currently there are many different boundary-layer parametrization schemes available. Unfortunately, many of these were designed to perform well during specific weather conditions, and not all of the parametrizations have been evaluated thoroughly. Thus such an evaluation is crucial for development of the scientific research agenda and for researchers and operational forecasters interested in weather prediction, transportation, agriculture, wind energy, air quality and climate modelling.

To improve the understanding and representation of the PBL in climate and NWP models, the Global Energy and Water Cycle Experiment (GEWEX) Atmospheric Boundary Layer Study (GABLS) has been performed (see Holtslag 2006; Holtslag et al. 2013. The third case study (GABLS3) focuses on the full diurnal cycle, interactions with the land surface and the LLJ. For the GABLS3 study, results from a single-column model (SCM) were compared to observations from the Cabauw tower (Bosveld et al. 2014a). Those results revealed the sensitivity of models to different boundary-layer parametrizations or land-use schemes (coupled to the surface). From the previous GABLS experiments it has been learnt that operational models tend to have too strong mixing at night, which leads to a boundary layer that is too deep and winds at 10 m that are too strong. The GABLS3 case study highlighted the problem of underestimating 2-m temperatures at night, while other studies reported overestimation (e.g. Svensson et al. 2011; Shin and Hong 2011).

The current study extends the previous GABLS3 research and focuses on the evaluation of a three-dimensional (3D) mesoscale model for the GABLS3 case as a complementary effort to the SCM experiment. Three-dimensional modelling represents also the large-scale processes and is in that respect more realistic than 1D modelling, as well as closer to the

operational forecasting practice. It allows for less academic cases, fully coupled to other processes. A 3D simulation gives the possibility to analyze mesoscale circulations that are lacking in a SCM analysis. Furthermore, it provides a 3D representation of spatially driven processes that are relevant to the GABLS3 case, which may provide additional insights.

Section 2 presents the model set-up and case description, and in Sect. 3 we evaluate the WRF model simulations using different PBL schemes of different complexity. Section 4 provides the results of the performed sensitivity analyses, and the final section summarizes and presents conclusions of the results from this study.

2 Case Description and Model Set-Up

2.1 Case Description

The case study covers the period from 1200 UTC 1 July to 1200 UTC 2 July 2006. It is centred at the Cabauw Experimental Site for Atmospheric Research (CESAR) located in the western part of the Netherlands (51.971°N, 4.927°E; surface elevation approximately −0.7 m). This area is relatively flat with little surface elevation variation (<1 m), with the surrounding environment dominated by grassland, fields, tree lines and scattered villages. The CESAR measurements include profiles of wind speed, wind direction, temperature and humidity at 10, 20, 40, 80, 140 and 200 m height. The energy budget components, the surface radiation, as well as soil thermal and water conditions are continuously monitored. Further description of the Cabauw tower, installed instruments and data calculation can be found in Van Ulden and Wieringa (1996), Beljaars and Bosveld (1997). Vertical profiles are available from soundings released twice a day (noon and midnight) at the KNMI weather station in De Bilt (The Netherlands, 52.1°N, 5.183°E; surface elevation approximately 1.9 m), located 25 km north-east of Cabauw. One must be aware that especially in the lowest part of the sounding this horizontal distance may provide differences with the actual profile at Cabauw.

The case selection was based on factors such as a clear sky, and absence of fog to limit the study case complexity. Another important aspect was availability of the observational data to validate the model. Additionally, a geostrophic wind forcing around 7 m s^{-1} (Baas et al. 2009) and expected moderate LLJ generated by inertial oscillations limited the possible case selection. Eventually, this specific case was selected because it was one of the few times in a 6-year period that was suitable for the GABLS3 experiment (Baas et al. 2010; Bosveld et al. 2014b).

Concerning the synoptic development, the period was characterized by a relatively stationary synoptic situation. There were no passing weather fronts over the Netherlands, an anticyclone was stationary over the north-east of Europe. At the start of the simulation period (1200 UTC) an air mass with relatively dry air was observed passing over the site. Between 0000 UTC and 0300 UTC that night, a relatively small synoptic disturbance was advected over the site causing relatively small variations in temperature, humidity and wind (Fig. 1).

Both satellite observations and surface radiation measurements indicated cloud-free conditions for the investigated period, and during the afternoon of 1 July a well-mixed convective boundary layer had developed. Based on LD40 ceilometer observations, the PBL depth was almost 2 km, which is unusually high for Cabauw. This was probably related to the relatively small amount of rainfall during the previous month (drier soil, increased sensible heat flux).

The time series of the basic meteorological parameters are presented in Fig. 1. Around 1600 UTC, the 2-m air temperature reaches its maximum of 27 °C (300 K), while rapid cooling is already observed before sunset near the surface (related to the strong surface

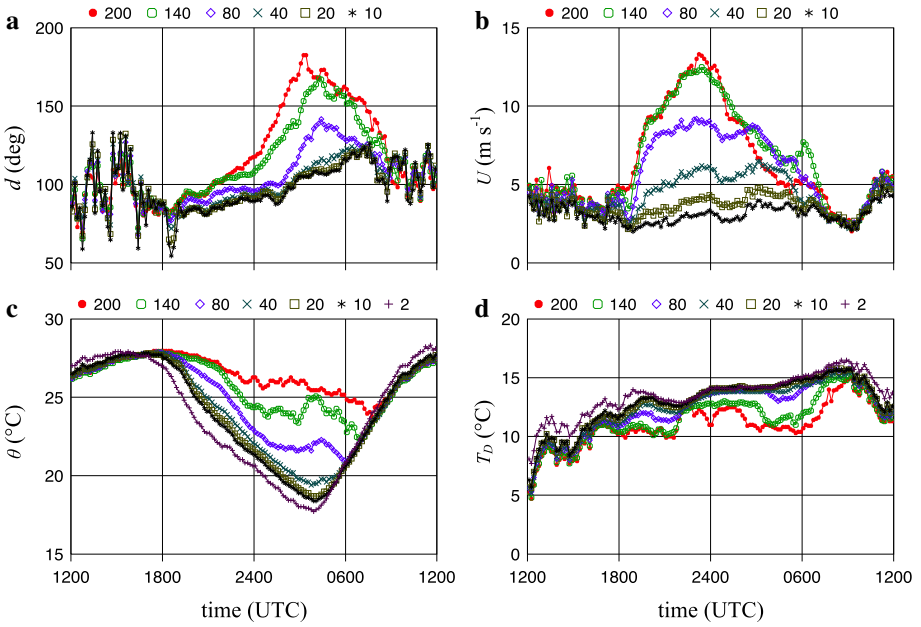

Fig. 1 Observed wind direction (**a**), wind speed (**b**), potential temperature (**c**) and dewpoint temperature (**d**) at 200, 140, 80, 20, 10 and 2 m at the Cabauw tower site for the case period, figure from Bosveld et al. (2014b)

evaporation). The potential temperature difference between the 2 and 200 m level increased up to 9 K through the night. The rapid decay of turbulence in the former mixed layer initiates a clear inertial oscillation in the wind. The LLJ reaches its maximum strength shortly before midnight with a 12 m s^{-1} wind speed. The height of the wind maximum is located around 200 m above the ground, which is confirmed by observations from different sources (e.g. wind profiler). On 2 July until 1000 UTC, the convective boundary layer grows gradually until it reaches the capping inversion at a similar level as the previous day.

2.2 Model Set-Up

In our study, we used the Advanced Research Weather forecasting model (WRF) version 3.4.1 (Skamarock and Klemp 2008), and configured with three nested domains with a grid size of 27, 9 and 3 km (Fig. 2). All model domains were centered at the Cabauw tower and have 61×61 horizontal grid points to ensure high resolution. We used 34 terrain-following (eta) levels in the vertical with more points focused in the lowest part of the PBL (15 levels in the lowest 1,000 m) to increase the vertical resolution of the SBL. The first model level was set at about 9 m. For the initial and boundary conditions, we used the European Centre for Medium-Range Weather Forecasts (ECMWF) operational analysis with 91 sigma levels, $0.5° \times 0.5°$ spatial resolution and 6-hourly timesteps. To confirm our model results, the experiment has been repeated using the NCEP Final Analysis (FNL from the Global Forecast System) boundary conditions. Their results generally supported our findings for runs with the ECMWF boundary conditions, although we observed slightly higher biases in the wind speed and the 2-m temperature (see Sect. 4). As a source for land-use categories and orography for the model simulations, we used the 24-category USGS land-use database (U.S. Geological Survey 2011), which is the default in the WRF model.

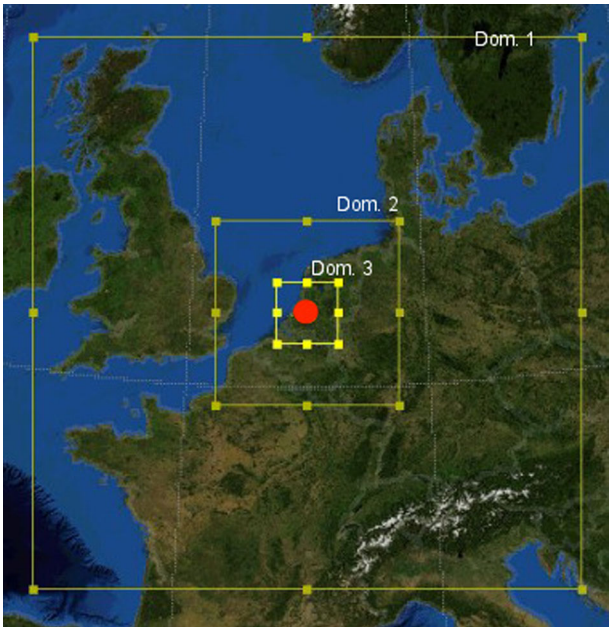

Fig. 2 Model domain arrangement, the *red dot* indicates the centre of each domain representing the Cabauw tower

In order to perform the model experiment, sufficient spin-up time should be taken into account. Jankov et al. (2007) and Skamarock and Klemp (2008) found that to prevent the occurrence of instabilities in the mesoscale NWP model, a minimum of 12 h spin-up time should be used. Unfortunately there is a lack of consensus on the spin-up time that ensures the best results for the simulation, and it likely depends on the state of the soil and quality of the model input fields. Our default spin-up time amounts to 24 h, and in Sect. 4 we describe the results of a sensitivity study to the selected spin-up time.

The physical parametrization schemes used in all model domains and default runs are the rapid radiative transfer model (RRTM, Mlawer et al. 1997), the Dudhia shortwave radiation scheme (Dudhia 1989) and the unified Noah land-surface model (Chen and Dudhia 2001; Ek et al. 2003). In order to reflect on our results, we performed a sensitivity analysis using different model resolutions, grid sizes, spin-up times and land-use schemes.

2.3 Planetary Boundary-Layer Schemes

This section briefly reviews the formulations of the PBL schemes evaluated herein. We performed the simulations with most of the available PBL schemes in WRF version 3.4.1.

The first of the selected PBL schemes is the Yonsei University (YSU) scheme, which is a non-local scheme (Eq. 1,Table 1), widely used and a default PBL scheme in the WRF model. The PBL depth in the YSU scheme is estimated by using the critical bulk Richardson number method. To allow for non-local vertical fluxes, a non-local term γ_ξ is added, and any variable ξ changes in time according to,

$$\frac{\partial \xi}{\partial t} = \frac{\partial}{\partial z}\left[K_\xi\left(\frac{\partial \xi}{\partial z} - \gamma_\xi\right) - \overline{(w'\xi')}_h\left(\frac{z}{h}\right)^3\right], \tag{1}$$

Table 1 Overview of the first-order PBL schemes used in this study

Scheme	Basic physics	Additional information	Primary reference
YSU	First-order closure scheme $K_M = \kappa \omega_s z \left(1 - \frac{z}{h}\right)^2$ $\omega_s = \left(u_*^3 + 7\kappa\omega_*^3 \frac{z}{h}\right)^{\frac{1}{3}}$ $K_H = Pr^{-1} K_M$ $h = Rib_{cr} \frac{\theta_{vs}\lvert U(h)\rvert^2}{g(\theta_v(h)-\theta_s)}$ Rib_{cr} critical bulk Richardson number θ_s appropriate temperature near the surface	Used with MM5 Monin–Obukhov surface-layer scheme Additional non-local gradient adjustment term in Eq. 1	Troen and Mahrt (1986), Hong and Pan (1996), Hong et al. (2006)
VH96	Similar to YSU except for: $h = Ri_{cr} \frac{\theta_{vs}((u_h-u_s)^2+(v_h-v_s)^2+100u_*^2)}{g(\theta_v(h)-\theta_{vs})} + z_s$ (2)		Vogelezang and Holtslag (1996)
ACM2	First-order closure scheme $K = \kappa \omega_s z \left(1 - \frac{z}{h}\right)^2$ $\omega_s = \frac{u_*}{f_{sb}\left(\frac{z_s}{\ell_m}\right)}$, $z_s = \begin{cases} \min(z, 0.1h) & \text{for unstable conditions} \\ z_s = z & \text{for stable conditions} \end{cases}$ f_{sb} stability functions from Dyer (1974) PBL height estimation similar to YSU scheme	Used in this study with MYJ surface-layer scheme	Pleim (2007a,b)

K eddy diffusion coefficient, M momentum, H heat, ℓ_m mixing length scale, h PBL height, z height, s - first model level, ω_s velocity scale, u_* friction velocity, ω_* convective velocity, κ von Karman constant, Pr Prandtl number, θ_v virtual potential temperature, U horizontal wind speed, u, v wind components in the x, y directions

where h is PBL height, z is height, $\overline{w'\xi'}$ is the vertical turbulent flux for ξ. Following Vogelezang and Holtslag (1996) we implemented modifications to the YSU parametrization in WRF (referred to as the VH96 scheme throughout this study). VH96 reformulated the bulk Richardson number Ri_b and instead of using the surface as a reference level, the first atmospheric level z_s is used (Table 1). Other specifications of the scheme remain as in the YSU PBL scheme.

The other first-order PBL scheme selected was the Asymmetric Convective Model, version 2 (ACM2). This PBL scheme is a combination of the local and non-local mixing approach. The prognostic mean variables at layer i for ξ are given by,

$$\frac{\partial \xi_i}{\partial t} = f_{conv} Mu \xi_1 - f_{conv} Md_i \xi_i + f_{conv} Md_{i+1} \xi_{i+1} \frac{\Delta z_{i+1}}{\Delta z_i} + \frac{\partial}{\partial z}\left(K_\xi (1 - f_{conv}) \frac{\partial \xi}{\partial z}\right), \tag{3}$$

where Mu is the non-local upward convective mixing rate, Md_i is the non-local downward mixing rate from layer i to $i-1$, Δz_i is the layer thickness and f_{conv} is the weighting factor that controls the degree of the local versus non-local behaviour. The current formulation extends the first version of the ACM model, described in Pleim and Chang (1992), by the local component (last term of the Eq. 3). Specifications of the eddy diffusivity in the ACM2 can be found in Table 1.

The remaining PBL schemes used in this study are the Mellor–Yamada–Janjic (MYJ), Mellor–Yamada–Nakanishi–Niino (MYNN25), Quasi-Normal Scale Elimination (QNSE) and Bougeault–Lacarrere (BOUL) schemes, which are classified as turbulent kinetic energy

(TKE) closure schemes. For all the above mentioned schemes only local transport is allowed. The TKE (e) prognostic equation is expressed by,

$$\frac{\partial e}{\partial t} = -\frac{\partial}{\partial z}\left(\overline{w'e'} + \frac{1}{\rho}\overline{w'p'}\right) - \overline{w'u'}\frac{\partial u}{\partial z} - \overline{w'v'}\frac{\partial v}{\partial z} + \frac{g}{\theta_v}\overline{w'\theta_v'} - \epsilon, \tag{4}$$

where ρ is the density, p is the pressure, g is the gravitational acceleration, θ_v is virtual potential temperature, u, v are wind components in x, y directions and ϵ represents the rate of dissipation of TKE. For TKE closure schemes we can commonly express the diffusivity as,

$$K = S_c \ell_m e^{0.5}, \tag{5}$$

where ℓ_m is the mixing length and S_c is the proportionality coefficient. The four TKE schemes selected in this study differ in the definitions of ℓ_m and S_c, whose formulations are summarized in Table 2.

Concerning the surface-layer schemes that are used together with the PBL schemes, (see also Tables 1 and 2) the simulations with the YSU and VH96 PBL schemes were performed with the modified MM5 scheme (Jiménez and Dudhia 2012). For the MYJ, ACM2 and BouLac schemes we use the Janjic Eta Monin–Obukhov surface-layer scheme (Janjic 1996, 2002). For the ACM2 scheme we also decided to use the same surface parametrization as for the simulation with the MYJ scheme, however, tests performed with the ACM2 surface-layer scheme confirmed rather consistent results. The QNSE PBL scheme simulation is performed with the QNSE surface-layer scheme (Sukoriansky 2008) and the MYNN25 simulation was performed with the MYNN surface-layer parametrization (Nakanishi 2001). A comparison, along with a more detailed description of the different surface-layer schemes, can be found in e.g. Liu et al. (2012).

3 Results

3.1 Mesoscale and Synoptic Scale

During the spin-up period, a weak sea breeze was found in the model simulation and developed between the relatively cold North Sea (100 km west of Cabauw) and the warm land. After sunset on 30 June the wind direction changed from north and north-east to east. On the day of our simulation, an easterly wind with a small north-south variation was observed. That direction was relatively consistent over the whole Netherlands in all performed simulations. For the GABLS3 case study, the Cabauw site was not affected directly by cold-air advection from the sea.

From sunset (around 2000 UTC) onwards, we noticed in the simulations a nighttime cooling effect over land, with warmer areas near the coastline. The large-scale synoptic situation in the simulations confirmed observations, that there were no passing fronts over Cabauw (for a more detailed case description, see Bosveld et al. 2014b).

3.2 Radiation Balance

All model runs reveal a correct forecast of the incoming solar radiation, about 900 W m^{-2} at noon on 1 July (not shown). During the afternoon several PBL schemes, i.e., ACM2, MYNN25, VH96 and MYJ, present lower values of incoming solar radiation with a minimum

Table 2 Overview of the higher-order PBL schemes used in this study

Scheme	Basic physics	Additional information	Primary reference
BOUL	1.5-order closure scheme $K = 0.4\ell_{\mathrm{m}}e^{0.5}$ $\ell_{\mathrm{m}} = \min(\ell_{\mathrm{up}}, \ell_{\mathrm{down}})$ ℓ_{up}, ℓ_{down} depends on distance air parcel travels vertically	Does not have own surface-layer scheme, Monin–Obukhov (Janjic) scheme used in this study	Bougeault and Lacarrere (1989)
MYJ	1.5-order closure scheme $K = \ell_{\mathrm{m}}e^{0.5}f_{\mathrm{sb}}$ $\ell_{\mathrm{m}} = \ell_0 \frac{\kappa z}{\kappa z + \ell_0}$, $\ell_0 = \alpha \frac{\int_0^h z q \,\mathrm{d}z}{\int_0^h q \,\mathrm{d}z}$ $f_{\mathrm{sb}} = f\left(\ell_{\mathrm{m}}, e, \left(\frac{\partial u}{\partial z}\right)^2, \left(\frac{\partial v}{\partial z}\right)^2, \frac{\partial \theta_{\mathrm{V}}}{\partial z}\right)$	Monin–Obukhov (Janjic) surface-layer scheme	Mellor and Yamada (1974, 1982), Janjic (1990, 2002)
MYNN25	2.5-order closure scheme as a modification of MYJ similar to MYJ except: $\frac{1}{\ell_{\mathrm{m}}} = \frac{1}{\ell_{\mathrm{s}}} + \frac{1}{\ell_t} + \frac{1}{\ell_{\mathrm{b}}}$, $\ell_{\mathrm{s}} = \begin{cases} \kappa \frac{z}{3.7} & \text{when } \frac{z}{\ell_{\mathrm{m}}} \geq 1 \\ \kappa z \left(1 + 2.7 \frac{z}{\ell_{\mathrm{m}}}\right)^{-1} & \text{when } 0 \leq \frac{z}{\ell_{\mathrm{m}}} < 1 \\ \kappa z \left(1 - 100 \frac{z}{\ell_{\mathrm{m}}}\right)^{0.2} & \text{when } \frac{z}{\ell_{\mathrm{m}}} < 0 \end{cases}$ $\ell_{\mathrm{t}} = 0.23 \frac{\int_0^\infty q z \,\mathrm{d}z}{\int_0^\infty q \,\mathrm{d}z}$ $\ell_{\mathrm{b}} \begin{cases} \frac{q}{N} & \text{when } \frac{\partial \theta_{\mathrm{V}}}{\partial z} > 0 \text{ and } \frac{z}{\ell_{\mathrm{m}}} \geq 0 \\ \left[1 + 5\left(\frac{q_{\mathrm{c}}}{\ell_{\mathrm{t}} N}\right)^{\frac{1}{2}}\right] \frac{q}{N} & \text{when } \frac{\partial \theta_{\mathrm{V}}}{\partial z} > 0 \text{ and } \frac{z}{\ell_{\mathrm{m}}} < 0 \\ \infty & \text{when } \frac{\partial \theta_{\mathrm{V}}}{\partial z} \leq 0 \end{cases}$ q_{c} turbulent velocity scale	MYNN surface-layer scheme	Nakanishi and Niino (2004)
QNSE	1.5-order closure scheme Similar to MYJ except: $K = 0.55\alpha_{\mathrm{M,h}}\ell_{\mathrm{m}}e^{0.5}$ $\alpha_{\mathrm{M,h}}$ as a functions of local gradient Ri $\frac{1}{\ell} = \frac{1}{\ell_B} + \frac{1}{\ell_N}$, $\ell_B = \frac{\kappa z}{1 + \frac{\kappa z}{\lambda}}$ Blackadar scale, $\lambda = 0.0063 \frac{u*}{f_{\mathrm{c}}}$ $\ell_N = 0.75 \frac{e^{\frac{1}{2}}}{N}$ length scale limitation due to stable stratification	QNSE surface-layer scheme Accounts for the combined effects of turbulence and waves	Sukoriansky et al. (2005, 2006), Sukoriansky (2008)

K eddy diffusion coefficient, ℓ_{m} mixing length scale, h PBL height, e TKE and $q^2 = 2e$, κ von Karman constant, z height, N the Brunt–Väisälä frequency, u_* friction velocity, θ_{V} virtual potential temperature, f_{c} Coriolis parameter, u, v wind components in x, y

around 1300 UTC for ACM2, 1600 UTC for MYNN25, 1700 UTC for VH96 and 1800 UTC for MYJ. This behaviour was the result of small clouds forecast by the mentioned schemes. This was confirmed by a non-zero cloud water vapour mixing ratio (approximately 0.2 g kg^{-1}) around 1,300 m, for the same times as the minima in the shortwave downward radiation (not shown). The difference in the cloud occurrence times may be explained by the different entrainment rates, where the ACM2 scheme shows the weakest entrainment and thus the earliest onset of clouds (not shown).

All the schemes overestimate the observed shortwave upward radiation ($S\uparrow$) of 190 W m^{-2} with the highest bias at noon of ≈15 W m^{-2} on 1 July and of ≈20 W m^{-2} at the end of the research period (noon, 2 July). The positive bias may be due to a slightly higher albedo in the model (0.23) than in reality (around 0.21). We find a similar decrease in $S\uparrow$ as was reported for the shortwave downward radiation and caused by the resolved PBL clouds. In the morning of 2 July all the schemes underestimate $S\uparrow$ at the Cabauw tower by ≈11 W m^{-2}, while after 0900 UTC all schemes reveal a positive bias in $S\uparrow$.

In this study, we also found the previously reported problems with modelling the longwave downward radiation ($L\downarrow$) (Guichard et al. 2003; Zhong et al. 2007; Van der Velde et al. 2010; Steeneveld et al. 2011). All the schemes have a negative bias of ≈20 W m^{-2}, which after midnight increases up to 40 W m^{-2}, except for small episodes in the afternoon of 1 July (Fig. 3b), during the presence of resolved clouds. This underestimation suggests that (i) we can expect a dry or cold bias in the forecast profiles (see Sect. 3.3) or (ii) there are deficiencies in the parametrization of the radiation scheme, or (iii) there is a limitation due to the limited vertical resolution, or (iv) there is a combined effect (see Sect. 4.2.2). WRF model simulations represent the general signature of the longwave upward ($L\uparrow$) radiation relatively well (Fig. 3a), which indicates that the surface temperature (T_S) is also rather well reproduced in the model. We have to note that the YSU and VH96 schemes slightly overestimate $L\uparrow$ at noon on 1 July, while the other schemes underestimate daytime values. In the morning of 2 July all the schemes underestimate $L\uparrow$, although after 0600 UTC, YSU and VH96 schemes revealed the least negative bias (due to their relatively strong mixing that raises T_S). The lowest longwave upward radiation (strongest bias) during daytime was found in the simulations with the QNSE and ACM2 schemes, with model values about 20 W m^{-2} lower than the observations (which in general confirms the results of the 1D SCM study). At night, the WRF model runs produce a negative bias (10–20 W m^{-2}) with all scheme permutations. The BOUL and QNSE simulations show the smallest bias. These nighttime biases suggest a substantial underestimation of T_S at night, where a negative bias of a maximum 5 K is found (not shown), consistent with the 1D results using the WRF model (Bosveld et al. 2014a).

The modelled net radiation (Q^*) follows the observations reasonably well (Fig. 3c). However, most of the schemes slightly overestimate Q^* during the afternoon of 1 July, and underestimate Q^* in the morning of 2 July. In the case of the simulations with the YSU and VH96 schemes, we find 30 W m^{-2} lower values at noon than are observed at Cabauw. During daytime, the simulations with higher order schemes are in better agreement with the observations. We notice a small disturbance caused by the PBL clouds, previously mentioned and described earlier, which results in a decrease in Q^*, around 1300 UTC in the case of the ACM2 scheme, 1600 UTC for the MYNN25 scheme, 1700 UTC for the VH96 scheme and 1800 UTC for the MYJ scheme. All of the permutations of schemes underestimate the observed Q^* of −45 W m^{-2} during nighttime, by ≈10 W m^{-2} for the YSU and VH96 schemes and up to 20 W m^{-2} for the TKE-based schemes. These radiation biases highlight the key aspect of a proper representation, as well as validation of the SBL (Morcrette and Geleyn 1985).

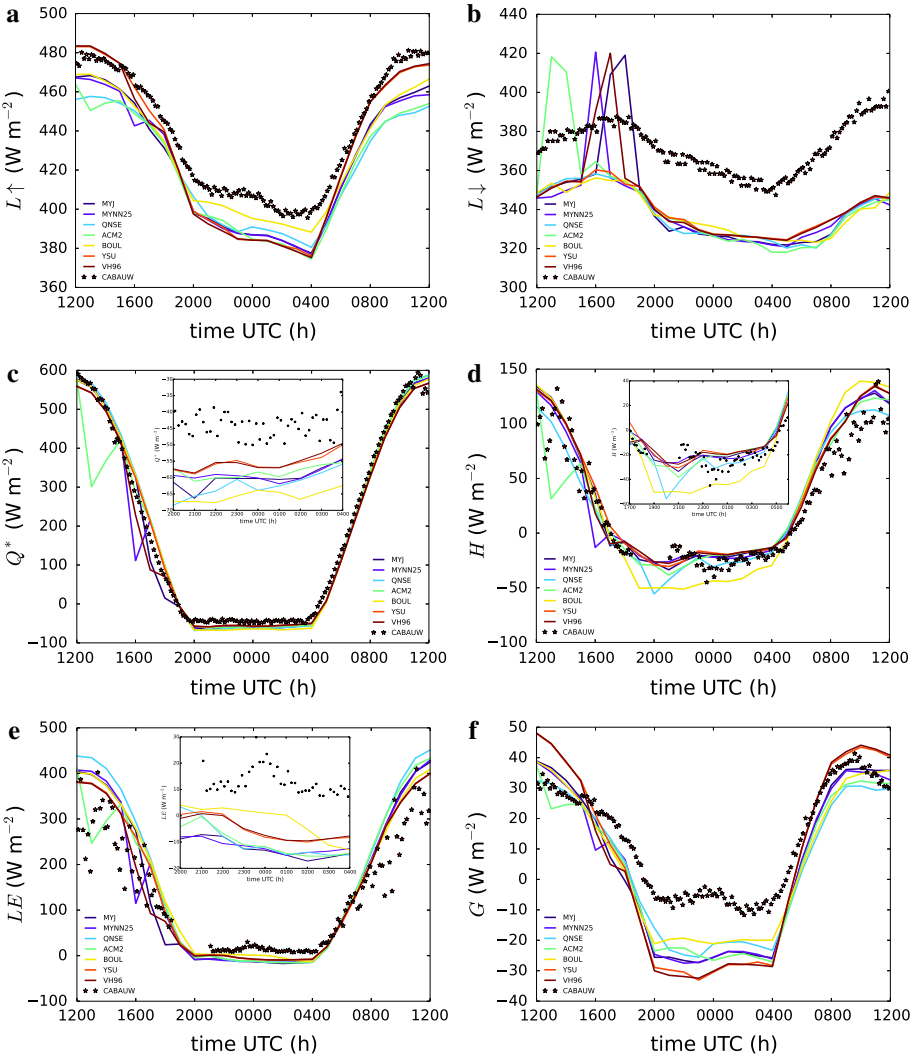

Fig. 3 Time series of observed (*stars*) and simulated (*solid lines*) longwave upward radiation (**a**), longwave downward radiation (**b**), net radiation (**c**), sensible heat flux (**d**), latent heat flux (**e**) and ground heat flux (**f**) at the Cabauw tower site, for a period of 24 h of the GABLS3 case

3.3 Turbulent Surface Fluxes

Considering the momentum flux, all the schemes substantially overestimate (on average by 0.1 m s^{-1}) the surface friction velocity (u_*, Fig. 4a). After 0400 UTC our simulations showed a substantial increase of positive bias, the strongest (roughly 0.2 m s^{-1}) around noon on 2 July. At night, all the schemes except BOUL present a similar signature of the momentum flux with a positive bias of about 0.05 m s^{-1}. Overestimation of u_* is a well-known issue that was reported in e.g. Zhang and Zheng (2004) and also in Bosveld et al. (2014a) for the GABLS3 SCM study. For the BOUL scheme the bias at night is the highest and is almost

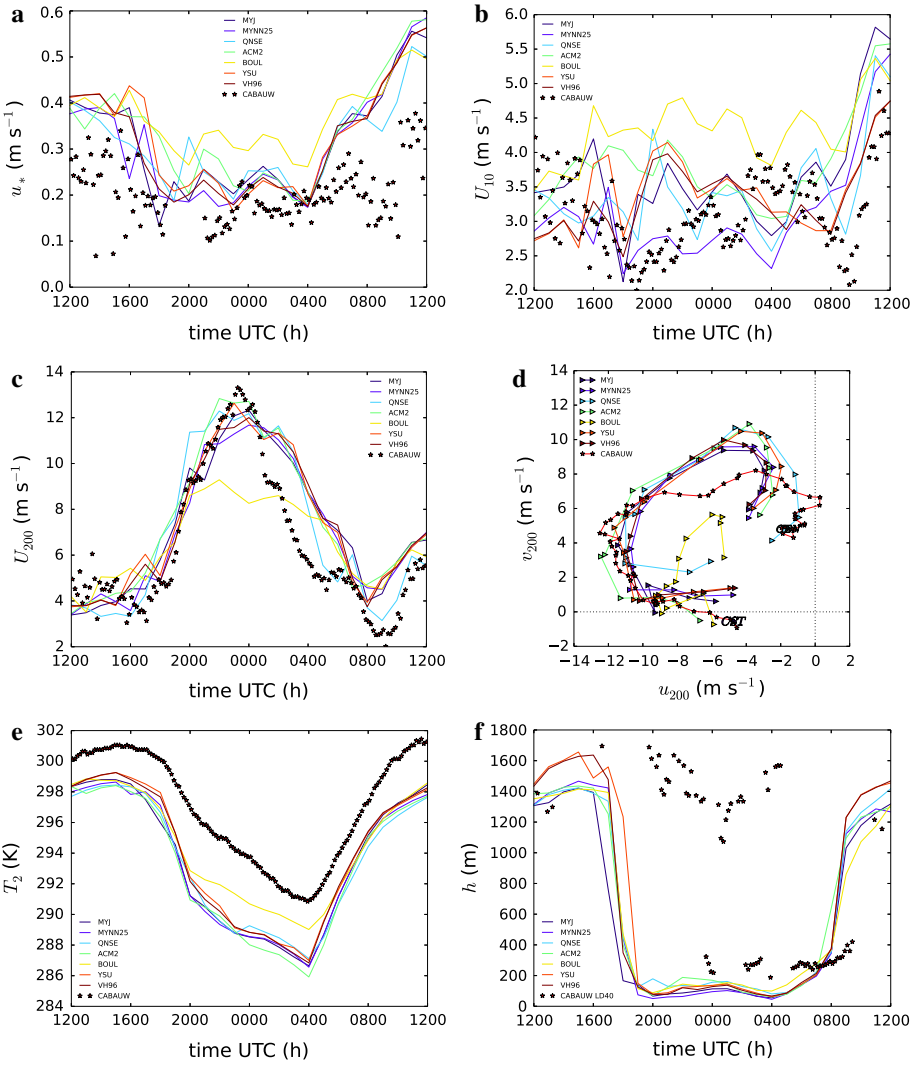

Fig. 4 Time series of observed (*stars*) and simulated (*solid lines*) friction velocity (**a**), 10-m wind speed (**b**), 200-m wind speed (**c**), hodograph for 200-m wind (**d**), temperature at 2 m (**e**) and PBL height (**f**) at the Cabauw tower site, for a period of 24 h of the GABLS3 case

0.15 m s^{-1}. The biases of u_* have a direct influence on the 10-m wind speed, which is generally overestimated by most PBL schemes (described in more detail later, Fig. 4b), as well as biases in the heat fluxes that are described below. We would like to note that we did not confirm the typical overestimation of mixing during nighttime by the YSU scheme. That may be explained by an error in the stability function in the earlier YSU scheme versions that was found by Heather Richardson and Sukanta Basu from North Carolina State University (personal communication, 2013), and corrected. Prior to the version 3.4.1, the YSU scheme contained wrong calculations of the diffusivity coefficients for the SBL. The wind profile function used in the calculation of the mixed-layer velocity scale, which in the case of stable

conditions should be equal to $\frac{u_*}{(1+5\frac{z}{L})}$, was by mistake (in the code) reduced to u_*. This correction resulted in a lower LLJ altitude and a stronger LLJ speed (described in more detail later), as the surface-atmosphere decoupling became stronger (Sect. 4.2.2).

The modelled sensible heat flux (H) captures rather well the observations from Cabauw (see Fig. 3d). All schemes substantially overestimate H in the second part of the simulation, after 0400 UTC on 2 July, where also an increase in the positive bias in the representation of u_* was noted. The highest positive bias with a magnitude of about 30 W m^{-2} was observed at noon. At night BOUL and QNSE schemes show values lower by 5–10 W m^{-2} than other schemes. The biases at night in the BOUL simulations have a direct connection with the overestimated u_*. The best model score in the representation of observed H was for the simulation with the QNSE PBL scheme.

The modelled latent heat flux (LE) shows a spread within the different PBL schemes (Fig. 3e). At night, when the observations at the Cabauw site indicate small positive values of LE (10–20 W m^{-2}), the models provide negative values, and LE ranges between -15 W m^{-2} to zero. Negative values of LE suggest that the model was able to produce such a subtle and complex process as dew. This was supported by the underestimated T_S at night and overestimated atmospheric water vapour concentration (not shown, mentioned in Sect. 4.4). However, the observations at Cabauw do not suggest any dewfall. In general the soil around Cabauw is rather moist, which supports evaporation rather than dew. Around noon the TKE-based boundary-layer schemes forecast $LE \approx 20$ W m^{-2} (and in the case of the QNSE scheme even 50 W m^{-2}) higher than the lower order schemes. The model results for LE from the first-order schemes are in the closest agreement to the observed values at the Cabauw tower.

The ground heat flux (G) comparison between the model results and the observations reveals a strong negative bias, especially at night (Fig. 3f). During nighttime, when the observed values of G oscillate around -5 W m^{-2}, all permutations of schemes present a much greater flux with values around -25 W m^{-2}. The nighttime bias was supported by a much lower T_S, up to 5 K (not shown). During daytime, the YSU and VH96 schemes present an overestimation of the observed values of G of about 20 W m^{-2} at noon, where for the other models the bias is up to 10 W m^{-2}. Daytime values produced by the QNSE scheme came closest to the observations. The modelled values of G are in fact a response to a coupled land and the PBL.

Time series of the 10-m wind speed reveal a wide spread between the parametrizations (Fig. 4b). Generally, during daytime, the YSU and VH96 schemes produced slightly lower wind speeds than other PBL schemes (despite the overestimated u_* described earlier), while around sunset their wind-speed values rapidly increased. Except for the BOUL scheme, which shows the strongest positive bias, all the TKE-based schemes results are slightly closer to the observed values during nighttime than the first-order schemes. The general signature reveals that none of the PBL schemes was able to reproduce the minimum in the 10-m wind-speed observed at Cabauw around 1900 UTC ($\approx$2 m s^{-1}); also the timing appears to be inaccurate (model reacts sooner on change in the wind strength). The wide spread between the different PBL schemes is not as vivid when we consider the 200-m wind speed (Fig. 4c). In the first part of the simulation, almost all the schemes were able to estimate the observed wind speed relatively well. Nevertheless after midnight, around 0100 UTC, all the simulations overestimate the measurements by approximately 3 m s^{-1}. The 200-m wind speed bias may be explained by an advection of momentum (Bosveld et al. 2014b), which the 3D model simulations do not realize. At night the worst score was revealed in the simulation with the BOUL PBL scheme, with a wind speed of around 9 m s^{-1}. The other schemes were able

to reproduce the observed 13 m s^{-1} more accurately. The correct timing of the minimum in the 200-m wind speed around 0900 UTC was captured only by the QNSE scheme, although with a small positive bias. Other schemes show the increase of the wind speed around 0800 UTC, much earlier than the observations. The wind direction is relatively well represented in all models (not shown); model simulations, especially at night, appear to react sooner to the wind direction change than is observed at the Cabauw site. The differences between first- and higher-order schemes results are much smaller, in comparison to the previous WRF versions, which is probably due to the YSU scheme error correction (Sect. 3.3).

Following nighttime biases in u_* and wind speed, we expected a different representation of inertial oscillations at night, visible in the hodographs (Fig. 4d). Figure 4d compares the 200-m wind components from the model with the Cabauw tower observations between 1900 UTC (approximately sunset, CST on the figures) and 0700 UTC (CEN on the figures). The initial wind values at 1900 UTC are overestimated by all the model simulations. Most of the selected PBL schemes present rather similar values, although in the case of the QNSE simulation till 2100 UTC, the v component is overestimated. Nevertheless, the QNSE scheme represents the observed morning wind-speed values best. The BOUL scheme was not able to reproduce the inertial oscillations that were responsible for the LLJ development (see Sect. 4.2.2). Additionally the vertical profiles of the BOUL scheme (see Fig. 5) suggest a PBL deeper than 200 m. Also the above mentioned 200-m wind-speed bias, which started around 0100 UTC, is clearly visible in Fig. 4d. None of the tested PBL schemes was able to capture the negative advection of v-momentum, which resulted in the overestimated wind speed.

Examining the temperature at 2 m (T_2, Fig. 4e) we confirm the findings of previous studies e.g. Steeneveld et al. (2008), Hu et al. (2010), García-Díez et al. (2013) that T_2 is underestimated during the day. At night we notice a substantial underestimation, consistent with the 1D model simulations of Bosveld et al. (2014a). Overall the general signature of the temperature evolution is well captured by all the PBL schemes. Simulations with the TKE-based boundary-layer schemes give slightly higher temperature biases, which during the day oscillate around 2 K and at night substantially increase to 4 K, than is the case with first-order schemes (except the BOUL scheme, which presents a consistent bias of about 2 K). Those biases are consistent with biases at the higher levels of the PBL. In Sect. 4.3 we discuss the results of the simulations with the thermal diffusion (TD, also known as the 5-layer slab land model) land-surface scheme in which the negative bias became smaller.

Finally we present in Fig. 4f the PBL height calculated by following the Vogelezang and Holtslag (1996) formulation (Eq. 2, see Table 1). However, by using the VH96 formulation the estimated PBL top at night is lower than expected from θ profiles. For observations we use the lidar ceilometer measurements (LD40) that have a measurement uncertainty of 75-100 m and, for the selected case, strongly fluctuate. For the GABLS3 case the LD40 captures mainly the residual-layer height and only a small sample reports the SBL top (see Sect. 5 and de Haij et al. 2006). In general, during daytime, with the YSU and VH96 schemes we find a PBL top higher by approximately 200 m than with the other schemes. At 1800 UTC the PBL top in most model simulations decreases, which is expected from models following the evening transition (transition from unstable to stable conditions). However, using the MYJ scheme we find a decrease in the PBL height 1 h earlier (1700 UTC), while with the YSU scheme the PBL top decreased 1 h later than in other simulations (1900 UTC). At night, a similar PBL height is found in all model simulations, however the simulation with the MYNN25 scheme presents the lowest PBL height. The growth of the boundary layer after sunrise (around 0400 UTC) was similar in most schemes. The strong increase of the PBL height in the morning was captured 1 h earlier (around 0800 UTC) than in the observations.

3.4 Vertical Profiles

Concerning the vertical profiles at the beginning of the simulation, i.e. noon 1 July (not shown), all schemes were able to reproduce the general signature of the potential temperature (θ) relatively well. Nevertheless, most schemes underestimate θ by roughly 1 K, except the simulations with the YSU and VH96 schemes that correspond well with the observations in the lowest 1,400 m. Those differences are much greater at noon on 2 July (not shown), where models underestimate the sounding measurements by $\approx$2 K in the lower part of PBL. Inspecting the modelled θ profiles for the midnight on 2 July (Fig. 5a), we find that all the schemes above 200 m show a substantial cold bias of nearly 2 K, compared to the sounding from De Bilt. That underestimation may be related to the negative longwave downward radiation bias (Sect. 3.2). The TKE-based schemes present a lower $\theta \approx 1$ K in the lower part of the PBL compared to YSU and VH96 simulations, which is consistent with the literature (e.g. Hu et al. 2010; Shin and Hong 2011). Around 250 m, the highest negative bias is roughly 4 K. In the BOUL simulation we find an inversion around 200 m with lower θ. The SBL profile for that scheme is less stratified, which explains the insufficient magnitude of inertial oscillations described above (Fig. 4d). The overall signature of θ is well captured in most PBL schemes.

The simulated specific humidity (q_V) in general overestimates the observations. The strongest positive bias was observed in the profiles from noon 1 July, where PBL schemes overestimated q_V by around 4 g kg^{-1} (not shown). The difference between the observations and model results was also reported in Bosveld et al. (2014a), although the origin of this bias is unknown. At 1200 UTC on 2 July 2006, the bias is smaller compared to the results from the first day and amounts to $\approx$2 g kg^{-1}. The positive q_V bias is greater in the TKE-based boundary-layer schemes, especially near the surface. Vertical q_V profiles for 0000 UTC (Fig. 5b) show that the overall structure of q_V is better captured by the YSU and VH96 schemes. Also in the upper part of the atmosphere, the YSU and VH96 schemes remain closest to the observations, while the higher order schemes have a positive bias of 1 g kg^{-1}.

In the vertical profiles of the wind speed at noon a strong bias, with respect to the sounding from De Bilt, was found (not shown). However, we have to note that at 1200 UTC on 2 July the general signature of the wind speed corresponds relatively well with the additional wind-profiler observations from Cabauw. As mentioned in Sect. 2.1, the distance of 25 km between De Bilt and Cabauw may be a reason for the different representation of the vertical profiles. At midnight the wind speed is estimated better by the TKE-based schemes and seems to correspond well the data captured by the sounding in De Bilt and the Cabauw tower (Fig. 5c). The MYJ, QNSE and ACM2 schemes were able to reproduce the LLJ of about 12.4 m s^{-1} at 200 m. The YSU scheme underestimates the altitude at which the maximum wind speed occurs and suggests this at 150 m, although with the correct wind speed (in contradiction to the results reported with older versions of the WRF model). The lower altitude and stronger LLJ in the YSU scheme may be explained by too strong a stability (decoupling), caused by the correction in the velocity scale, mentioned in Sect. 3.3. In previous versions of the WRF model, the YSU scheme was known as a first-order scheme with too strong a nighttime mixing, causing e.g. an overestimation of the LLJ altitude. Since version 3.4.1, the corrected velocity scale limited the strong mixing, although our results indicate that the SBL became apparently too stable. The lowest skill was presented by the BOUL scheme, which was not able to reproduce the LLJ due to the lack of inertial oscillations necessary for its development, as previously discussed.

The wind direction profiles in general show rather small variation between the different simulations. At 1200 UTC on both 1 and 2 July, a strong bias in all permutations of schemes

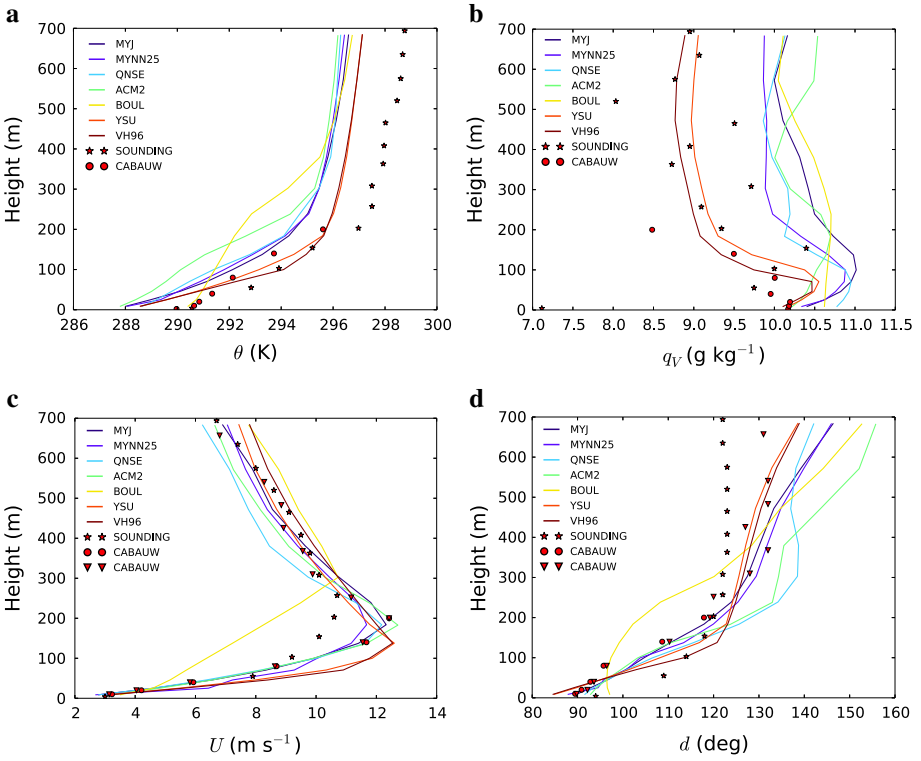

Fig. 5 Vertical profile for 0000 UTC 2 July 2006, of observed at the Cabauw tower (*circles* and *triangles*), De Bilt (*stars*) and simulated (*solid lines*) potential temperature (**a**), water vapour mixing ratio (**b**), wind speed (**c**) and wind direction (**d**)

was found around inversion level (roughly 1,500 m, not shown). Additionally stronger variability between simulations was revealed around that altitude. At midnight, the near-surface wind direction is relatively well captured, although in the upper part of the PBL the model simulations present a bias of approximately 20° of veering (Fig. 5d).

4 Sensitivity Study

In the following section we discuss the results from a sensitivity analysis regarding a few aspects that may influence the model performance. A more detailed description of the performed simulations can be found in Table 3. The first sensitivity study concerns the model domain size. The selected domain size is an important aspect as discussed in e.g. Seth and Rojas (2003) and Vannitsem and Chomé (2005). A smaller model domain may lead to limitations that prevent the model from developing detailed mesoscale processes.

The second test concerned the horizontal resolution of the model. As stated in studies such as Mass et al. (2002), Gego et al. (2005), Carvalho et al. (2012) the horizontal resolution is one of the factors that may affect model performance. An increase of the horizontal resolution results in more precise, better resolved, small-scale processes, although also increases the model numerical costs.

Table 3 Sensitivity study description and labelling

Test	Specification	Performed runs	Labels and figures
Model domain	Different size of the model domain	31×31 grids of 27×27 km 61×61 grids of 27×27 km (reference) 91×91 grids of 27×27 km	PBL_Hz × Gs, Fig. 6
Horizontal resolution	Different horizontal resolution used in the model with similar domain size	31×31 grids of 54×54 km 61×61 grids of 27×27 km (reference) 91×91 grids of 18×18 km	PBL_Hz × Gs, Fig. 7
Vertical resolution	Higher vertical resolution in the model. Lower vertical resolution of ECMWF input files	ECMWF input files with 91 vertical levels and 34 vertical sigma level in the model (reference) ECMWF input files with 91 vertical levels and 45 vertical sigma level in the model ECMWF input files with 61 vertical levels and 34 vertical sigma level in the model	PBL_X ECMWF_Y sigma, X,Y respond to amount of vertical levels in ECMWF input files and model respectively, Fig. 8
Land-use scheme	Impact of land-use scheme	Run with NOAH land-use scheme (reference) Run with TD land-use scheme	PBL_LU, where *LU* land-use scheme, Fig. 9
Spin-up time	Different spin-up time	Simulation with 24 h spin-up time (reference) Simulation with 18 h spin-up time Simulation with 12 h spin-up time All mentioned spin-up time is a prior to presented results in this study	PBL_s X, X Amount of spin-up time, Fig. 10

Universal terminology used: *PBL* used PBL scheme, *Hz* number of horizontal grid boxes (in all nested domains both north-south, east-west direction, which are equal), *Gs* outer domain grid size (nested domain has factor 3 grid size decrease)

Studies by Chou (2011) and Carvalho et al. (2012) showed that a higher vertical resolution can improve the accuracy of NWP models. Along with the higher vertical resolution in the model we performed a simulation with a lower vertical resolution in the ECMWF input files to investigate its influence on the model performance. This was the third sensitivity test.

As stated by e.g. Jin et al. (2010), temperature especially is very sensitive to the selected land-surface scheme. In this study following the biases in T_2 (see Sect. 3.3, Fig. 4e), we performed a simulation with the relatively simple thermal diffusion (TD) scheme, scale in search of a better estimation of the 2-m air temperature in the model.

NWP models require a spin-up time to develop proper mesoscale and large-scale circulations (Skamarock 2004; Jankov et al. 2007; Yang et al. 2011), if they are not operated run in a data-assimilation cycle. Finally in this study we used 24, 18 and 12 h of spin-up time for which the results are discussed below (fourth sensitivity test).

Results were compared to the observations from the Cabauw tower, De Bilt and the reference runs of the YSU and MYJ schemes presented in the previous section (the labels contain 'ref' respectively). We selected the YSU and MYJ schemes from all PBL schemes, since those are most commonly used and relatively less complex. The applied labels are consistent with the discussed changes throughout the sections (see Table 3). Concerning the vertical profiles we focused on the representation of the LLJ as one of the goals of the

GABLS3 exercise. Because of the limited amount of available soundings, we compared vertical profiles for midnight on 2 July.

4.1 Model Domain

In this section we discuss the model performance depending on the model domain size. As mentioned above and in Table 3, all the simulations were performed with three nested domains for which the grid sizes decreased by a factor of three (27 km, 9 km and 3 km respectively). The presented results originate from the innermost domain.

We find that the general structure of the surface variables is not influenced much by using different domain sizes. For the shortwave radiation similar results are found in reference to both YSU and MYJ permutations. The negative biases (20–40 W m^{-2}) within the longwave downward radiation discussed in Sect. 3.2 did not improve. However, on the afternoon of 1 July until sunset (around 1900 UTC), the simulations with the smallest domain present an even larger bias (about 30 W m^{-2}) compared to the other runs at that time. We also find similar biases as previously reported in the longwave upward radiation. until 0400 UTC there is no systematic difference between the different permutations, although both PBL schemes show lower $L\uparrow$ in the simulations with a smaller domain. After sunrise (around 0400 UTC) we find that the simulations with a smaller domain size produce a stronger longwave upward radiation and in the case of the YSU scheme with the smallest domain size the bias at noon, of approximately 10 W m^{-2} (reported in Sect. 3.2), disappears. This tendency is noted in the surface temperature (not shown). Until 0400 UTC $L\uparrow$ is generally lower in the simulations with the smallest domain, whereas afterwards it is opposite.

In the representation of u_*, we find that the model runs with a smaller domain size have the tendency to decrease the momentum flux (Fig. 6b). In the second part of the simulation, e.g. at noon 2 July, the positive bias of around 0.2 m s^{-1} decreased by roughly 0.1 m s^{-1}. At night, the lowest u_* is found in the runs with the YSU scheme, in combination with the smallest domain. There is no systematic behaviour in the representation of H and LE (not shown), although the variation between the simulations with different domain sizes is visible. In the second part of the simulation, in the morning of 2 July, we find smaller H (by about 30 W m^{-2}) in the simulations with the smallest domain, than in the other performed runs (Fig. 6a). Additionally, we find that in the second part of the simulation, around noon, the simulations with the smallest domain produce a stronger LE. Furthermore the representation of the soil heat flux also remains fairly similar. After the sunrise we notice a gentle increase of G (not shown) in the simulation with the smaller domain size, which is consistent with the changes in $L\uparrow$.

A rather large variation between the model runs appears in the 10-m wind speed in comparison to the Cabauw observations (Fig. 6c). After midnight, generally lower wind speeds are produced for the smaller domain sizes. This trend is partially carried into the 200-m wind speed (Fig. 6d). Around 0000 UTC the simulations with the smallest domain were not able to reproduce the observed wind maximum. The YSU and MYJ simulations with the largest domain set-up model sudden decrease in the wind speed around 0200 UTC, for which the origin was not identified and remains a challenge for future study. It is worth noting that despite their underestimation of the wind speed around 0000 UTC, the runs with the smallest domain come closest to the observed values in the second part of the simulation (after 0000 UTC). By comparing the hodographs for 200 m (Fig. 6e) we find a weaker and substantially underestimated inertial oscillations in the simulations with the smallest domain size. That difference is most likely the origin of the previously mentioned underestimation of the wind-speed maxima around 0000 UTC. In the simulations with the largest domain size, we find

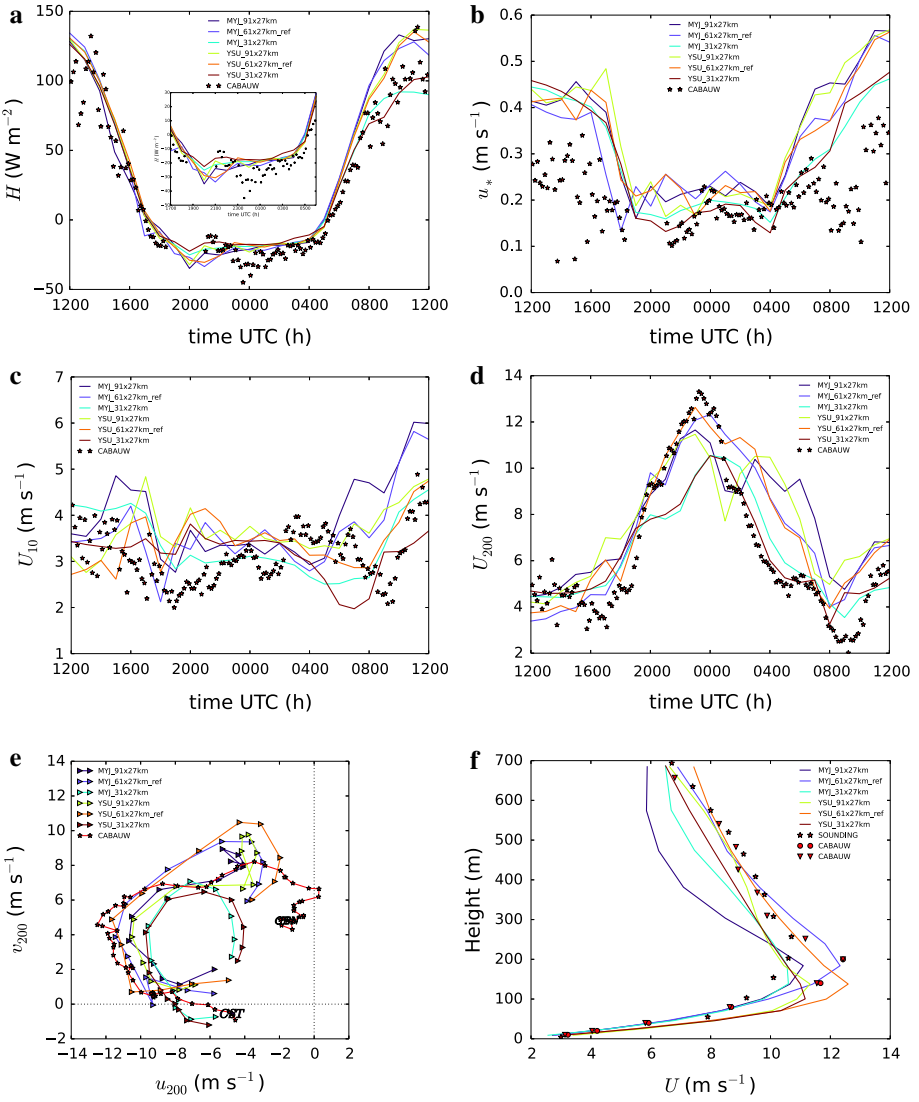

Fig. 6 Domain size sensitivity analysis. Time series of observed (*stars*) and simulated (*solid lines*) sensible heat flux (**a**), friction velocity (**b**), 10-m wind speed (**c**), 200-m wind speed (**d**) and hodograph for 200-m wind (**e**) for a period of 24 h of the GABLS3 case, along with wind speed vertical profile (**f**) for 0000 UTC 2 July 2006 with observations at Cabauw (*circles* and *triangles*) and De Bilt (*stars*)

a very specific wind vector turning around 0200 UTC, of which the direct effects were not found in other variables. In the 2-m temperature profiles we find that the simulations with the smallest domain showed the least negative bias in the second part of the simulation (1 K greater than in the other model simulations). However, in the afternoon of the first day, a smaller domain led to a lower 2-m temperature.

Considering the θ profiles we notice that the simulations with the smallest domain result in a warmer atmosphere indicating that the cold bias in the default run is created within the WRF

model itself. The lowest wind-speed biases were found in the reference runs. Simulations with a domain size different than the reference run, underestimated the wind-speed and forecast a lower altitude of the LLJ (see Fig. 6f).

4.2 Resolution

In the following section we investigate the importance of the horizontal resolution in the WRF model; a description of the performed simulations can be found in Table 3. In general we selected two additional (to the reference run) horizontal resolution values keeping the domain size similar. The second part of this section describes the results of the vertical resolution sensitivity study. By increasing the WRF model vertical resolution from 34 to 45 eta levels (with more levels in the PBL, the lowest 1 km) we studied how a higher vertical resolution affects the model simulations. Additionally we performed a simulation where instead of input files with a high resolution (91 vertical levels), as used in the other simulations, the ECMWF input files with a coarser resolution (61 vertical levels) were applied.

4.2.1 Horizontal Resolution

Considering the radiation balance components, we find that the shortwave radiation is not directly affected by a horizontal resolution change in the model. This is consistent, since radiation schemes are expected to be most sensitive to the vertical resolution. We notice that, in the case of the simulations with the coarsest horizontal resolution, the clouds described above in Sect. 3.2 also appear in the simulation with the YSU scheme. The appearance of clouds in the YSU results in a sudden decrease of incoming shortwave radiation, with a minimum around 1700 UTC. The disturbance caused by the clouds appearance is visible in the other radiation components as well. However, these additional clouds have negligible effect on the following night. For $L\uparrow$ and $L\downarrow$ we do not find any systematic behaviour. Additionally, the difference between the simulations with different horizontal resolution is rather negligible (Fig. 7a).

Regarding the momentum flux, we find that a higher resolution results in a slightly higher momentum flux during the day (see Fig. 7b). Additionally, we do not find a direct influence of the horizontal resolution on the modelled sensible and latent heat fluxes. However, around 1800 UTC we find in the representation of H lower values in the simulation with the coarsest horizontal resolution. In the simulation with the YSU PBL scheme and the coarsest horizontal resolution, we find a local minimum in LE around 1700 UTC. A similar bias was also found in the temperature at the surface (not shown), which may be caused by the resolved clouds. In the 10-m wind speed we find a substantial variation between the different simulations (Fig. 7c). After the sunrise the simulations with a higher horizontal resolution produce higher wind speeds. Concerning the 200-m wind speed the main difference is found around 0800 UTC where the timing and magnitude of the wind speed show a rather strong sensitivity to the selected horizontal resolution (Fig. 7d). The simulations with a lower horizontal resolution were able to more accurately reproduce the wind-speed minimum observed at Cabauw around 0800 UTC. Furthermore in the hodograph for the 200-m wind speed (Fig. 7e) we note a similar general signature of the wind vectors for all runs. Although the sensitivity of the wind speed on the horizontal resolution is visible, a clear, substantial signature is absent. For the 2-m temperature the biases between the different model simulations are rather negligible, however with a small sensitivity (not shown).

In the vertical structure of θ at midnight of 1 July we do not find any significant differences between the simulations with different horizontal resolutions. Specific humidity

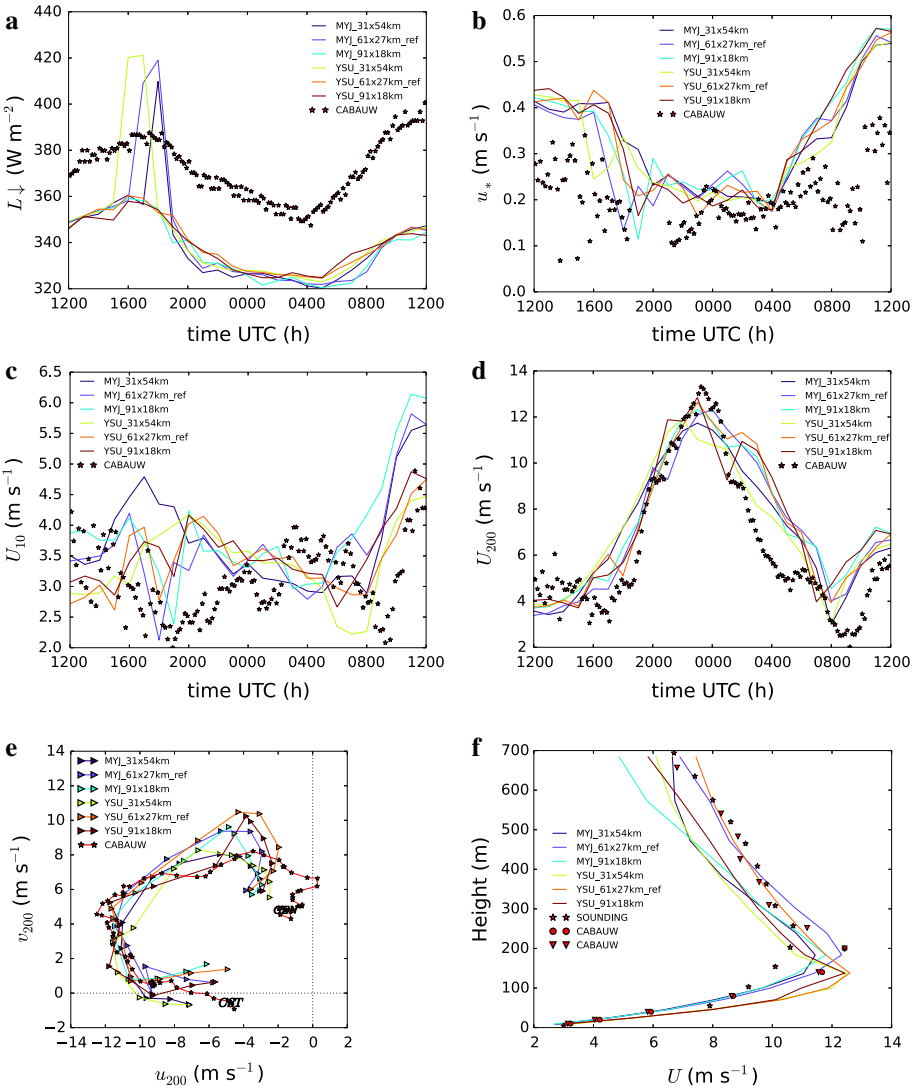

Fig. 7 Model horizontal resolution sensitivity analysis. Time series of observed (*stars*) and simulated (*solid lines*) longwave downward radiation (**a**), friction velocity (**b**), 10-m wind speed (**c**), 200-m wind speed (**d**) and hodograph for 200-m wind (**e**) for a period of 24 h of the GABLS3 case, along with wind speed vertical profile (**f**) for 0000 UTC 2 July 2006 with observations at Cabauw (*circles* and *triangles*) and De Bilt (*stars*)

profiles likewise did not reveal a systematic behaviour or sensitivity. The main difference is found in the vertical wind-speed profiles (Fig. 7f). For our simulations, the reference run (Table 3, 61×61 of 27 km) was in best agreement with the observations. Furthermore, the LLJ magnitude becomes weaker in the simulations with the MYJ scheme and both coarser and finer resolution set-ups. Possible simulation with coarser resolution omits some small-scale processes, additionally thermal contrasts are smoothed. From simulations with finer resolution we expect greater wind-profile details resolved by the model, which leads to addi-

tional biases. This is supported by the initial states of the SBL, Fig. 7e, where each simulation has a different initial wind vector.

4.2.2 Vertical Resolution

This section describes the results of two sensitivity tests both related to the vertical resolution. In one simulation we increased the number of vertical levels in the model, in the other simulation we used the ECMWF input files with a lower vertical resolution than in the reference simulation. More details are described above and in Table 3.

Regarding the radiation components, we find that for the shortwave radiation neither the vertical resolution of the ECMWF input nor the WRF model vertical model resolution has a substantial influence on the model performance. In the afternoon of the first day, the YSU scheme with 45 sigma levels captured clouds that were not visible in a reference run (see Sect. 3.2), though this was not the case for the MYJ simulation with increased resolution. The longwave downward radiation still shows a consistent negative bias of 20–40 W m^{-2} in all simulations. Apparently neither horizontal nor vertical resolutions are the critical cause of the $L\downarrow$ bias. When looking at the modelled values of longwave upward radiation we notice that the runs with a higher model vertical resolution present a slightly lower $L\uparrow$ than in the other simulations.

Considering the momentum flux (see Fig. 8a) we find that a higher vertical resolution in the model results in a smaller momentum flux at night. This effect is stronger and clearer in the case of the MYJ scheme (1.5-order, TKE-based). The simulations with coarser input files did not reveal any systematic behaviour for this case. Nevertheless, especially at night, we do find the WRF model to be sensitive to the input file resolution. Simulations with different vertical resolutions, and different resolution of the input files, do not have a substantial influence on LE. In the case of H (Fig. 8b), during daytime the simulations with a coarser resolution of the input files produce a slightly stronger sensible heat flux. During nighttime a higher vertical resolution in the model results in a smaller H due to stronger stratification. The 10-m wind speed reveals a substantial variation between the different simulations, which indicates a strong sensitivity to the model vertical resolution and the selection of the input files (Fig. 8c). At night, the simulation with the MYJ PBL scheme and 45 sigma levels produces the lowest wind speed. Generally the simulation with the MYJ scheme seems to be more sensitive to the selection of the input file resolution and the amount of vertical levels in the WRF model. In the representation of the 200-m wind speed we do not find such a strong sensitivity as at lower levels (Fig. 8d). The hodographs for 200-m wind vectors (Fig. 8e) reveal a different direction of the oscillation around sunset. The general signature of the inertial oscillations between the different simulations is rather similar. For the 2-m temperature, a substantial difference between the simulations was not found; at night slightly lower values are produced in the simulations with a higher model vertical resolution.

Regarding the θ profile at 0000 UTC (not shown), we find that by using the coarser resolution ECMWF input files, the cold bias around 200 m in YSU increases, and decreases in the case of MYJ; values of θ in other parts of the PBL between the different simulations are fairly similar. As mentioned in the previous paragraph the MYJ PBL scheme reveals a stronger sensitivity to the input file resolution in the representation of the LLJ (Fig. 8f). By using the coarser ECMWF input files the altitude becomes lower (by 50 m) and strength of the LLJ becomes weaker. The YSU scheme seems less sensitive to the input file resolution, however the LLJ does become slightly stronger. The increase of the model vertical resolution did not affect the representation of the LLJ in the MYJ simulations significantly, however, the vertical wind profile at 0000 UTC became more stable. For the YSU PBL scheme we find a

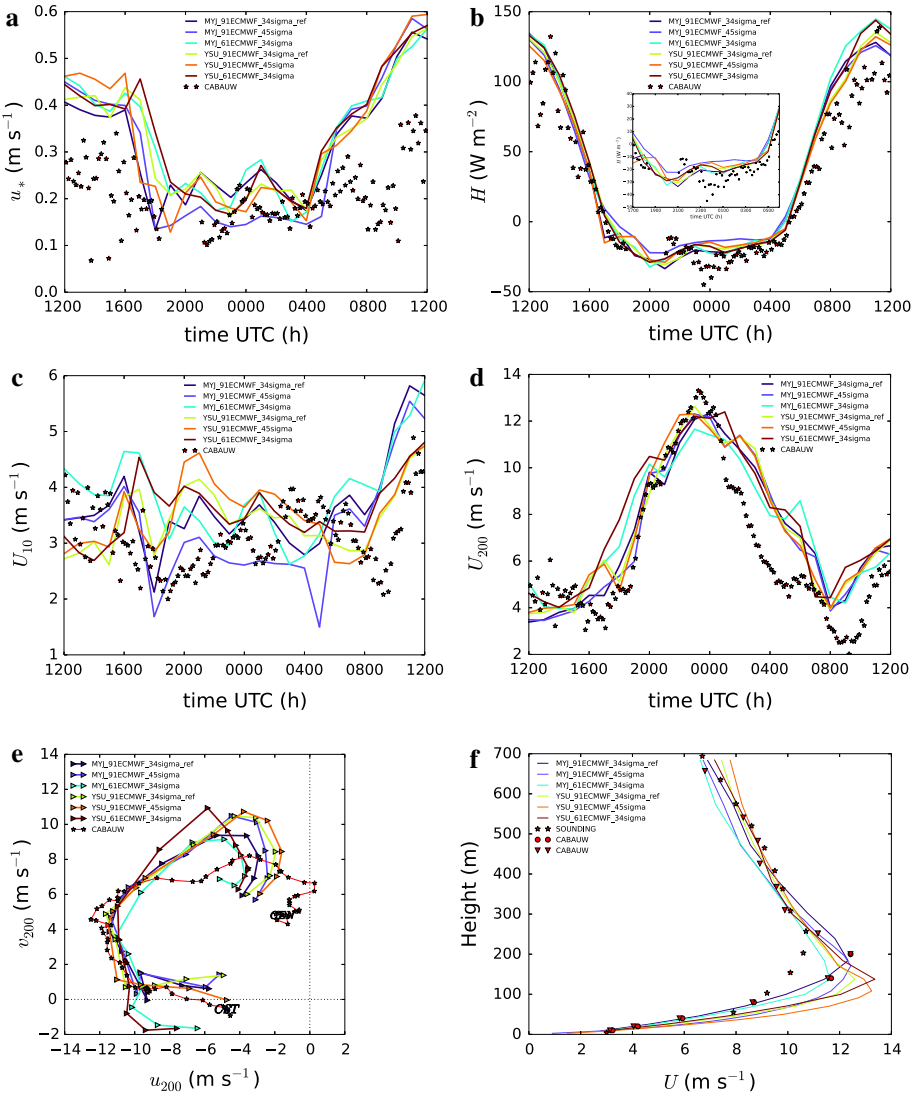

Fig. 8 Model and ECMWF input files vertical resolution sensitivity analysis. Time series of observed (*stars*) and simulated (*solid lines*) friction velocity (**a**), sensible heat flux (**b**), 10-m wind speed (**c**), 200-m wind speed (**d**) and hodograph for 200-m wind (**e**) for a period of 24 h of the GABLS3 case, along with wind speed vertical profile (**f**) for 0000 UTC 2 July 2006 with observations at Cabauw (*circles* and *triangles*) and De Bilt (*stars*)

minor strengthening of the LLJ (of the same magnitude as when decreasing the resolution of input files) and it is located at a lower altitude, along with a more stable vertical wind profile.

4.3 Land-Use Scheme

Following the biases in the 2-m temperature we performed a sensitivity study for different land-surface parametrizations. In our research we used the most commonly applied Unified

Noah land-surface model (NOAH, Ek et al. 2003), which was used as a reference case in this section, and a less complex thermal diffusion scheme (Table 3).

Considering the radiation balance components, both reference runs produce fairly similar results for the shortwave downward radiation, disregarding the land-surface schemes used. Around 0500 UTC of 2 July in the simulation with the MYJ_TD configuration we find a small disturbance (caused by resolved fog), which is also visible in the other radiation components. The modelled $S\uparrow$ behaved similar to the described shortwave downward radiation. Concerning $L\downarrow$, we note fairly analogous results regarding the selected land-use scheme (higher θ is compensated by lower q_V therefore the $L\downarrow$ bias does not decrease, see also below). Around 0500 UTC, the simulation with the MYJ_TD configuration indicates a local maximum of the longwave downward radiation, as mentioned above, due to fog (confirmed by non-zero cloud water vapour mixing ratio near the ground, not shown). The bias in $L\uparrow$ described in Sect. 3.2 becomes weaker in the simulations with the TD scheme, which is especially visible at night (not shown). The bias reduction can be explained by the higher values of the surface temperature in the simulation with the TD scheme. Nevertheless, after 0400 UTC, the simulation with the YSU_NOAH configuration reproduces $L\uparrow$ best, while other simulations reveal a similar, stronger negative bias.

The representation of u_* remains rather similar, regardless of the selected land-use scheme. During nighttime, in the simulation with the YSU_TD configuration, we find a slightly greater momentum flux that indicates stronger mixing. For the latent heat flux representation, the simulations with the TD scheme reveal the weakest bias, smaller than for the runs with the NOAH scheme, compared to the observations from Cabauw. At noon the bias between the simulations with the different land-use schemes is the strongest (see Fig. 9a). The simulation with the YSU_TD configuration during the day reveals a strong, positive bias in the H representation, up to 90 W m^{-2} at noon (Fig. 9b). At night the simulation with the YSU_TD configuration predicts the lowest sensible heat flux. Around 2000 UTC this simulation reached about -55 W m^{-2}, where in other simulations H oscillates around -25 to -15 W m^{-2}. The lower values at night may be related to the stronger momentum flux for nighttime in the simulation with the YSU_TD configuration. The difference in the PBL height (Fig. 9c) at night is rather negligible. During daytime the YSU scheme reveals a stronger dependence on the selected land-use scheme, by estimating a 200 m higher PBL top in the simulation with the TD scheme than with the NOAH scheme. For the 10-m wind speed we find a rather small sensitivity on the land-use scheme selection (Fig. 9d). At 200 m the general signature of the wind speed becomes more clear, with slightly higher wind speeds in the YSU_TD configuration, related to the slightly stronger inertial oscillations (not shown). In the simulation with the MYJ_TD configuration we find the opposite effect, i.e. lower wind speeds, confirmed by slightly weaker inertial oscillations. The 2-m temperature bias significantly decreased in the simulations with the 'simple' thermal diffusion land-use scheme (Fig. 9e). The YSU PBL scheme shows an improvement during the whole study period, while MYJ improves especially at night.

Considering the vertical structures we find that the simulation with the YSU_TD configuration results in a much warmer and drier atmosphere at the higher levels (above 200 m). The simulation with the YSU_TD configuration was able to reproduce the θ sounding from De Bilt most accurately. As briefly mentioned in Bosveld et al. (2014b), models in general required higher initial specific humidity to correctly estimate the other variables. This is confirmed in this study where the bias of specific humidity (not shown), around noon 1 July is present. Nevertheless the YSU_TD configuration reveals weaker positive bias, which also results in a drier atmosphere at higher levels. In the vertical representation of the wind speed we find that the altitude of the estimated LLJ becomes higher in the simulations with the TD scheme

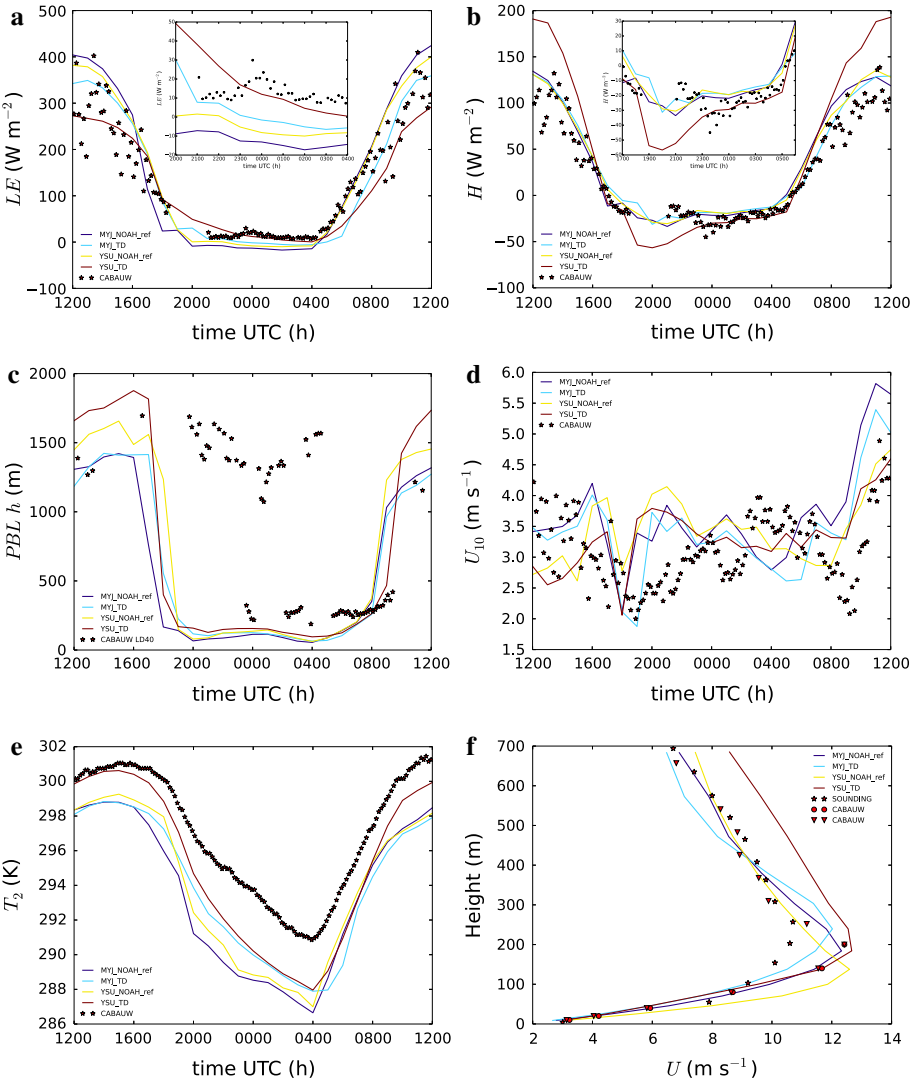

Fig. 9 Land-surface scheme sensitivity analysis. Time series of observed (*stars*) and simulated (*solid lines*) latent heat flux (**a**), sensible heat flux (**b**), PBL height (**c**), 10-m wind speed (**d**) and temperature at 2 m (**e**) for a period of 24 h of the GABLS3 case, along with wind speed vertical profile (**f**) for 0000 UTC 2 July 2006 with observations at Cabauw (*circles* and *triangles*) and De Bilt (*stars*)

(Fig. 9f). As a result the wind-speed maxima and altitudes are much more accurately represented in the YSU_TD configuration compared to the YSU_NOAH configuration. Above 200 m the simulation with the YSU_TD configuration overestimates the wind speed in comparison to other simulations. Additionally, the simulation with the MYJ_TD configuration overestimates the altitude and slightly weakens the strength of the LLJ found in the simulation with the NOAH scheme. The best model score regarding the LLJ representation is achieved by the simulation with the MYJ_NOAH configuration.

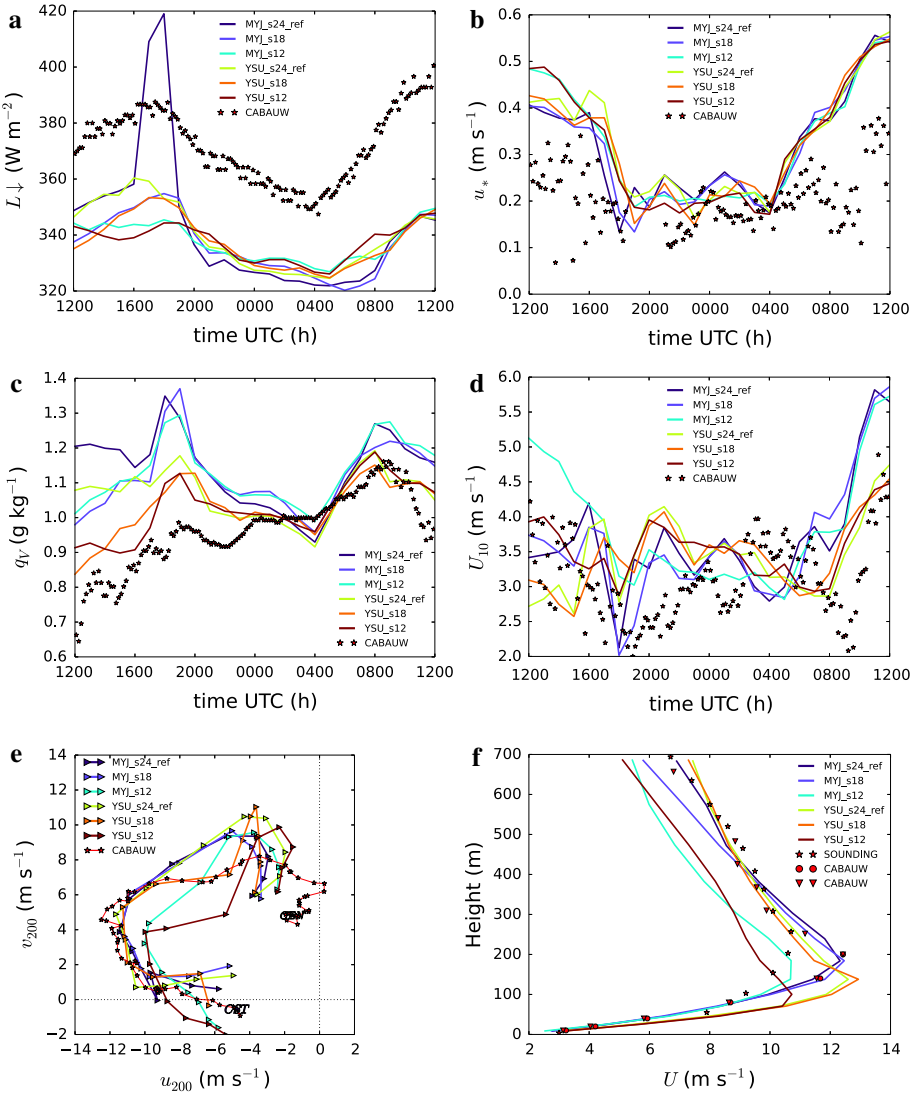

Fig. 10 Spin-up time sensitivity analysis. Time series of observed (*stars*) and simulated (*solid lines*) longwave downward radiation (**a**), friction velocity (**b**), water vapour mixing ratio (**c**), 10-m wind speed (**d**) and hodograph for 200-m wind (**e**) for a period of 24 h of the GABLS3 case, along with wind speed vertical profile (**f**) for 0000 UTC 2 July 2006 with observations at Cabauw (*circles* and *triangles*) and De Bilt (*stars*)

4.4 Spin-Up Time

To justify our choice for the spin-up time (24 h) prior to the GABLS3 period, we performed additional simulations with 18 h and 12 h as spin-up times (see Table 3 for more details). Our simulations do not reveal a significant influence on the representation of the shortwave radiation components. In the longwave downward radiation, as previously discussed in Sect. 3.2, the bias of 20–40 W m^{-2} does not improve (see Fig. 10a). Nevertheless, the modelled

$L\downarrow$ is sensitive to the spin-up time length, with the biggest variation between the simulations especially in the first few hours of the study period. This may be due to the amount of time the model requires to develop proper meso and large scale circulations. The modelled longwave upward radiation remains fairly similar regardless of the length of spin-up time.

The modelled u_* is similarly estimated in all the simulations (Fig. 10b). Around noon on the first day we distinguish slightly higher values in the simulations with the shortest spin-up time (about 0.1 m s^{-1} higher) than in the other permutations. The general signature for all the simulations remains similar. Both YSU and MYJ simulations show a similar sensitivity to a longer spin-up time in the heat flux representation (not shown). In the first part of the simulations, the longer spin-up time results in higher values of H of approximately 16 W m^{-2} for the YSU scheme and 30 W m^{-2} for the MYJ scheme. After sunset (around 1900 UTC), the differences between the PBL schemes and the simulations with a different spin-up time disappear. In the modelled LE we find a consistent negative bias between runs with 24 h spin-up and simulations with a shorter spin-up time. During the first day the bias exceeds 80 W m^{-2} or less for the simulation with the MYJ scheme. After sunset the differences between the PBL schemes diminishes. Differences around noon 1 July may partially be caused by the strong bias in the specific humidity at that time (see also previous section). By decreasing the spin-up time the modelled water vapour mixing ratio is closer to the observations (Fig. 10c). The main difference in the 10-m wind speed is found in the afternoon of 1 July, where runs with only 12 h spin-up produced a greater wind speed. At night, the simulations with the shortest spin-up produce much less variation in wind speed, which may suggest a weaker inertial oscillations (Fig. 10d). At 200 m the behaviour of the models remains rather analogous. Around 0100 UTC the simulations with a 12 h spin-up along with the YSU scheme and 18 h spin-up present a sudden decrease of wind speed for which the origin was not discovered and this remains a challenge. The hodograph for 200-m wind reveals weaker inertial oscillations in the simulations with the shortest spin-up time (Fig. 10e). In the 2-m temperature time series there is still a consistent bias, as mentioned in the Sect. 3.3.

Considering the vertical structure of the potential temperature the cold bias decreased by 1 K when using a shorter spin-up time. For the LLJ development a longer spin-up time results in a better representation of the strength and altitude of the LLJ (Fig. 10f). The simulation with only 12 h spin-up time results for the YSU scheme in a weaker LLJ at lower altitude, while for the MYJ scheme only the strength seems to be underestimated by approximately 2 m s^{-1}.

5 Discussions and Outlook

We performed WRF mesoscale model simulations using different boundary-layer schemes to extend the GABLS3 inter-comparison study. A selection of analogous cases was suggested by Baas et al. (2010), where the focus was on improving the performance of single-column models by averaging the forcings over selected cases. To strengthen our results and to generalize the conclusions, our study should be repeated for similar cases. Additionally, with such an extension the model performance could be statistically analyzed, and would help to refine the origin of many biases in coupled models. Work presented in our study may be considered as a simplified guideline for sensitivity tests.

Our results confirm the previously reported bias in $L\downarrow$ regardless of the PBL scheme, which highlights the possible deficiencies in the radiation scheme. The inaccurate temperature and moisture values that are generated by the PBL mixing and serve as an input for the radiation scheme can also contribute to the differences in $L\downarrow$. Negative bias is found in the model

simulations for the GABLS3 case despite the use of different longwave parametrizations (Marina Sterk, 2013 personal communication). However, the modelled net radiation is in better agreement with the observations than are the radiation balance components, which suggests that the model compensates individual biases. The deficiencies in the radiation fluxes result in the underestimation of the temperature tendencies.

Additionally we find a sudden decrease in the the 2-m temperature rise in the model simulations after 0800 UTC 1 July, despite different spin-up times. This temperature decrease should be further investigated since the simulations with different PBL parametrizations show a consistent bias. Moreover, for the simulation with the YSU_TD configuration, we show that the model requires a much larger H than that observed in order to reduce the general negative bias in the 2-m temperature. This was already found by Steeneveld et al. (2011), who reported the correct thermodynamic profiles in the boundary layer but obtained using much larger sensible heat fluxes than those observed by either scintillometry or eddy covariance. We have to note that the surface energy balance closure, for Cabauw observations, still remains an open issue. Surprisingly in our case study, overall the simulations with the TD land-use scheme gave weaker biases for most variables than was the case with NOAH land-use scheme simulations. We performed additional sensitivity tests with additional spin-up of soil parameters to validate our land-use initialization. For our study, additional iteration of the land-use parameters (three times) did not improve the results significantly (not shown). We have to note that for simulations with more complex terrain and/or weather situation, the additional spin-up for land use is strongly recommended and can improve the results.

The simulations highlight the importance of proper and accurate observations to correctly validate a model. As described by de Haij et al. (2006), the measurements using lidar ceilometers for boundary-layer height estimation still require further improvements especially for nighttime. At the start of the night, the accumulation of aerosols is smaller below the top of the SBL than in the residual layer, leading to a difference in the PBL height estimation.

Our study confirms that the TKE schemes better simulate nighttime observations while the daytime is better represented by the first-order schemes. This was also found in the previous GABLS studies (Cuxart et al. 2006; Svensson et al. 2011). In our study, however, the results of the first-order schemes are closer to the results obtained by the higher order schemes. We are aware that some schemes are intended for specific conditions, for instance, the Bougeault–Lacarrere PBL scheme was designed mostly for convective conditions, however for studies with longer experiments stable conditions are unavoidable.

6 Conclusions

This study evaluates the 3D WRF mesoscale model version 3.4.1 for the GABLS3 experiment. The study extends the work by Bosveld et al. (2014a) who evaluated 1D simulations for a range of PBL schemes for the same site. We studied the 24-h period between the noon on the 1 July and noon 2 July 2006, with different, most commonly used, PBL schemes available in the WRF model. We compared the results to the observations from the Cabauw tower facility and De Bilt sounding data. Overall, for the utilized configurations in this study, the WRF model results gave a rather good representation of the near-surface variables and vertical profiles, without additional use of observations (e.g. data assimilation). We were able to distinguish different model behaviour depending on the choice of the PBL scheme. In addition we found substantial bias in the simulations with the Bougeault–Lacarrere PBL scheme ("BOUL" in the text), which should be further investigated.

In all model permutations we found a consistent negative bias of about 20–40 $\mathrm{W\,m^{-2}}$ in the longwave downward radiation, which was also reported in Van der Velde et al. (2010) among others. All the PBL schemes underestimated the daytime 2-m air temperature by approximately 2 K. At night, a substantial cold bias of about 4 K was found regardless of the chosen PBL scheme. In general, those results confirm the biases reported in the single-column model study by Bosveld et al. (2014a). Moreover, the YSU scheme captures the low-level jet at a lower altitude and with higher wind speed compared to the other schemes, contradicting studies with WRF versions previous to the version 3.4.1. This difference may be due to the newly implemented correction to the stability function within the YSU scheme that results in a more stably stratified boundary layer.

We performed a number of sensitivity tests on several technical features. To minimize the bias in the 2-m temperature we performed a sensitivity test with the thermal diffusion land-surface scheme. It appears that the simple 5-layer slab (thermal diffusion) scheme combined with the YSU PBL scheme, reduces the cold bias in the 2-m temperature. Similar to the results in Steeneveld et al. (2011), the smaller bias was achieved by much higher sensible and lower latent heat fluxes for the same site.

In the sensitivity study to the domain size, we found that the selected domain sizes did not directly influence the general signature of most surface variables. In the case of the smallest domain we found the strongest variability, especially for the wind speed components. A too small domain size limits the development of the inertial oscillations and wind speeds at 200 m. The worst model skill was achieved by a model set-up with the smallest domain. The vertical wind structure appears to be strongly influenced by the domain size. In the representation of the LLJ, the reference simulation presented the smallest bias. In the case of a too small domain, the modelled profiles rely too much on the boundary conditions delivered by the ECMWF. Wind speed was also the most sensitive variable to the different model horizontal resolution, and systematic behaviour was not revealed.

Furthermore, we performed simulations with a lower vertical resolution of the ECMWF input files that resulted in an increase of the biases. Therefore, we confirm that the improvement of the quality of the input files is rather important for the model performance. By increasing the amount of vertical grid points within the WRF model we did not find substantial improvements. The strongest sensitivity on the model vertical resolution was found in the 10-m wind-speed representation.

Finally, we studied how a shorter spin-up time affects the model results. The differences between the simulations are rather small for most of the variables, however, we found model sensitivity especially in the beginning of our simulation. In this case a longer spin-up time resulted in a stronger potential temperature bias, but it improved the vertical wind-speed profile. Simulations with the shortest spin-up time provided the weakest inertial oscillations along with the most degraded vertical structure representation, in comparison to the observations.

Acknowledgments The first author acknowledges funding from the NWO-ALW project "Quantifying the role of nocturnal small-scale orographic wave drag on near-surface weather and climate over land", project number 820.01.004 and the project NWO-Computer time SH-060-12. The contribution by G.J. Steeneveld has partly been sponsored by the NWO contract 863.10.010 (Lifting the fog). We would like to thank KNMI, especially Fred Bosveld for making the observations available and Jordi Vilà and Marina Sterk for their valuable comments. We also thank the reviewers for their constructive suggestions.

References

Atlaskin E, Vihma T (2012) Evaluation of NWP results for wintertime nocturnal boundary-layer temperatures over Europe and Finland. Q J R Meteorol Soc 138(667):1440–1451

Baas P, Bosveld FC, Holtslag AAM (2009) A climatology of nocturnal low-level jets at cabauw. J Appl Meteorol Climatol 48(8):1627–1642

Baas P, Bosveld FC, Lenderink G, van Meijgaard E, Holtslag AAM (2010) How to design single-column model experiments for comparison with observed nocturnal low-level jets. Q J R Meteorol Soc 136(648):671–684

Banta R (2008) Stable-boundary-layer regimes from the perspective of the low-level jet. Acta Geophys 56:58–87

Beljaars ACM, Bosveld FC (1997) Cabauw data for the validation of land surface parameterization schemes. J Clim 10(6):1172–1193

Bosveld F, Baas P, Steeneveld GJ, Holtslag A, Angevine W, Bazile E, de Bruijn E, Deacu D, Edwards J, Ek M, Larson V, Pleim J, Raschendorfer M, Svensson G (2014a) The GABLS third intercomparison case for boundary layer model evaluation: part B: results and process understanding. Boundary-Layer Meteorol (this issue)

Bosveld F, Baas P, van Meijgaard E, de Bruijn E, Steeneveld GJ, Holtslag A (2014b) The GABLS third intercomparison case for boundary layer model evaluation: part A: case selection and set-up. Boundary-Layer Meteorol (this issue)

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

Carvalho D, Rocha A, Gómez-Gesteira M, Santos C (2012) A sensitivity study of the WRF model in wind simulation for an area of high wind energy. Environ Modell Softw 33:23–34

Chen F, Dudhia J (2001) Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity. Mon Weather Rev 129:569–585

Chimonas G, Nappo C (1989) Wave drag in the planetary boundary layer over complex terrain. Boundary-Layer Meteorol 47:217–232

Chou S-H (2011) An example of vertical resolution impact on WRF-Var analysis. Electron J Oper Meteorol 12:1–20

Cuxart J, Holtslag A, Beare R, 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

de Haij M, Wauben W, Klein Baltink H (2006) Determination of mixing layer height from ceilometer backscatter profiles. In: Proceedings of SPIE, vol 6362, Stockholm, Sweden

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

Dyer A (1974) A review of flux–profile relationships. Boundary-Layer Meteorol 7(3):363–372

Ek MB, Mitchell KE, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley JD (2003) Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. J Geophys Res Atmos 108(D22):8851

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

Gego E, Hogrefe C, Kallos G, Voudouri A, Irwin J, Rao S (2005) Examination of model predictions at different horizontal grid resolutions. Environ Fluid Mech 5(1–2):63–85

Grachev AA, Fairall CW, Persson POG, Andreas EL, Guest PS (2005) Stable boundary-layer scaling regimes: the SHEBA data. Boundary-Layer Meteorol 116(2):201–235

Guichard F, Parsons DB, Dudhia J, Bresch J (2003) Evaluating mesoscale model predictions of clouds and radiation with SGP ARM data over a seasonal timescale. Mon Weather Rev 131(5):926–944

Holtslag A (2006) GEWEX Atmospheric Boundary-Layer Study (GABLS) on stable boundary layers. Boundary-Layer Meteorol 118(2):234–246

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

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, 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

Janjic ZI (1990) The step-mountain coordinate: physical package. Mon Weather Rev 118(7):1429–1443
Janjic ZI (1996) The surface layer in the NCEP Eta Model. In: 11th Conference on numerical weather prediction. American Meteorological Society, Boston, pp 354–355
Janjic ZI (2002) Nonsingular Implementation of the Mellor-Yamada Level 2.5 Scheme in the NCEP Meso model. Technical report, NCEP
Jankov I, Gallus WA Jr, Segal M, Koch SE (2007) Influence of initial conditions on the WRF-ARW Model QPF response to physical parameterization changes. Weather Forecast 22(3):501–519
Jiménez PA, Dudhia J (2012) Improving the representation of resolved and unresolved topographic effects on surface wind in the WRT model. J Appl Meteorol Climatol 51(2):300–316
Jin J, Miller NL, Schlegel N (2010) Sensitivity study of four land surface schemes in the WRF model. Adv Meteorol 2010:1–11
LeMone MA, Tewari M, Chen F, Dudhia J (2012) Objectively determined fair-weather CBL depths in the ARW-WRF model and their comparison to CASES-97 observations. Mon Weather Rev 141(1):30–54
Liu G, Liu Y, Endo S (2012) Evaluation of surface flux parameterizations with long-term ARM observations. Mon Weather Rev 141(2):773–797
Lothon M, Lenschow DH (2010) Studying the afternoon transition of the planetary boundary layer. Eos Trans AGU 91(29):253–254
Mahrt L (1999) Stratified atmospheric boundary layers. Boundary-Layer Meteorol 90:375–396
Mass CF, Ovens D, Westrick K, Colle BA (2002) Does increasing horizontal resolution produce more skillful forecasts? Bull Am Meteorol Soc 83(3):407–430
Mellor GL, Yamada T (1974) A hierarchy of turbulence closure models for planetary boundary layers. J Atmos Sci 31:1791–1806
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
Morcrette JJ, Geleyn JF (1985) On the influence of different radiation parametrizations on model-generated radiation fields. Q J R Meteorol Soc 111(468):565–585
Nakanishi M (2001) Improvement of the Mellor–Yamada turbulence closure model based on large-eddy simulation data. Boundary-Layer Meteorol 99(3):349–378
Nakanishi M, Niino H (2004) An improved Mellor–Yamada level-3 model with condensation physics: its design and verification. Boundary-Layer Meteorol 112(1):1–31
Nieuwstadt FTM (1984) The turbulent structure of the stable nocturnal boundary layer. J Atmos Sci 41(14):2202–2216
Pleim JE (2007a) A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: model description and testing. J Appl Meteorol Climatol 46(9):1383–1395
Pleim JE (2007b) A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: application and evaluation in a mesoscale meteorological model. J Appl Meteorol Climatol 46(9):1396–1409
Pleim JE, Chang JS (1992) A non-local closure model for vertical mixing in the convective boundary layer. Atmos Environ Part A 26(6):965–981
Savijärvi H (2006) Radiative and turbulent heating rates in the clear-air boundary layer. Q J R Meteorol Soc 132(614):147–161
Seth A, Rojas M (2003) Simulation and sensitivity in a nested modeling system for South America. Part I: reanalyses boundary forcing. J Clim 16(15):2437–2453
Shin H, Hong S-Y (2011) Intercomparison of planetary boundary-layer parametrizations in the WRF model for a single day from CASES-99. Boundary-Layer Meteorol 139(2):261–281
Skamarock WC (2004) Evaluating mesoscale NWP models using kinetic energy spectra. Mon Weather Rev 132(12):3019–3032
Skamarock WC, Klemp JB (2008) A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. J Comput Phys 227(7):3465–3485
Steeneveld GJ, Vilà-Guerau de Arellano J, Holtslag A, Mauritsen AAM, Svensson G, de Bruijn EIF (2008) Evaluation of limited-area models for the representation of the diurnal cycle and contrasting nights in CASES-99. J Appl Meteorol Climatol 47(3):869–887
Steeneveld GJ, Tolk LF, Moene AF, Hartogensis OK, Peters W, Holtslag AAM (2011) Confronting the WRF and RAMS mesoscale models with innovative observations in the Netherlands: evaluating the boundary layer heat budget. J Geophys Res Atmos 116(D23):D23114
Sterk HAM, Steeneveld GJ, Holtslag AAM (2013) The role of snow-surface coupling, radiation, and turbulent mixing in modeling a stable boundary layer over Arctic sea ice. J Geophys Res Atmos 118(3):1199–1217

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, Developmental Testbed Center
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
Sukoriansky S, Galperin B, Perov V (2006) A quasi-normal scale elimination model of turbulence and its application to stably stratified flows. Nonlinear Process Geophys 13(1):9–22
Svensson G, Holtslag A, Kumar V, Mauritsen T, Steeneveld G, Angevine W, Bazile E, Beljaars A, Bruijn E, Cheng A, Conangla L, Cuxart J, Ek M, Falk M, Freedman F, Kitagawa H, Larson V, 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
Teixeira J, Stevens B, Bretherton CS, Cederwall R, Klein SA, Lundquist JK, Doyle JD, Golaz JC, Holtslag AAM, Randall DA, Siebesma AP, Soares PMM (2008) Parameterization of the atmospheric boundary layer: a view from just above the inversion. Bull Am Meteorol Soc 89(4):453–458
Troen IB, Mahrt L (1986) A simple model of the atmospheric boundary layer: sensitivity to surface evaporation. Boundary-Layer Meteorol 37(1–2):129–148
U.S. Geological Survey (2011) U.S. Geological Survey home page. Accessed 6 May 2011
Van de Wiel BJH, Moene AF, Steeneveld GJ, Baas P, Bosveld FC, Holtslag AAM (2010) A conceptual view on inertial oscillations and nocturnal low-level jets. J Atmos Sci 67(8):2679–2689
Van der Velde IR, Steeneveld GJ, Holtslag AAM (2010) Modeling and forecasting the onset and duration of severe radiation fog under frost conditions. Mon Weather Rev 138(11):4237–4253
Van Ulden AP, Wieringa J (1996) Atmospheric boundary layer research at Cabauw. Boundary-Layer Meteorol 78(1–2):39–69
Vannitsem S, Chomé F (2005) One-way nested regional climate simulations and domain size. J Clim 18(1):229–233
Vogelezang DHP, Holtslag AAM (1996) Evaluation and model impacts of alternative boundary-layer height formulations. Boundary-Layer Meteorol 81:245–269
Yang Y, Uddstrom M, Duncan M (2011) Effects of short spin-up periods on soil moisture simulation and the causes over New Zealand. J Geophys Res Atmos 116(D24):D24108
Zhang D-L, Zheng W-Z (2004) Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parameterizations. J Appl Meteorol 43(1):157–169
Zhong S, In H, Clements C (2007) Impact of turbulence, land surface, and radiation parameterizations on simulated boundary layer properties in a coastal environment. J Geophys Res Atmos 112(D13):D13110