ARTICLE

The Third GABLS Intercomparison Case for Evaluation Studies of Boundary-Layer Models. Part B: Results and Process Understanding

Fred C. Bosveld · Peter Baas · Gert-Jan Steeneveld · Albert A. M. Holtslag · Wayne M. Angevine · Eric Bazile · Evert I. F. de Bruijn · Daniel Deacu · John M. Edwards · Michael Ek · Vincent E. Larson · Jonathan E. Pleim · Matthias Raschendorfer · Gunilla Svensson

Abstract We describe and analyze the results of the third global energy and water cycle experiment atmospheric boundary layer Study intercomparison and evaluation study for

F. C. Bosveld (✉) · P. Baas · E. I. F. de Bruijn
Royal Netherlands Meteorological Institute, De Bilt, The Netherlands
e-mail: fred.bosveld@knmi.nl

G.-J. Steeneveld · A. A. M. Holtslag
Wageningen University, Wageningen, The Netherlands

W. M. Angevine
CIRES, NOAA Earth System Research Laboratory, University of Colorado, Boulder, CO, USA

E. Bazile
Meteo France, Toulouse, France

D. Deacu
Environment Canada, Gatineau, Canada

J. M. Edwards
MetOffice, Exeter, UK

M. Ek
National Centers for Environmental Prediction, Washington, DC, USA

V. E. Larson
University of Wisconsin, Milwaukee, WI, USA

J. E. Pleim
U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

M. Raschendorfer
Deutsche Wetterdienst, Offenbach, Germany

G. Svensson
Department of Meteorology and Bolin Centre for Climate Research,
Stockholm University, Stockholm, Sweden

single-column models. Each of the nineteen participating models was operated with its own physics package, including land-surface, radiation and turbulent mixing schemes, for a full diurnal cycle selected from the Cabauw observatory archive. By carefully prescribing the temporal evolution of the forcings on the vertical column, the models could be evaluated against observations. We focus on the gross features of the stable boundary layer (SBL), such as the onset of evening momentum decoupling, the 2-m minimum temperature, the evolution of the inertial oscillation and the morning transition. New process diagrams are introduced to interpret the variety of model results and the relative importance of processes in the SBL; the diagrams include the results of a number of sensitivity runs performed with one of the models. The models are characterized in terms of thermal coupling to the soil, longwave radiation and turbulent mixing. It is shown that differences in longwave radiation schemes among the models have only a small effect on the simulations; however, there are significant variations in downward radiation due to different boundary-layer profiles of temperature and humidity. The differences in modelled thermal coupling to the land surface are large and explain most of the variations in 2-m air temperature and longwave incoming radiation among models. Models with strong turbulent mixing overestimate the boundary-layer height, underestimate the wind speed at 200 m, and give a relatively large downward sensible heat flux. The result is that 2-m air temperature is relatively insensitive to turbulent mixing intensity. Evening transition times spread 1.5 h around the observed time of transition, with later transitions for models with coarse resolution. Time of onset in the morning transition spreads 2 h around the observed transition time. With this case, the morning transition appeared to be difficult to study, no relation could be found between the studied processes, and the variation in the time of the morning transition among the models.

1 Introduction

The global energy and water cycle experiment (GEWEX) atmospheric boundary-layer study (GABLS) focuses on the representation of the diurnal cycle and the stable boundary layer (SBL) in atmospheric models (Holtslag 2006; Holtslag et al. 2013). This is of importance for applications including weather forecasting (Richardson 2009), climate studies (Walsh et al. 2008), atmospheric transport (Nappo et al. 2008), agriculture (Prabha and Hoogenboom 2008), wind engineering (Porté-Agel et al. 2011), and aviation and public transport (Gultepe et al. 2009). Our current understanding of the relevant processes in the atmospheric boundary layer (ABL) and the way they can be represented in an atmospheric model still have significant limitations (e.g., Beljaars and Viterbo 1998; Mahrt 1999; Edwards et al. 2006; Steeneveld et al. 2006; Sterk et al. 2013; Holtslag et al. 2013).

One of the main goals of GABLS is to provide a platform for the atmospheric boundary-layer research community through the organization of model intercomparisons. Here, we focus on the performance of single-column models (SCMs) that include research models and these derived from operational weather and climate models. Two SCM intercomparison case studies have so far been performed in the context of GABLS: one was a highly idealized case over snow with prescribed surface temperature (Cuxart et al. 2006), and the second was based on observations taken during the CASES-99 SBL experiment, also with prescribed surface temperature (Svensson et al. 2011). In these previous studies it was found that the uncertainties

in the large-scale forcings and the exclusion in the models of interaction with the land surface make it difficult to confront the models directly with observations. Moreover, the transitions at sunset and sunrise were found to be difficult to simulate correctly. In separate studies, Holtslag et al. Holtslag et al. (2007) and Baas et al. (2008) showed that the spread in terms of SBL structure, when using various SBL parametrizations, tends to decrease when there is interaction with the surface instead of using prescribed surface flux or surface temperature as a lower boundary condition.

The third GABLS case addresses the issues of the large-scale forcings, the interaction with the surface, evening and morning transitions and the direct evaluation of models with observations. The case was derived from the long-term dataset at Cabauw, The Netherlands. The specific characteristics of the Cabauw site, e.g. its flat topography and sufficient homogeneity (Ulden and Wieringa 1996; Beljaars and Bosveld 1997), make it well-suited to study momentum decoupling around sunset, low-level jet (LLJ) formation, and the morning transition (Angevine and Klein Baltink 2001). The case covers the 24-h period starting at 1200 UTC 1 July 2006, noting that local solar mean time is within 20 min of UTC. This is an (almost) cloud-free period with near-constant geostrophic wind speed over time $\approx 7\,\text{m s}^{-1}$. Daytime well-mixed conditions changed to a turbulent SBL overnight, with a pronounced temperature decrease in time. A well-developed LLJ at around 200 m height was caused by an inertial oscillation. To make a valid comparison with observations possible, care was taken to prescribe realistic geostrophic forcings and dynamic tendencies to the SCMs. These were estimated from both local observations and hindcasts of several three-dimensional numerical weather prediction models. The description of the third GABLS SCM case, details of the selection criteria and the composition of the large-scale forcings have been documented in Bosveld et al. (2014).

Nineteen models from eleven institutes participated in the intercomparison; twelve of these models had participated in GABLS2. The models varied with respect to application, resolution and parametrization of the fundamental processes. Some of the simulations differed only by the choice of turbulence scheme, while other aspects of the models remained the same. The SCMs were run with full physical interaction, e.g. interaction with their own land-surface and radiation schemes.

The intercomparison and evaluation are done in two steps: first the focus is on a comparison of time series of the main features of the model runs and observations, e.g. 2-m air temperature, 200-m wind speed and SBL height. Then we attempt to explain the differences in the behaviour of the models. This is found to be challenging because of the strong interactions between turbulence mixing, radiation and the soil-vegetation system. To unravel these interactions we make use of sensitivity runs performed with one of the models and combine this information in so-called process diagrams that reveal and explain correlations between carefully chosen pairs of parameters.

In Sect. 2 we summarize the processes that play a significant role in the formation of the SBL in particular. After describing the models and the observations used for the evaluation in Sect. 3, we focus in Sect. 4 on a characterization of the differences among models and between models and observations. In Sect. 5 differences among models are analyzed in terms of the basic processes.

2 The Relevant Processes and their Modelling

In this study we focus on the ability of models to represent the time development of the vertical profiles of wind and temperature in the SBL, the growth rate of the SBL and the tran-

sitions around sunset and sunrise. The dynamics of the SBL is governed by many processes (Mahrt 1999). However, for horizontally homogeneous conditions with reasonably strong geostrophic forcing and no fog formation, the dominant processes are turbulent mixing, longwave radiative cooling and thermal coupling to the land surface. First we describe these three processes, and then the evening and morning transitions are explained from the interplay of these processes.

2.1 Turbulent Mixing

As stability increases in a weakly stable ABL, turbulence can be maintained through shear generation in the lowest layers of the atmosphere, which will lead to an increase in the downward sensible heat flux (Holtslag and Nieuwstadt 1986; Mahrt 1999). This results in a relatively slow decrease of the air temperature in the lowest few m of the SBL, and is in contrast with the very stable ABL where turbulence ceases and rapid cooling close to the surface occurs (Wiel et al. 2003). At the top of the SBL turbulence intensity is weak, but in the special case when a low-level jet develops, the increased shear below the jet maximum may result in downward transport of momentum (Smedman 1988; Mahrt 1999).

Complexity of turbulent mixing parametrization schemes ranges from first-order, through 1.5-order to second-order schemes. First-order schemes use a specified profile of exchange coefficients. 1.5-order schemes employ the turbulent kinetic energy (e) equation and combine that with a parametrization of the turbulent mixing length (l), hence they are called $e - l$ schemes. Higher-order closure models are evidently of importance for convective conditions where turbulent transport over large depths is significant. For stable conditions the importance is probably less, since the turbulent transport length scale is limited in extent (Derbyshire 1999).

Baas et al. (2008) show that for stable stratification $e - l$ closure follows local scaling and is therefore approximately equivalent to first-order closure. An advantage of $e - l$ schemes over first-order schemes is that they automatically fulfil the constraint that buoyancy destruction cannot exceed shear production in the steady state. Higher-order closure models are of importance, for example, when the flow is rapidly distorted by topography. For horizontal homogeneous and stable conditions the difference with $e - l$ schemes is limited (Weng and Taylor 2003). Thus, we expect that for a horizontal homogeneous turbulent SBL the behaviour of the different turbulence schemes is characterized by the stability dependency of the simulated dimensionless gradients of momentum and heat. Beljaars and Viterbo (1998) show that introducing a varying turbulent Prandtl number with stability has a significant impact on the development of the SBL in models. Theoretical evidence for an increasing Prandtl number with stability is given by Zilitinkevich et al. (2007), but from an observational point of view important uncertainties remain (Grachev et al. 2007; Sorbjan 2010).

2.2 Radiative Cooling

The main influence of radiation on the SBL is exerted indirectly through the cooling of the land surface and subsequent turbulent transport of heat from the atmosphere to the colder surface. Direct exchange of longwave radiation by absorption and emission in the air has a smaller but still significant effect on the energy budget of the SBL, especially during the evening transition (Steeneveld et al. 2010). The SBL exchanges longwave radiation with the land surface, with the atmosphere above the SBL and through the atmospheric window with outer space. In general, all these three processes lead to cooling under stable conditions.

To quantify the radiation divergence in the SBL we rely on modelling since observations are scarce and very difficult to make. Estournel and Guedalia (1985) estimated radiative cooling for a geostrophic wind speed of 3 and 10 m s^{-1}, and found cooling rates of 0.1 and 0.5 K h^{-1} respectively. Tjemkes and Duynkerke (1989) estimated, on the basis of Cabauw observations and model simulations, that radiative cooling rates in the middle of the SBL for two specific cases amounted to 0.2 K h^{-1}, which was 20 % of the turbulent cooling rates. They suggested that the most important effect of radiative cooling is weakening of the temperature inversion, which results in a 25 % deeper SBL. Ha and Mahrt (2003) found, on the basis of CASES-99 observations and model runs, the same typical cooling rates. They also considered the influence of inversion strength and inversion height and found that cooling rates vary approximately linearly with inversion strength. Dependence on the inversion depth is weak since it only affects the curvature-related part of the radiative cooling, and Ha and Mahrt (2003) found that cooling rates in the lowest 10 m are strongly influenced by the temperature difference with the surface. For each 1 K increase in temperature difference the cooling rate in the lowest 10 m increases by 0.3 K h^{-1}.

In the same study it was found that humidity in the SBL influences cooling rates as well. Each g kg^{-1} increase in specific humidity results in 0.05 K h^{-1} extra cooling. This rate was found at 25 m above the surface whereas Steeneveld et al. (2010) found much larger humidity effects closer to the surface. Savijärvi (2006) and Edwards (2009) arrive at similar conclusions but also report longwave heating at $z < 1$ m in the SBL.

Due to differences in radiation schemes numerical models differ in the amount of downward longwave radiation at the top of the SBL, and in the dependence of the longwave radiation components at the top and bottom of the SBL on the changing temperature structure of the developing SBL (Guichard et al. 2003; Steeneveld et al. 2008). Vertical resolution of models also influences simulated radiation (Räisänen 1996; Niemelä et al. 2001).

2.3 Land-Surface Thermal Coupling

The thermal coupling of the SBL to the land surface has a significant impact on the development of the SBL Holtslag et al. (2007). This impact holds for all the three regimes, i.e. fully turbulent, intermittent and radiation dominated (Steeneveld et al. 2006). Transitions between these regimes also depend crucially on the characteristics of the vegetation-soil system (Wiel et al. 2003). From the perspective of the atmosphere, the vegetation-soil system serves as the lower boundary condition and regulates the surface fluxes. The relevant processes are momentum transport as determined by the roughness of the interface between atmosphere and vegetation, sensible and latent heat flux that regulate energy partitioning at the interface and longwave upward radiation that depends on the temperature at the interface. These are all processes that adapt quickly to changing conditions. The soil heat flux is associated with longer time scales into the system, since heat that is stored during the previous day is released during the night. The magnitude of soil heat flux is affected by soil type and soil water content (Garratt 1994).

Many parametrization schemes for the soil-vegetation-atmosphere interaction have been described previously (e.g. Viterbo and Beljaars 1995; Sellers et al. 1996; Ek et al. 2003). Generally speaking, emphasis has been more on the hydrological daytime aspects (Santanello et al. 2009) than on the thermodynamic nighttime aspects of the canopy-soil system. In many atmospheric models these processes are simplified and described in terms of roughness lengths for momentum and heat, a skin layer and its temperature, and the resistance for heat transport from the skin layer to the soil. Important differences exist among models. Some models use a skin layer with zero heat capacity as an interface (Cox et al. 1999), while other

models directly couple the top soil layer to the atmosphere (Noilhan and Planton 1989). This results in important variations in thermal coupling strength among models and in differences in the timing of the transition periods (Betts et al. 1993; Betts and Ball 1996).

2.4 Evening Transition

The change of sign of the sensible heat flux in the late afternoon induces a significant change in the ABL. Already before this moment of transition the convective intensity in the ABL decreases and the convective-scale turbulence decays (Nieuwstadt and Brost 1986; Pino et al. 2006). After the moment of transition the air overlaying the land surface becomes stably stratified and turbulence intensity is further reduced. This stable layer grows at a rate that is controlled by the geostrophic wind and radiative forcing; the early growing stage is characterized by significant divergences in turbulent fluxes (Grant 1997). The turbulence intensity decreases with height and at a certain height the wind field becomes effectively decoupled, in terms of momentum transport, from the layers below. This induces an inertial oscillation, with the actual wind vector rotating around the geostrophic wind vector (Wiel et al. 2010).

The sign change of the sensible heat flux is induced by the changing surface radiation balance at sunset. Its precise moment in time depends on the properties of the land surface. The interplay between the thermal properties of the surface and the strength of turbulent mixing just after the transition determines the development of the SBL and the timing of momentum decoupling. The magnitude of the ageostrophic wind component at the time of momentum decoupling determines the strength of the LLJ later in the night. This magnitude depends on the evolution of the convective boundary layer and the subsequent early stages of the SBL (Mahrt 1981). The combined evolution of turbulent mixing and the surface fluxes during the transition period often results in a peak in 2-m specific humidity and an inflection point in 2-m temperature (Acevedo and Fitzjarrald 2001).

The correct simulation of the evening transition by atmospheric models is a difficult task given the imperfections in the parameterization of the basic processes. Besides observations, large-eddy simulation (LES) studies can help in understanding the deficiencies of parametrization schemes in numerical models (Edwards et al. 2006). The LES of the evening transition is becoming feasible despite serious demands on computer power needed to resolve the SBL turbulent length scales in a domain that encompasses the convective boundary layer (CBL) (Beare et al. 2006).

2.5 Morning Transition

The morning transition begins when the surface energy budget is first affected by solar incoming radiation sunrise. The quasi-steady state in which the downward sensible heat flux is balanced by longwave cooling is disturbed, and a rise in surface temperature follows immediately. Due to flux divergence the 2-m air temperature also rises. A convective layer begins to grow by entrainment into the SBL of the preceding night, and when the entrainment layer has reached the top of the SBL the growth of the convective layer generally accelerates due to the weaker stability in the residual layer.

Angevine and Klein Baltink (2001) define three moments in time relating to the morning transition: sunrise, crossover of the surface virtual temperature, and onset, the moment that convection has reached the 200-m level. For typical clear days at Cabauw they found times from sunrise to crossover of 2.0 ± 0.5 h and from crossover to onset of 1.2 ± 1.0 h. Surface sensible heat flux integrated over the transition period was found not to be sufficient to explain

the temperature rise in the 200-m air column, which led them to the conclusion that warming by entrainment is significant. The observations also suggest that higher wind speeds lead to a shorter time to onset. The importance of entrainment is confirmed by Beare (2008) on the basis of LES. During the transition the lower atmosphere is in a mixed convective–stable state where the entrainment ratio is much larger then the typical 20 % found in the fully developed CBL. The study clearly demonstrated that higher wind speeds leads to more efficient shear-driven entrainment. Parametrization schemes of turbulent mixing have difficulties to resolve this mixed CBL-SBL state.

3 Models and Observations

In Table 1, the nineteen models that comprise the intercomparison are listed, and the acronyms given here are used throughout to identify the models. A comprehensive description of the models is not feasible here, but for each model a reference is given. Here we concentrate on model aspects that are particularly relevant to the current study. The intercomparison includes operational global models with coarse vertical resolution and research models with fine vertical resolution; eight of the models have a skin layer at the surface of their land-

Table 1 The models with their characteristics

Name	Institute	Nlev	1st full level (m)	ABL Scheme	Skin	(Longwave) Radiation	References
ALADIN	Meteo France	41	17	$e - l$	No	RRTM	Baldauf et al. (2011)
AROME	Meteo France	41	17	$e - l$	No	RRTM	Seity et al. (2011)
MUSC	KNMI	41	17	$e - l$	No	RRTM	Seity et al. (2011)
GLBL38	Met Office	38	10f, 20t	K (LT)	Yes	Other	Martin et al. (2006)
UK4L70	Met Office	70	2.5f,5t	K (ST)	Yes	Other	Martin et al. (2006)
ACM2	USEPA	155	2	K+NL (ST)	No	RRTM	Pleim (2007a,b)
WRF YSU	CIRES	61	55	K+NL (LT)	No	RRTM	Skamarock et al. (2008)
WRF MYJ	CIRES	61	55	$e - l$	No	RRTM	Skamarock et al. (2008)
WRFTEMF	CIRES	61	55	Total $e - l$	No	RRTM	Angevine et al. (2010)
GFS	NCEP	57	20	K+NL (LT)	Yes	RRTM	Hong and Pan (1996)
MYJ	NCEP	57	20	$e - l$	Yes	RRTM	Janjic (2002)
YSU	NCEP	57	20	K+NL (LT)	Yes	RRTM	Hong et al. (2006)
D91	WUR	91	1.7	K (ST)	Yes	Other	Duynkerke (1991)
COSMO	DWD	41	10	Higher order	No	Other	Baldauf et al. (2011)
MIUU	MISU	65	1	2nd order	No	Other	Andrén (1990)
RACMO	KNMI	80	10	$e - l$	Yes	RRTM	van Meijgaard et al. (2008)
GEM	Env. Canada	89	40	$e - l$	No	Other	Bélair et al. (1999)
CLUBB	UWM	250	20	Higher order	No	Other	Golaz et al. (2002), Larson and Golaz (2005)
C31R1	ECMWF	80	10	K (LT)	Yes	RRTM	ECMWF (2006)

Here NL stands for non-local mixing under convective conditions, LT and ST stand for long-and short-tail dependence of mixing on Richardson number respectively

surface scheme. Nine of the models have a first-order turbulence closure scheme, eight have a $e-l$ scheme and two of the models have higher-order schemes. The operational models tend to have relatively strong mixing under stable conditions, whereas research models adhere more to weaker mixing that is more in line with results from field experiments (Cuxart et al. 2006; Svensson et al. 2011). Most of the models employ the Rapid Radiative Transfer Model (RRTM) (Mlawer et al. 1997) for longwave radiation calculations, and most of the remaining models have comparable schemes that also rely on the k-distribution method (Fu and Liou 1992). Differences may occur due to the prescribed profiles of relevant atmospheric constituents.

Several of the models operate in different configurations (Table 1). Three varieties of the Advanced Research Weather and Forecast version 3.0 model are used, i.e. WRFYSU, WRFMYJ, and WRFTEMF, and they only differ in the boundary-layer scheme. Sterk et al. (2013) note that the WRFYSU version had at that time a coding error in the stability function that led to more mixing than is intended for stable stratification. Also three varieties of the National Centers for Environmental Prediction model are operated, all with different boundary-layer schemes, i.e. GFS, YSU, MYJ. AROME, ALADIN and MUSC have the same land-surface scheme, but ALADIN has a smaller soil heat capacity than AROME. All these three models differ in the way convection is treated: RACMO and C31R1 only differ in their boundary-layer scheme. Two versions of the UK MetOffice model are operated, i.e. GLBL38 and UK4L70; they have the same land-surface and radiation schemes, but differ in vertical resolution, 38 and 70 levels respectively. They also differ in their first-order turbulent diffusion formulation—the GLBL38 version sustains significantly more efficient mixing at higher Richardson number than the UK4L70 version.

Observations for the evaluation are taken from the continuous observational program of KNMI at the Cabauw Experimental Site for Atmospheric Research (CESAR), (http://www.cesar-observatory.nl), the Netherlands (51.971°N, 4.927°E); see Bosveld et al. (2014). These include profiles of wind speed, wind direction, temperature and humidity from the 200-m tower at 2, 10, 20, 40, 80, 140 and 200 m. A wind-profiler/RASS provides wind speed, wind direction and (sonic) temperature above the tower. Radiosoundings at midday and midnight are taken from the synoptic station De Bilt 25 km north-east of Cabauw. Downward longwave and shortwave radiative fluxes are obtained from the Cabauw Baseline Surface Radiation Network (BSRN) site, a site dominated by grassland. Upward radiative fluxes are taken from the meteorological field close to the main tower, while sensible and latent heat flux observations are taken from eddy-correlation measurements 200 m to the south. Soil heat-flux observations are made with soil heat-flux plates, and together with soil temperature observations they are used to derive the surface soil heat flux (Beljaars and Bosveld 1997). Friction velocity is derived from the wind speed at 10-m height and the same regional scale roughness length as prescribed in the case set-up. Observed sensible heat flux is used to correct for stability effects. Taking the eddy-correlation observations would result in too low a friction velocity, induced by the local low roughness length. This friction velocity would not be representative of the scale of the boundary layer (Bosveld et al. 2014). All data are processed to 10-min mean values and archived in the CESAR database (http://www.cesar-database.nl).

When comparing model results with observations, it is important to assess the uncertainty in the observations. Measurement uncertainties are small for the state variables, wind speed (1 %), temperature (0.1 K) and somewhat larger for humidity (3 %). Measurement uncertainties in BSRN downward radiation fluxes are small, typically 1 % for shortwave downward radiation and 3 W m^{-2} for longwave downward radiation under clear-sky conditions. Upward radiation fluxes are derived from older instruments having somewhat higher uncertainties: 2

% for shortwave and 5 W m^{-2} for longwave radiation. Failure in closing the surface energy budget from micrometeorological observations (Oncley et al. 2007; Foken 2008 and Roode et al. 2010) forces us to attribute a large uncertainty to the observed surface heat fluxes. For daytime, the uncertainty is typically 10–15 % of the available energy. At nighttime relative uncertainties can be even higher. For the nighttime period 2100–0300 UTC of the current case the observed available energy (net radiation minus soil heat flux) was -36 W m^{-2} and observed turbulent heat flux was -29 W m^{-2}.

Ideally, for an SCM case, the upwind fetch should be homogeneous over a distance that guarantees equilibrium between surface exchange and atmospheric profiles to the top of the ABL: typically this distance will be several tens of kilometres. In the absence of such homogeneity, there will be representation errors, where the locally observed values of relevant parameters close to the surface are not representative of the profiles aloft. At Cabauw, the dominant vegetation (grassland), soil type (clay) and water management regime extend for about 20 km in the (easterly) upwind direction: this suggests that representation uncertainties on this scale will be small. One of the few quantitative measures of homogeneity available is obtained by comparison with data from surrounding automatic weather stations, typical situated at distances of about 50 km from the Cabauw site, for which variations in minimum temperature are found to be within 1 K in the present case.

On the local scale (0.1–1.0 km), there is a problem of representation, since all observations are made over grassland, whereas, on the larger scale, there are rows of trees and small villages. This is most clearly seen in the friction velocity (Beljaars 1982; Verkaik and Holtslag 2007) for which we have made provision, but the problem also exists to a lesser extent for the upward surface fluxes.

4 Results

In this study we are more interested in the variation between models than in the specific behaviour of any individual model; for this reason, and for clarity in the figures, the legend for the models is displayed only once in Table 2.

Table 2 Legends referring to the models as used in the figures

Lines for time series, symbols are used in the process diagrams. Open and closed symbols refer to various classifications used in the process diagrams. Black dot is used for observations
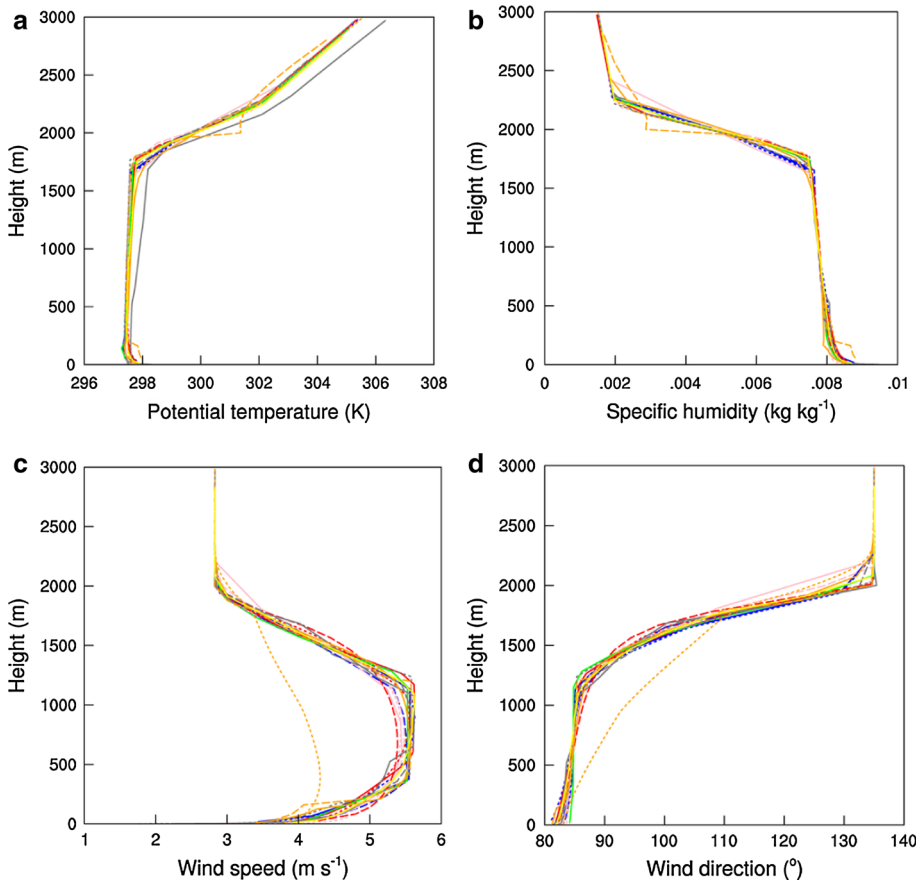
Fig. 1 Vertical profiles of **a** potential temperature, **b** specific humidity, **c** wind speed and **d** wind direction 10 min after initialization (1210 UTC). See Table 2 for legend of model lines

4.1 Initial Conditions and Daytime Simulation

The initial conditions for this case comprises a typical convective situation, with a moderate wind speed and a well-developed CBL. Figure 1 shows profiles of potential temperature, specific humidity, wind speed and direction at 10 min after initialization. Variations in the daytime simulation between models will have an influence on the simulation of the transition at sunset and the subsequent development of the SBL. In this section, we describe the variation among model simulations for the daytime period prior to the evening transition. No figures are shown, but in some cases, where appropriate, reference is made to figures from the following sections.

Even in the first hour of the simulation, there is a substantial variation in the downward shortwave radiation calculated by different models. Whilst the observed flux is 880 W m^{-2}, the simulated fluxes range from 850–950 W m^{-2}. Although some models produce small cloud fractions, as observed, this is insufficient to explain the range of variation in downward shortwave radiation. The variation must be attributed to differences in the clear-sky shortwave radiation calculations, possibly related to different aerosol concentrations. Differences in

downward longwave radiation are much smaller and the simulated range is 355–375 W m^{-2} (see Fig. 7a). Advection of dry air around midday (Bosveld et al. 2014), not represented in the simulations, has a profound effect on the downward longwave radiation and consequently a comparison with observations could be made only in late afternoon, at which time the observed values lie towards the top end of the range of modelled values.

Modellers were asked to adjust the soil moisture in their initial conditions, so that a Bowen ratio equal to the observed value of 0.33 was obtained in the early afternoon. The modelled Bowen ratios vary from 0.25 to 0.40, showing that a precise adjustment of the models was not trivial. In some models, the initialization of the soil temperature profile led to an imbalance with the surface energy fluxes (see Fig. 6a). In these cases, the heat flux into the soil is in general too high and a significant adjustment is seen in the first few hours of the simulation. To study the impact of such an imbalance on the simulation, a test was carried out using the RACMO SCM but with increased thermal coupling between the soil and the atmosphere, so that a significant imbalance occurred. As a balanced soil temperature profile for this configuration, we took the profile from the end of this simulation. The simulation was then repeated using this balanced temperature profile as the initial condition. Compared with the balanced simulation, the unbalanced one showed unrealistically large soil heat fluxes in the first few hours. By the time of the evening transition, the differences had become small, suggesting that a model intercomparison of the SBL can be carried out despite such initial imbalances in some of the models.

The modelled sensible heat fluxes ranged from 80–140 W m^{-2}, whereas the observed value was 120 W m^{-2} one hour after noon (see Fig. 4a). The sensible heat flux is an important determinant of the growth of the CBL. The modelled boundary-layer height in the first few hours varies between 1800 and 2200 m, which is close to the value of 2000 m prescribed at initialization, but diverges from the observed value of 1500 m because of the effect of dry-air advection around midday, as mentioned above. Latent heat fluxes range from 300–400 W m^{-2}, while a value of 365 W m^{-2} was observed. As expected, models with high latent heat fluxes tend to have relatively low sensible heat fluxes.

The observed friction velocity is 0.4 m s^{-1}, while modelled values range from 0.35–0.5 m s^{-1} (see Fig. 4b). Although a specific roughness length for momentum is prescribed, the simulated bulk transfer coefficients for momentum between 10 m and the surface range from 8–19 $\times 10^{-3}$, while the observed value was 13 $\times 10^{-3}$.

4.2 State and Structure of the Stable Boundary Layer

By midnight a well-defined SBL has developed, as can be seen from the profiles displayed in Fig. 2. A surface-based inversion has formed, the strength of which is overestimated by most of the models. The profile of specific humidity shows little structure, but the tower observations suggest a specific humidity above 140 m somewhat lower than that predicted by the models. The maximum wind speed for the LLJ is at 200 m in the observations. Many models show a more gradual variation in the wind field with height than do the observations, but the turning of the wind with height is reasonably captured by most of the models. Differences between the radiosounding and the tower observations must be attributed to their spatial separation of 25 km.

To judge how models simulate the time evolution, we consider time series of the state variables. Figure 3a shows modelled and observed time series of the 2-m air temperature. The general characteristics of the temperature change are well captured by the models, i.e. an initial fast decrease, followed by a more gradual decrease in the subsequent hours, and then, from one hour before midnight, slightly faster cooling. At the time of the minimum
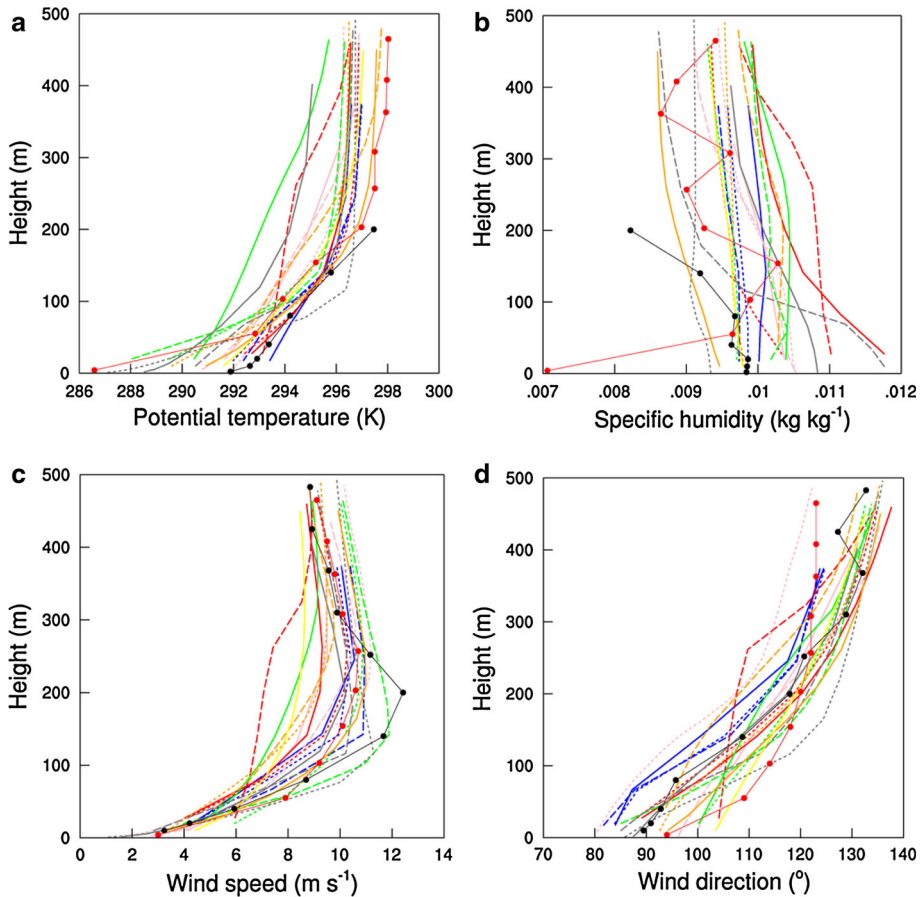
Fig. 2 As in Fig. 1 but at midnight (0000 UTC) and up to 500 m. *Black dots* are observations from tower and wind profiler, *red dots* are radio sounding data. See Table 2 for legend of model lines

temperature, seven out of the 19 models yield values within 1.5 K of the observations, while the remaining models give temperatures up to 5 K lower than observed. The specific humidity at 2 m (Fig. 3b) shows peaks around the evening and morning transitions. These peaks are due to significant latent heat fluxes at times when mixing from the surface to higher levels is still limited. Especially around sunset, these peaks are more pronounced in most of the models than in the observations.

Winds around the 200-m level are shown in Fig. 3c. For each model, the level closest to 200 m was chosen, giving a range of 180–220 m. The 200-m level is of interest because, in the observations, it is well decoupled from the surface in terms of momentum exchange, and exhibits a clear inertial oscillation. In the first hour after momentum decoupling the observed flow accelerates more strongly than the modelled flow. The inertial oscillation is significantly affected by horizontal momentum advection especially after midnight (Bosveld et al. 2014), and is clearly seen in most of the simulations, which show a sharp decrease in wind speed after midnight, much sharper than would be expected if no advection was prescribed. The 200-m wind speed in all models peaks at 2300 UTC, but all show values lower than observed, although more than half of the models show peaks within 2 m s^{-1} of the observed values.
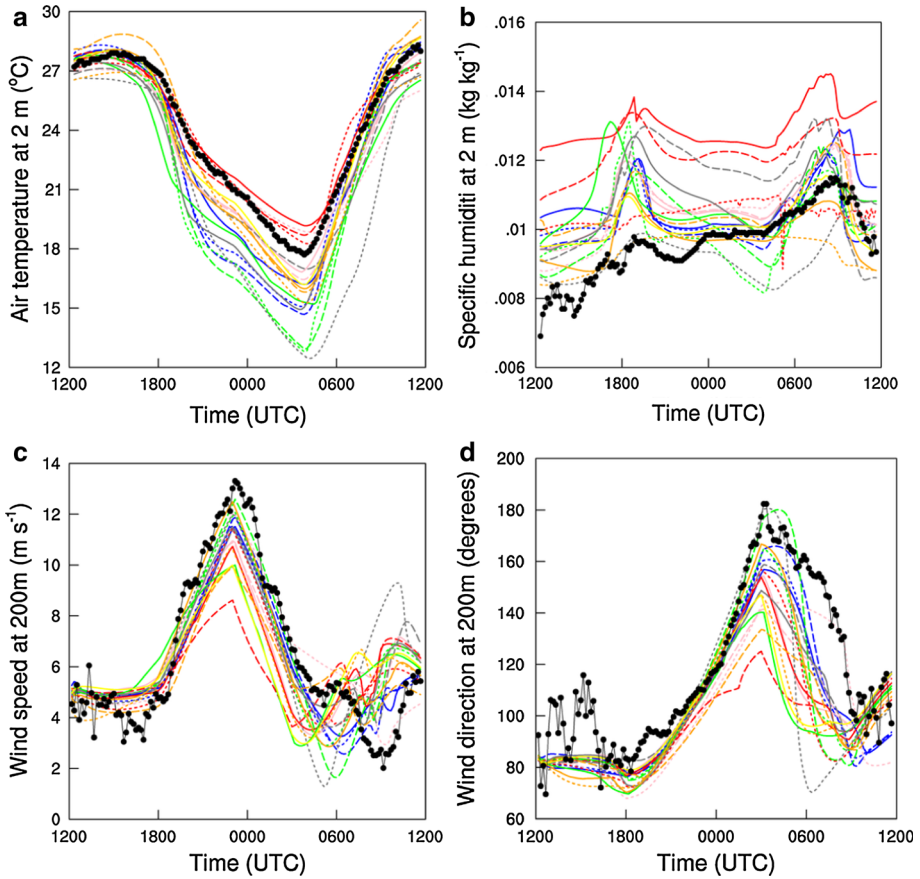
Fig. 3 Time series of **a** air temperature at 2 m, **b** specific humidity at 2 m, **c** wind speed at 200 m and **d** wind direction at 200 m. Observations are in *black*. *Coloured lines* are the models. See Table 2 for legend of model lines

Around and after sunrise, models start to differ from each other and from the observations. At the 80-m level (not shown), which is well within the turbulent layer, a few models show wind speeds higher than observed. For the wind direction (Fig. 3d) similar behaviour is observed, with a good correspondence in the time of maximum veering, but with many models giving less turning of the 200-m wind direction after midnight.

Some models show excessive downward sensible heat flux at night (Fig. 4a), but most are quite close to each other and to the observations. Models tend to overestimate the friction velocity before midnight (Fig. 4b) but underestimate after midnight. Finally we consider the boundary-layer height. Models differ in the way in which the SBL height is calculated internally. A robust and uniformly applicable method is to define the boundary-layer height as that at which the air temperature attains its maximum value. Figure 5 shows the evolution of the SBL height defined in this way. Note that for the period from 2200–0400 UTC no observational estimate could be made because of the limited height of the tower, except for one instant at 0000 UTC when a radiosonde was available. A check was performed to ensure that all the values represent surface-based inversions. Two-third of the models give
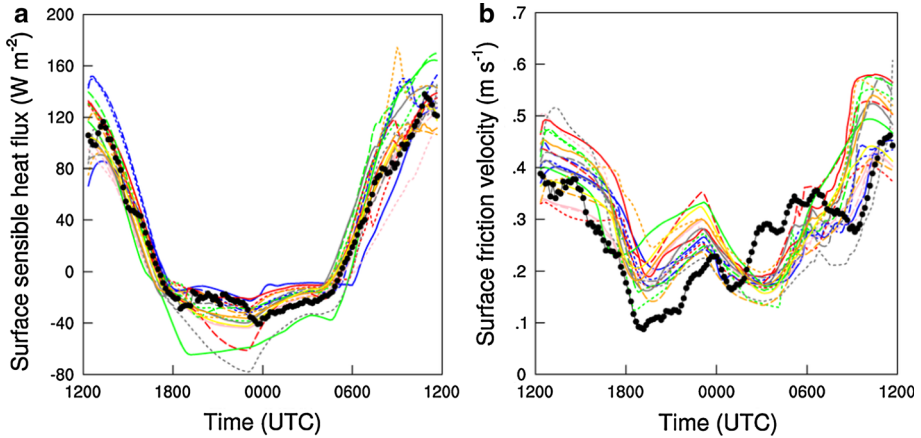
Fig. 4 Time series of **a** sensible heat flux and **b** friction velocity. Observations are in *black*. *Coloured lines* are the models. See Table 2 for legend of model lines

Fig. 5 Stable boundary-layer height, based on air temperature maximum, as function of time. Between 2200 and 0400 UTC no temperature maximum could be detected along the tower. The *black dot* at 0000 UTC is derived from the radiosounding. See Table 2 for legend of model lines
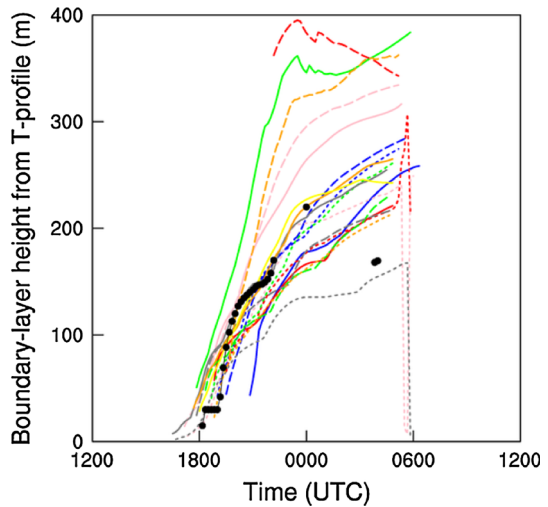
satisfactory results compared with the observations, but the remaining models overestimate SBL height by as much as a factor of two.

5 Analysis

Time series of atmospheric parameters, as shown in the previous section, have already revealed interesting information about the performance of the models and their relation to the observations. To facilitate further interpretation in terms of the representation of the dominant physical processes in the models, we reduced the raw results in such a way as to illustrate the essential properties of the models. Examples of such data reduction are the minimum 2-m temperature, the maximum wind speed at 200 m, the temperature decrease

over the night, the mean longwave cooling after midnight, the time of the evening transition, and the time of the morning transition. Below, we plot various combinations of these data, together with the results from model sensitivity runs, in process diagrams, and take these diagrams as a starting point for further analysis.

The sensitivity runs have been performed with one of the participating models (RACMO). Each run is designed to affect one of the relevant processes in the SBL in such a way as to encompass a significant part of the variation in that specific process between models. In the sections below, it will become clear how these process-specific variations among models are quantified. It will be shown that the chosen range for the sensitivity runs does indeed match these variations. Thermal coupling to the land-surface/soil system was changed by varying the thermal conductance (Λ) between the skin layer and the soil. Typical published values are between 2.5 and 10 W m^{-2} K^{-1}. To achieve the desired effect, it was necessary to choose a much larger range. We used $\Lambda = 0.5$, 5 and 50 W m^{-2} K^{-1}, where the middle value is that used in the standard run. These runs are termed *coupling*, which refers to thermal coupling to the land surface. The efficiency of turbulent transport between the top of the vegetation and the lowest model layer is expected to vary little among models. Roughness lengths are prescribed and variations in flux-profile relations will not be large at the moderate stabilities occurring in this case. Therefore, it is more interesting to change the activity of turbulent mixing above the lowest model level. In RACMO, with its $e - l$ scheme, this can be accomplished by varying the coefficients that relate the turbulent length scales for heat and momentum to e and the stability of the flow (Baas et al. 2008). To obtain the desired variation, we chose $c_h = 0.1$, 0.2 and 0.4, where c_h is the coefficient for heat. This changes the turbulent length scale by a factor of two relative to the standard run. These runs are termed *mixing*. Longwave radiation can be varied in the model by changing the concentration of one of the greenhouse gases. By increasing the concentration of water vapour, cloud formation was induced in the latter part of the morning of the second day, and evaporation may be influenced by such a perturbation. These problems are not encountered when performing simulations with higher or lowered concentration of CO_2. Simulations with concentrations of 50 and 1500 ppm were performed to give changes of 5 W m^{-2} in the downward longwave radiation at the surface relative to the standard run. These runs are labelled *radiation*. The sensitivity runs were repeated with a second model (D91) to investigate the generality of the results. In general both models showed the same sensitivity in the relevant parameters for changes in *coupling*, *mixing* and *radiation*, so only the sensitivity results of RACMO will be presented.

The first three parts of this section focus on these three dominant processes and their relation to the characteristics of the SBL, and we examine combinations of parameters that are likely to be significantly influenced by the respective processes. The last two sections are concerned with the evening and morning transitions.

5.1 Thermal Coupling with the Land Surface

Time series of the surface soil heat flux are displayed in Fig. 6a. Large differences are observed between models, and part of the large variations in the first few hours after initialization are due to thermal imbalances, as explained in Sect. 4. The diurnal amplitude in the observations is small compared to that in most of the models. The surface soil heat flux is, to a large extent, determined by the evolution of the skin temperature, but it is also affected by the penetration of radiation through the vegetative canopy to the top of the first soil layer. Duynkerke (1991) estimated the fraction of downward radiation that reaches the soil layer as 3 % at Cabauw, which suggests that the influence of radiation is minor. From a long times series of Cabauw
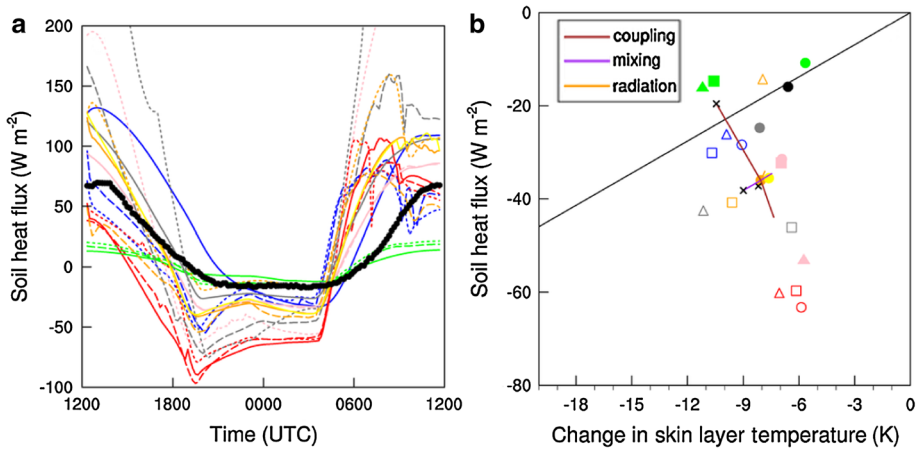
Fig. 6 **a** Soil heat flux as a function of time. **b** Nighttime (0000–0300 UTC) surface soil heat flux as function of skin temperature change over the previous evening (1800–0000 UTC). The *black line* represents points with the same thermal coupling strength as observed. *Closed symbols* represent models with a skin layer. The legend represent the sensitivity lines. The *coloured lines* connect the points of the sensitivity runs (see text). *Crosses* at one end of the sensitivity lines indicate the sensitivity runs with the weakest surface thermal coupling, weakest turbulent mixing and the lowest downward longwave radiation respectively. See Table 2 for legend of model points

observations we found a good correlation between the decrease in skin temperature between 1800 and 2400 UTC and the mean surface soil heat flux over the period 0000 to 0300 UTC the next day, with a constant of proportionality of 2.3 W m^{-2} K^{-1}. This coefficient quantifies the strength of the thermal coupling between the atmosphere and the land surface at the Cabauw site.

For each model, Fig. 6b shows the decrease in skin temperature between 1800 and 2400 UTC and mean surface soil heat flux after midnight. The black dot represents the observation for this case. The black line represents all the points in the diagram that have a thermal coupling strength of 2.3 W m^{-2} K^{-1} as found above from the longer time series. The black dot lies almost on the black line. A considerable spread among models is observed, mainly perpendicular to the line of equal thermal coupling strength. Also shown are the results of the tests of sensitivity. The red line for *coupling* sensitivity connects the points representing the RACMO simulation with low, standard and high land surface thermal coupling strengths. This line has the same orientation as the main spread among the models and covers a significant part of that spread. This justifies the range in values of the thermal coupling strength chosen for the sensitivity test. In the same way, sensitivitiy lines are drawn for changing *radiation* (orange) and *mixing* (blue). These lines are short, showing that changes in longwave radiation and mixing have little impact on the relation between the skin temperature change and the soil heat flux. This is confirmed by considering groups of models which have the same surface schemes, but different turbulent mixing schemes. Members in the groups (WRFTEMF, WRFYSU, WRFMYJ), (UK4L70, GLBL38), (AROME, MUSC, ALADIN) and (C31R1, RACMO) are close to each other in the process diagram.

These results clearly demonstrate that differences in modelled thermal coupling play a significant role in explaining the differences in surface soil heat flux and the nocturnal decrease in near-surface air temperature between models. It is found that those models with a skin layer (closed symbols) are clustered in the region between the observations and the standard simulation with RACMO. The group of models without a skin layer shows a somewhat larger
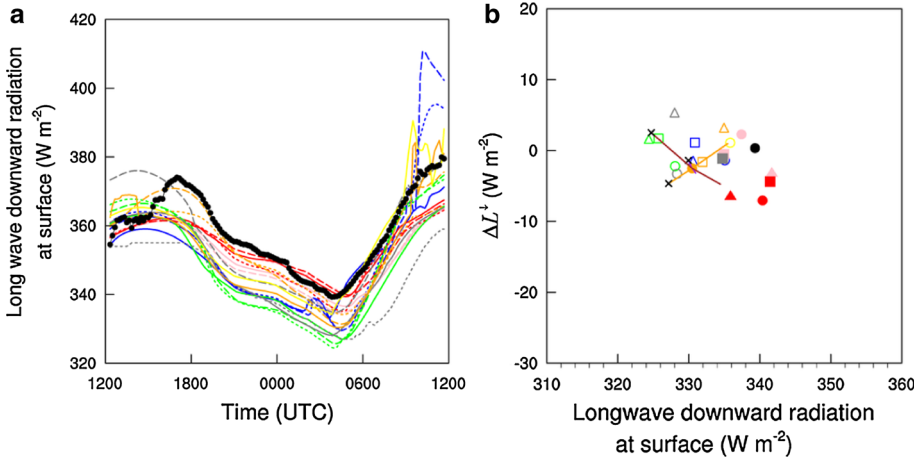
Fig. 7 **a** Longwave downward radiation at the surface as function of time. **b** Deviation of longwave downward surface radiation from the reference radiation as function of longwave downward surface radiation itself at time of minimum temperature (0400 UTC). In graph **b** models with strong thermal coupling have *closed symbols*. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model lines and points

variation in the diagram, although some of them (MUSC, AROME, ALADIN, GEM) are also clustered near the group with a skin layer. The observations suggest that most models overestimate the thermal coupling to the surface. These results enable a classification into strongly and weakly thermally coupled models that will be used in the next section.

That it is not only the skin temperature that is strongly influenced by thermal coupling but also the 2-m temperature is shown in Fig. 8, which will be fully discussed in the next subsection. For the present purpose, it shows that the spread in 2-m temperature between models is comparable to the length of the sensitivity line for thermal coupling, and that strongly and weakly thermally coupled models, as indicated by closed and open symbols respectively, form distinct groups in the diagram.
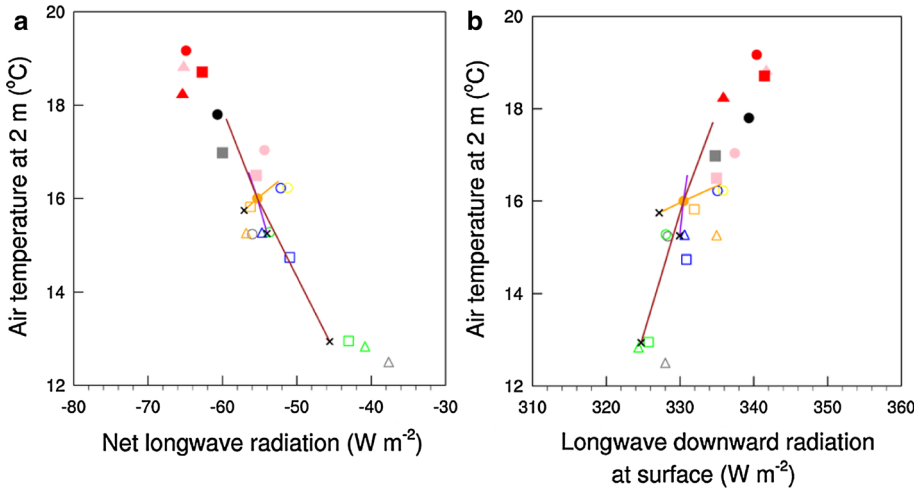
5.2 Radiation

Figure 7a indicates a general underestimation of downward longwave radiation at the surface ($L^\downarrow$) by the models). Larger variations at the end of the simulation are caused by the formation of convective clouds in some models. Variations in $L^\downarrow$ between models may originate from differences in the formulation of the radiation schemes and also from differences in prescribed green house gas concentration and aerosol profiles. Note that the latter are not prescribed in the case set-up. Differences in radiation can also occur because of differences in the simulated thermodynamic profiles. Typically, more than half of the downward longwave radiation reaching the surface originates within the lowest hundreds of the atmosphere (Ohmura 2001), so longwave radiation is intimately connected to the representation of the temperature profiles and, as such, to the dynamics of the SBL. To estimate the effect of deviations in thermodynamic profiles on deviations in $L^\downarrow$ among models we developed an extension of Brunt's (1932) formula that includes the effect of the lapse rate in the lowest few hundreds of the atmosphere (see Appendix). This reference downward longwave surface radiation, $L^\downarrow_{ref}$, is a function of the air temperature at 2 m and 200 m and of the humidity at 200 m. Using this reference $L^\downarrow_{ref}$ we can isolate variations in $L^\downarrow$ between models due to differences in the radiation scheme by correcting for the effect of variations in the thermodynamic profiles.

We define $\Delta L^{\downarrow}$ as the difference between $L^{\downarrow}$ and the reference $L^{\downarrow}_{ref}$. Thus we expect that the influence of the thermodynamic profile will be absent in $\Delta L^{\downarrow}$. Figure 7b shows $\Delta L^{\downarrow}$ as a function of $L^{\downarrow}$ itself at 0400 UTC, which is the time of the minimum 2-m temperature. Modelled values of $L^{\downarrow}$ differ by between -15 and $2\,\mathrm{W\,m^{-2}}$ from the observations, and values of $\Delta L^{\downarrow}$ lie between -7 and $5\,\mathrm{W\,m^{-2}}$. As expected, the observed value of $\Delta L^{\downarrow}$ is very close to zero. Thus, the spread among models is decreased by taking into account the influence of the simulated thermodynamic profiles on longwave radiation and the spread is more centred on the observed value. We found no correlation between $\Delta L^{\downarrow}$ and the minimum 2-m temperature. Note that the sensitivity line for radiation lies in approximately the 1:1 direction, showing that the reference $L^{\downarrow}_{ref}$ indeed retains differences in the radiation schemes. The amplitude of the radiation sensitivity line covers a significant range of the variation in $\Delta L^{\downarrow}$ variation between models. A significant sensitivity to *coupling* remains, seen both from the sensitivity line and from the position of strongly coupled models in the diagram (closed symbols). This indicates the difficulty in isolating the radiation process because variations induced by the different radiation schemes seem to be relatively small. Models with the RRTM scheme did not show behaviour systematically different from that of the other models.

We now turn to the relation between surface radiation budget and the 2-m temperature. Net longwave radiation at the surface is tightly coupled to the air temperature at 2 m. Figure 8a shows the 2-m temperature as a function of the net longwave radiation at the time of the minimum 2-m temperature. A clear correlation is found, with the strongest radiative cooling occurring in those models where air temperatures are high. The result can be understood by noting that one component of the net radiation, the upward longwave radiation, is directly coupled to the surface temperature. In Fig. 8b the upward longwave radiation is removed, and the 2-m air temperature is plotted against $L^{\downarrow}$. Again we see a strong correlation in the variation among models, but with the opposite sign. Thus, a high temperature is correlated with a high value of $L^{\downarrow}$. In both figures, the spread of the model points is aligned with the line representing the *coupling* sensitivity, while the lines representing the *radiation* and *mixing* sensitivity are short. This is confirmed by the classification into strongly (closed) and weakly coupled (open symbols) models as derived from Fig. 6b. Thus, the thermal coupling to the land surface is the major factor in explaining the variation in minimum temperatures among models. Weak thermal coupling leads to low minimum temperatures because soil heat flux is small. Both downward and upward longwave radiation decrease in response to falling temperatures, both at the surface and in the developing SBL. The decrease in upward radiation is larger by a factor of two than that in the downward radiation because the lowest temperature occurs at the surface, which explains the reversal in the sign of the correlation between the two figures.

The observed thermal coupling strength (Fig. 6b) is of the same magnitude as that of the most weakly coupled models. Moreover, the sensitivity line representing thermal coupling to the soil indicates that observed thermal coupling is relatively weak. Figure 8 can be interpreted as an indirect measure of the thermal coupling to the soil. Here we see that observations are comparable with the strongly coupled models and lie close to the strongly coupled end of the corresponding sensitivity line. In fact, the observed thermal coupling strength in Fig. 6 is 2.5 times weaker than that suggested by the observational point in Fig. 8. An underestimation of the observed thermal coupling strength could be due either to an overestimation of the decrease in observed skin temperature drop or to an underestimation of the magnitude of the observed surface soil heat flux. Given the fundamental problems in closing the surface energy budget, it is not possible to be certain of the soil heat flux observations. Certainly,

Fig. 8 2-m air temperature at 0400 UTC as function of **a** net longwave radiation and **b** longwave downward radiation. In both frames models with strong land-surface thermal coupling are represented with *closed symbols*. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model points

larger absolute values of the soil heat flux would result in better closure of the surface energy budget (Jacobs et al. 2008).

It is clear that $L^{\downarrow}$ is not an external parameter to the SBL: its value depends significantly on the evolution of the SBL. As such, the net longwave cooling of the surface is less sensitive to changes in surface temperature than it would be if $L^{\downarrow}$ were a true external parameter. For one of the models (D91) it was we found that, during nighttime, the radiative loss from the lowest 200 m of the atmosphere was 16 W m^{-2}. This was largely due to the divergence of upward longwave radiation. The divergence of the downward radiation was much smaller. The divergence of the radiative flux in the SBL will be largest for those models with low minimum temperature (strong inversion) and this gives rise to a positive feedback, further decreasing the minimum temperature. However, since this flux divergence is relatively small, it only partly counteracts the direct negative feedback on further cooling due to the strong relation between surface temperature and longwave upward radiation.

5.3 Turbulent Mixing

Variations in mixing efficiency between models will be expressed variations in the boundary-layer height, the turbulent fluxes, and the speed of the LLJ, as will become clear in this section. Here, we seek a measure of turbulent mixing efficiency that can be derived in a uniform way from the output of the models. We note that a comparison of diffusion coefficients would not be meaningful, since even models with the same mixing scheme will differ in their stratifications, because of differences in other aspects of the model (thermal coupling and radiation). Following Louis (1979), we assume that the diffusion coefficient, K, of a quantity X (either temperature or momentum), depends on the height, shear and stability,

$$K_X = \ell_X^2(z) F_X(Ri_G) \frac{\Delta U}{\Delta z} = - <wx> \frac{\Delta z}{\Delta X}. \tag{1}$$

Here, $\langle wx \rangle$ is the vertical turbulent flux, positive when directed upward, Δz is the vertical grid spacing, ΔX the difference of X over the vertical distance Δz, l_X is a mixing length for neutral
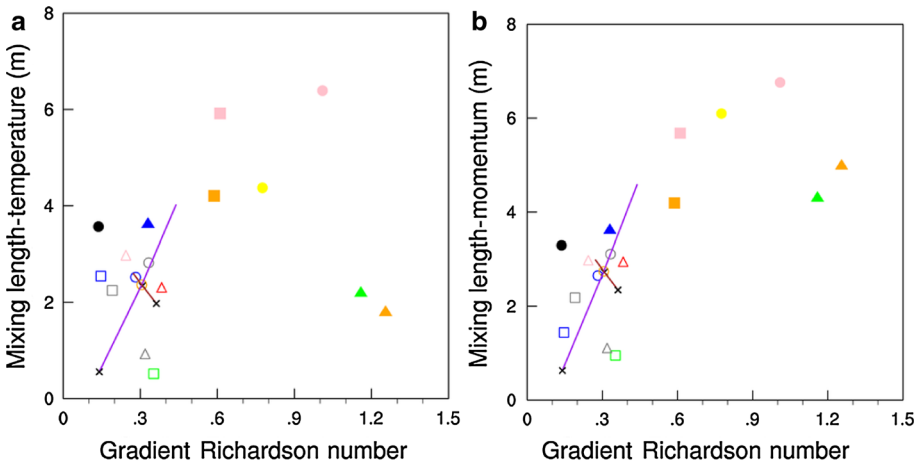
Fig. 9 Mixing length scales as function of gradient Richardson number at the 60-m level for **a** temperature and **b** momentum around midnight (2300–0100 UTC). In both graphs *closed symbols* represent models which are labelled as strong mixing. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model points

conditions and will be a function of height, ΔU is the magnitude of the vector difference in the wind field over Δz and F is a dimensionless function of the gradient Richardson number, which is 1 for neutral conditions. The gradient Richardson number is a measure of local stability and is defined by,

$$Ri_G = \frac{g\Delta z}{\theta}\frac{\Delta\theta}{(\Delta U)^2}, \qquad (2)$$

where g is the acceleration due to gravity, and θ is absolute potential temperature. In both relations, we already anticipate application to numerical models and observations by writing finite difference forms. The expression on the right-hand side of Eq. 1 can be calculated from the output of the models. From the middle expression we can derive a stability-dependent mixing length by combining l_X and the stability function,

$$L_X(z, Ri_G) = \ell_X(z)\sqrt{F_X(Ri_G)} = \sqrt{-<wx>\frac{(\Delta z)^2}{\Delta X \Delta U}}. \qquad (3)$$

The expression on the right-hand side of the equation can be calculated from the output of the various models to yield the mixing length in a consistent way. Near the surface, z will dominate and the length scale is expected to approach κz, where κ is the von Karman constant. With sufficient stratification, z will become less important at greater heights and the length scale will largely be determined by Ri_G. For a given z, we can compare the mixing efficiency deduced from the models and from observations.

In Fig. 9 the mixing lengths for heat and momentum around midnight, deduced from Eq. 3, are plotted as a function of the gradient Richardson number. Data from the models are shown for the grid level closest to the 60 m. The 60-m level is well within the turbulent SBL and above the lowest grid level of all models. For models with a staggered grid, where the state variables and fluxes are on different levels, the calculations were straight forward. Some models had a different grid configuration and an extended version of Eq. 3 was applied, allowing for different levels for momentum and temperature. Finally, for some models no fluxes were available and the calculation could not be done. The derived mixing lengths shown in Fig. 9 are substantially smaller than the value of 24 m that would be the expected at

this height under neutral conditions. Therefore, we anticipate that although the chosen model levels do not, in general, coincide with a height of 60 m, a fair comparison of the mixing efficiency can still be made.

From Fig. 9 we observe a significant spread in combinations of $L_{H,M}$ and Ri_G. It is found that models with large mixing lengths also have higher stabilities, which may seem counter-intuitive at first sight. However, this result is confirmed by the *mixing* sensitivity line, the length of which is comparable to the range spanned by the models. A larger mixing length will result in stronger mixing. Assuming, to first order, that the gradients of temperature and wind are decreased by the same factor as a result of this strong mixing, then stability will increase. This increased stability will partly, but not fully, counteract the increased mixing. This plot enables a classification into strongly and weakly mixing models to be made. On the basis of the diagram, models were labelled strongly mixing (closed symbols) and weakly mixing (open symbols). Note that the models CLUBB and YSU are labelled "strongly mixing" even though they show much weaker mixing for temperature than for wind.

For those models with first order mixing schemes, the mixing strength can be determined directly from the functional form of F_X in Eq. 1. Baased on the dependence of mixing on Richardson number these functional forms are often described as long- or short-tail. From the model F_X's, we conclude that the models C31R1, GLBL38 and YSU can be labelled as long-tailed and the models D91, ACM2 and UK4L70 can be labelled short-tailed. Encouragingly, for five out of these six models, this corresponds with classification into strong mixing and weak mixing made here. The exception is UK4L70, which, though labelled as strongly mixing, has a short-tailed functional form. Note that UK4L70 is one of the models which has separate grids for temperature and momentum, which makes the derivation of the mixing length more complicated.

Note that the quantities on the two axis of Fig. 9 contain common factors (ΔU), but here no self correlation was encountered (see also Baas et al. 2006), since the model variables are not stochastic variables, in contrast to wind shear observations in a turbulent atmosphere.

Figure 10a shows mean nighttime sensible heat fluxes as a function of the boundary-layer height, derived from the temperature profile. Larger downward sensible heat fluxes are correlated with deeper SBLs, resulting from more efficient turbulent mixing, as is illustrated by the *mixing* sensitivity line. Given this correlation, it is to be expected that the rate of change of temperature in the SBL will not be very sensitive to the mixing efficiency, since, for larger downward sensible heat fluxes the heat is extracted from a deeper turbulent layer. This is confirmed by Fig. 10b where the minimum 2-m temperature is plotted as a function of the boundary-layer height. The *mixing* sensitivity lines suggest a weak influence on minimum temperature, but the spread among models seems to be dominated by *coupling*. The shot length of the *coupling* sensitivity line in Fig. 10a shows that land surface thermal coupling is of lesser importance for evolution of boundary-layer height. The closed symbols in Fig. 10 represent strongly mixing models as derived from Fig. 9. We observe, as expected, that these models have, in general, deeper boundary layers, except for the models CLUBB and YSU. Note that CLUBB and YSU have weaker mixing for temperature.

Figure 11a shows that deep boundary layers exhibit relatively low heights of the wind speed maximum (approximately 200 m). This illustrates the dominant influence of mixing on these parameters, as is confirmed by the *mixing* sensitivity line. The backing of the surface wind relative to the geostrophic wind is shown in Fig. 11b. Despite the scatter, a trend toward reduced backing in the case of deeper boundary layers can be seen: this can be attributed to the influence of mixing, as shown by the *mixing* sensitivity line. We note from the *coupling* sensitivity lines that the backing of the wind is not influenced by the strength of thermal coupling, whereas the wind speed maximum does show some sensitivity. Again, strongly
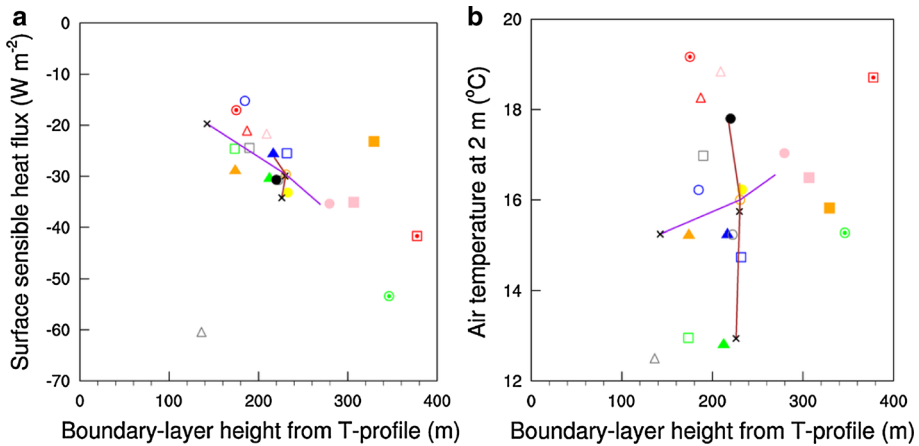
Fig. 10 Relation between boundary-layer height after midnight (0000–0200 UTC) and **a** nighttime sensible heat flux (2200–0200 UTC) and **b** minimum 2-m temperature (0400 UTC). In both graphs *closed symbols* represent strongly mixing models. Models that could not be classified have a dot in the *open symbol*. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model points
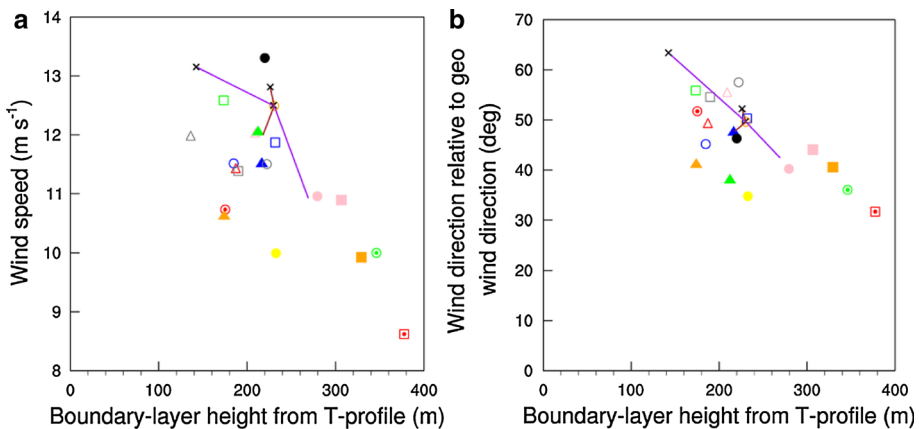

Fig. 11 Relation between boundary-layer height after midnight (0000–0200 UTC) and **a** nighttime wind speed maximum and **b** maximum backing of the surface wind relative to the geostrophic wind. In both graphs *closed symbols* represent strongly mixing models. Models that could not be classified have a dot in the *open symbol*. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model points

mixing models (closed symbols) are clustered on the right of the diagrams, except for the models CLUBB and YSU. Note that UK4L70 shows a relatively large boundary-layer height, corresponding to its classification as "strongly mixing," despite being judged as having a short-tailed mixing scheme.

5.4 Evening Transition

We define the crossover time during the evening transition as the moment when the buoyancy flux, expressed in terms of the surface virtual sensible heat flux, becomes negative and a stably stratified layer starts to develop close to the surface. Model crossover times vary between −0.5 and 0.5 h relative to the observation as can be seen in Fig. 4a. The dryness of the land
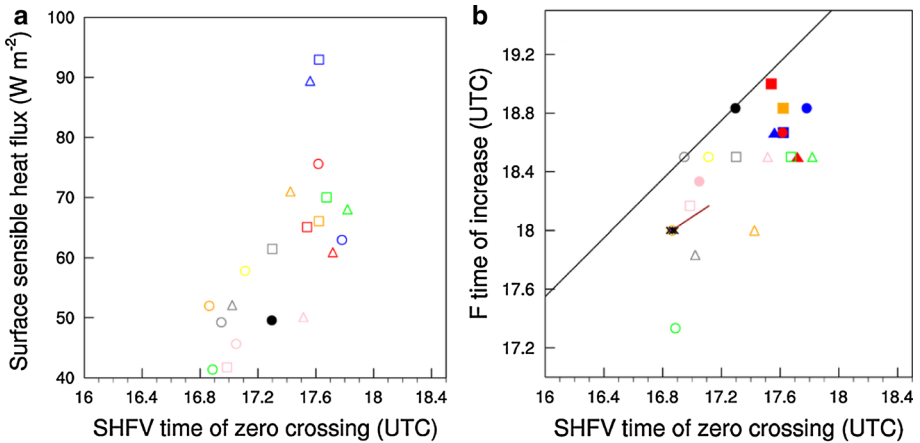
Fig. 12 **a** Mid-afternoon sensible heat flux (1500 UTC) as function of evening virtual sensible heat flux (SHFV) zero crossing time. **b** Time of 200-m wind speed acceleration as function of time of virtual sensible heat flux zero crossing at evening transition. In graph **b** *black line* crosses the observation and represents points with constant delay between the two. Models with coarse vertical resolution have *closed symbols*. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend of model points

surface expressed by the Bowen ratio is likely to influence crossover time, with low Bowen ratios in the afternoon leading to early crossover times. Figure 12a shows the sensible heat flux in the middle of the afternoon (1500 UTC) as a function of the evening crossover time. There is a tendency for models with high sensible heat fluxes to have late crossovers and vice versa. There is little correlation with the sensible heat flux at noon (not shown here), probably because of spin-up issues in the models. The model sensitivity lines are of little use for interpretation here, since during daytime other processes play a role. The latent heat flux in the middle of the afternoon showed no clear correlation with the crossover time, showing that other processes, such as differences in the evolution of temperature and humidity.

The flow acceleration at 200 m, which starts in the late afternoon (see Fig. 3c), is caused by the development of a stable layer with reduced turbulent exchange. Therefore, we expect a delay between the time of the crossover and the start of the acceleration, because of momentum decoupling and inertial oscillation. The start of the acceleration is defined here as the time at which 200-m wind speed has increased by 1 m s^{-1} relative to its value at 1500 UTC. Figure 12b shows that the observed delay is 1.5 h, while most models show shorter delay times, ranging between 0.5 and 1.5 h. The black line represents points with the same delay time as observed. The sensitivity lines show that thermal *coupling* plays a role in determing the delay time. Closed symbols represent models with coarse vertical resolution: these models tend to have late times of crossover and a tendency for the acceleration of the flow at 200 m to begin later.

5.5 Morning Transition

Angevine and Klein Baltink (2001) distinguish between the time of heat flux crossover and the time of the onset of convective turbulence at 200 m. Since many of the models do not have turbulent kinetic energy as a prognostic variable we use the 200-m temperature as an indicator of the onset. The time series of the observed 200-m temperature is shown in Fig. 13a. They have a marked minimum in the morning, which is related to the mixing of low-level cold air, when the entrainment zone of the mixed layer reaches the 200-m level. Here, we use the
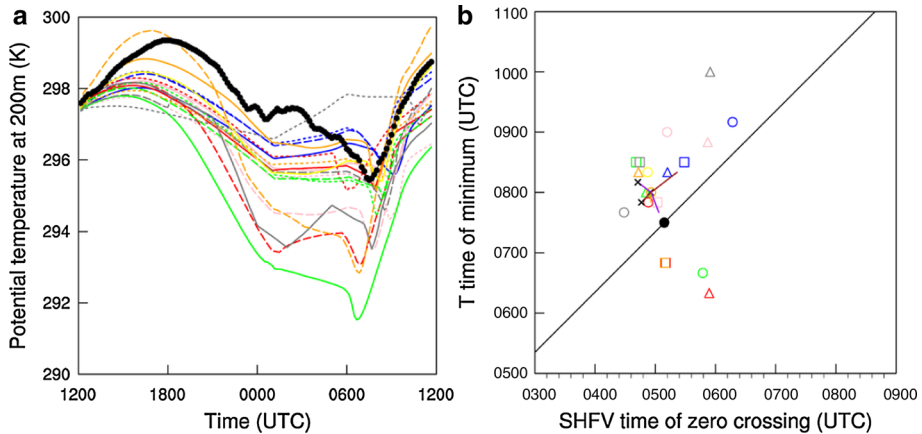
Fig. 13 **a** Time series of 200-m potential temperature. **b** time of minimum temperature at 200 m as function of virtual sensible heat flux (SHFV) zero crossing time during morning transition. The *black line* are points with the same delay between times as observed. Colouring of sensitivity lines as in Fig. 6. See Table 2 for legend

time when the 200-m temperature reaches its last minimum before noon as a substitute for the time of onset. Figure 13b shows the time of onset as a function of the morning crossover time. Sunrise is at 0330 UTC and the observed crossover time is 1.5 h later. The models show a 2-h spread around the observed value. The observed delay of the onset relative to the time of the crossover is more than 2 h. The black line in the process diagram connects points with the same delay as observed. Most models show a delay of more than 3 h. A minority of the models show a significantly less delay. None of the classifications used in the preceding process diagrams, like coupling, mixing and resolution, gives any structure to this diagram. Also, none of the sensitivity lines is of any significant length.

During typical summer days at Cabauw, the 200-m wind-speed in the morning exhibits a minimum caused by the inertial oscillation and by momentum recoupling of the atmospheric layers, when the stability in the lower atmospheric layer decreases after sunrise. Subsequently, a slight recovery of the wind speed often occurs as the CBL grows and high wind speeds aloft are mixed downward. Inspection of comparable mid-summer days at Cabauw shows that the 200-m speed wind-speed minimum generally occurs around 0900 UTC, as is observed in the current case. Figure 3c shows that wind speeds at 200 m vary substantially among models during this transition period. Any sign of a wind-speed minimum seems to be hidden beneath the tendency of the models to display a variation in the phase of oscillation of the 200-m wind speed during and after the transition period. Evaluation against observations is problematic here, since the actual dynamical tendencies of momentum during the morning transition are uncertain (Bosveld et al. 2014).

6 Discussion and Conclusions

Here we have compared a variety of numerical model boundary-layer schemes with Cabauw data. The Cabauw site with its flat and homogeneous terrain and its long observational record has enabled the selection of a relatively ideal 24-h case. By carefully prescribing the temporal evolution of the forcings to the vertical column, the single column model (SCM) results could be compared to observations in a meaningful way. The models are able to reproduce the gross

features of the observed SBL including the onset of momentum decoupling, signature of 2-m air temperature over time, and evolution of the inertial oscillation.

To unravel the interactions between the main processes in the SBL (coupling, radiation, mixing) specific analyses were performed by using process diagrams. These diagrams show relations between different parameters of the models, together with the observations. In the same diagram, models can be classified in terms of common characteristics. At the same time, results of sensitivity runs with one of the models (RACMO) are displayed contributing to the interpretation of variations among models. To ensure that the conclusions based on these sensitivity runs are on a firm basis a second model (D91) was also operated with variations in relevant, but of course different, model parameters to derive sensitivities with approximately the same amplitudes as in RACMO in the defining process diagrams, Figs. 6, 7 and 9. In these diagrams the other, supposedly irrelevant, sensitivity lines indeed had small amplitudes for both RACMO and D91 and their orientations were also comparable.

As a first step, the models are characterized in terms of land-surface thermal coupling, radiation and turbulent mixing by looking at parameter relations that isolate a specific process. For thermal coupling the relation between skin-temperature variation in the evening and surface soil heat flux around midnight was used. For radiation an equation was derived that relates longwave downward radiation at the surface with surface temperature, 200-m air temperature and 200-m humidity. To obtain a fair comparison of the radiation schemes this equation was used to correct for the influence of different thermodynamic profiles among models. A turbulent mixing length scale was derived as a function of gradient Richardson number for conditions around midnight at 60 m above the surface. For each of the three processes a variation among the models was found.

Traditionally, the focus of SBL modelling has been on turbulence parametrization. However, the results from GABLS3 show that thermal coupling to the soil has the largest influence on the variation of 2-m minimum temperature among models for this fully turbulent case. Models with a skin layer in their land-surface scheme tend to show a more realistic thermal coupling, although several models without a skin layer show similar behaviour. The weak observed thermal coupling is in contrast to the strong thermal coupling inferred from the relation between 2-m minimum temperature and longwave radiation. This hints at an underestimation of the observed soil heat flux.

Variations among models in surface longwave downward radiation during the night are substantial and dominated by the variations in the thermodynamic profiles as simulated by the models. A large portion of the surface longwave radiation originates from within the SBL, showing that SBL dynamics play an important role in the longwave radiation received at the surface. On the other hand, differences in the longwave radiation schemes among models seem to play a minor role for this case where turbulent transport dominates the exchange in the SBL. Note that in many models the same longwave radiation scheme is used (RRTM). Apart from the influence of the thermodynamic profiles the radiation schemes perform within a few W m^{-2} of the observed surface longwave downward radiation.

Models with strong turbulent mixing show too deep a boundary layer, too little turning of the wind, too large a downward sensible heat flux and too small a wind-speed maximum at 200 m. Height of maximum wind speed is greater for strongly mixing models and the maximum jet speed shows less variation than the maximum wind speed at 200 m. These conclusions are similar to those found for the previous GABLS studies (Cuxart et al. 2006; Svensson and Holtslag 2009; Svensson et al. 2011; Holtslag et al. 2013). The correlation between boundary-layer height and sensible heat flux leads to a weak dependence of minimum temperature on mixing intensity. Our definition of the boundary-layer height, which is based on the temperature profile, emphasises differences among models with respect to the mixing

of heat. Since many models use different mixing efficiencies for momentum and heat this may obscure deviations resulting from unrealistic momentum mixing. As an example, the model C31R1, which has significantly stronger momentum mixing than the mixing of heat, gives a reasonable boundary-layer height but its maximum wind speed at 200 m is close to the lowest.

The timing of the evening transition varies by 1.5 h among the models, with models having coarse resolution tending to have later transition times. At the evening transition the observed 200-m wind speed increases more rapidly than the modelled speeds. Modelled delay times between the evening transition and the start of the wind speed increase at 200 m are significantly smaller than observed. The surface humidity peaks at higher values around the evening transition in models than in the observations. This peak is determined by the subtle interplay between the fast declining evaporation rate and the appearance and growth of the SBL. The most probable cause of these peaks is that modelled latent heat fluxes are larger than the observed flux during the evening transition.

Modelled times of the morning transition crossover show a spread of 2 h around the observed value. The modelled times of the 200-m temperature minimum, taken as a proxy for the onset of convection, shows an even larger variation. The observed delay between the transition and the onset is 2.4 h, while most models have a delay of 3 h, suggesting excessively slow initial development of the CBL. No clear relationship between the variations in the two transition times is found amongst the models, neither is there any clear relationship with any of the model classifications used here (coupling, mixing, and resolution). Around the time of the morning transition, the 200-m wind speed varies greatly among models and also shows significant oscillations with varying phases. Simulations with the KNMI regional climate model, as described by Bosveld et al. (2014), did not show such oscillations in the 200-m wind speed. Also, the nine observed Cabauw cases described in the same paper did not show such behaviour, suggesting that it may be an artefact of the SCM simulations. Note that in a three-dimensional model, the neighbouring columns are coupled to each other by pressure variations induced by horizontal wind differences between the columns and by horizontal diffusion, probably suppressing the tendency of a single atmospheric column in such a model to exhibit inertial oscillations. In the SBL in the real atmosphere, the importance of such horizontal diffusion may be smaller than in these models, but some interaction through horizontal pressure and wind variations will certainly be present. These effects are absent in an SCM. Given the uncertainties in the prescribed dynamical tendencies during the morning transition and the long integration from initialization over different challenging regimes of the ABL, we must conclude that a definite evaluation of the morning transition based on the material of the current case is not possible. Differences in convective turbulence schemes between the models will also play a role in determining the timing of the morning transition. In the light of the limitations mentioned, this topic is not further pursued here. For an evaluation of the morning transition it is advisable to set the initialization time much closer to the moment of transition. Given these limitations no evaluation could be done of the findings of Angevine and Klein Baltink (2001) and Beare (2008) who show that shear over the entrainment layer plays an important role in the early stages of the convective boundary layer.

In cases with lower geostrophic winds, the relative sensitivities of the parameters relevant to thermal coupling, mixing and radiation may change. Thermal coupling, as defined here, is a model concept; the thermal coupling needed for optimal performance depends on the model structure (McNider et al. 2005). The significant differences in sensitivities found among models in this GABLS3 study are consistent with the conclusions of McNider et al. (2012) that the modelled SBL is at present is uncertain and that the modelled trends in surface parameters should be viewed with caution.

Here, we have shown that, by carefully selecting a case and carefully prescribing the forcings on the atmospheric column, it is possible to constrain the models in such a way that a useful comparison with observations is possible. Dynamical tendencies derived from numerical weather forecasting models are currently not accurate enough to be used for inter-comparisons based on a single case. But, by using composite model tendencies, formed over many comparable days, Baas et al. (2010) showed that a fruitful evaluation of SCM models with observations is possible. The advent of test beds where continuous model simulations and observations come together in a routine setting may contribute to this type of evaluative study (Neggers et al. 2012).

In hindsight, the case description could have been improved. In particular, models without a skin-layer show a significant spin-up of the soil heat flux during the first two hours after initialization. This also influences the temperature of the CBL. In such cases, this problem could have been mitigated by performing a preliminary simulation to set the soil profiles: the final soil profiles from the preliminary simulation would be taken as the initial profiles for the main simulation. The initial imbalance also hampered the tuning of the models to the prescribed Bowen ratio. It might have been better to set the Bowen ratio on the basis of the values of sensible and latent heat fluxes at 1500 UTC. A drawback of tuning the Bowen ratio is that the soil water may deviate from its actual value, with consequences for the thermal properties of the soil and thus for soil thermal coupling. Model profiles of upward and downward longwave radiation were not available for intercomparison; this information would have facilitated a more detailed analysis of the role of longwave radiation divergence in the SBL.

Appendix: A Simple Relation for Surface Longwave Radiation Including the Influence of Temperature Lapse Rate

To separate the effect of variations in thermodynamic profiles and the radiation schemes, we developed a parametrization of longwave downward radiation at the surface ($L^{\downarrow}$) that accounts for the characteristics of the SBL. This is based on the relation of Brunt (1932), who parametrizes $L^{\downarrow}$ in terms of the 2-m air temperature and humidity. Dürr and Philipona (2004) show that the temperature lapse rate in the lowest layers of the atmosphere has a significant impact on $L^{\downarrow}$. This is confirmed by experiments with one of the participating models (D91), where it was found that during the night of the third GABLS case approximately 50 % of $L^{\downarrow}$ originates from the lowest 200 m of the atmosphere. Ohmura (2001) arrived at even higher fractions, based on climatological thermodynamic profiles and a detailed radiation model. Here, we modify Brunt's relation to account for the temperature difference between 200 m and 2 m in terms of black-body radiation; this leads to a reference radiative flux,

$$L^{\downarrow}_{ref} = (a + b\sqrt{e_{200}})\sigma T_2^4 + (c + d\sqrt{e_{200}})\sigma(T_{200}^4 - T_2^4) + f, \qquad (4)$$

where e_{200} is the water vapour pressure in hPa at 200 m, T_2 and T_{200} are the absolute temperatures, at 2 and 200 m respectively, σ is the Stefan-Boltzman constant and a, b, c, d and f are regression coefficients. Notice that the latter coefficients are expected to depend on the site and the case.

We used a test set of cloud free episodes based on Boers et al. (2010), who studied different cloud-detection methods employed at Cabauw to estimate cloud fraction. They concluded that nubiscope observations were the best choice for low and mid-level clouds. By combining a nubiscope with a cloud radar, which is sensitive to cirrus clouds, a stringent test to determine clear-sky episodes can be made. We used their dataset to find approximately 2400 10-mm clear-sky periods in the period from 15 May 2008 to 14 May 2009. The modified Brunt relation was derived by trying several alternatives, among them using water vapour pressure at 2 m in the first or second term. The use of e_{200} in both terms gave the highest explained variance. No improvement was found from allowing for curvature in the temperature profile by introducing an extra regression term involving the 80-m temperature.

We found the following optimized values of the regression coefficients: $a = 0.640 \pm 0.008$; $b = 0.0499 \pm 0.0010$; $c = 1.003 \pm 0.030$; $d = -0.209 \pm 0.010$ and $f = -3.1 \pm 2.1$. The standard deviation of the residuals was 4.1 W m^{-2}. Note that d is negative, which shows that, at high levels of humidity, the 200-m temperature is less important in determining $L^{\downarrow}$. At a typical water vapour pressure of 16 hPa, the value of $c + d\sqrt{e_{200}}$, multiplying the temperature-difference term in Eq. 4 is 0.2. This relatively small factor confirms that emission from levels close to the surface make a significant contribution to $L^{\downarrow}$.

References

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

Andrén A (1990) Evolution of a turbulence closure scheme suitable for air-pollution applications. J Appl Meteorol 29:224–239

Angevine WM, Klein Baltink H, Bosveld FC (2001) Observations of the morning transition of the convective boundary layer. Boundary-Layer Meteorol 101:209–227

Angevine WM, Jiang H, Mauritsen T (2010) Performance of an eddy diffusivity–mass flux scheme for shallow cumulus boundary layers. Mon Weather Rev 138:2895–2912

Baas P, Steeneveld GJ, van de Wiel BJH, Holtslag AAM (2006) Exploring self-correlation in stably stratified conditions. J Atmos Sci 63:3045–3054

Baas P, de Roode SR, Lenderink G (2008) The scaling behaviour of a turbulent kinetic energy closure model for stably stratified conditions. Boundary-Layer Meteorol 127:17–36

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 136:671–684

Baldauf M, Seifert A, Förstner J, Majewski D, Raschendorfer M, Reinhardt T (2011) Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Mon Weather Rev 138(12):3887–3905

Basu S, Holtslag AAM, van de Wiel BJH, Moene AF, Steeneveld G-J (2008) An inconvenient "truth" about using sensible heat flux as a surface boundary condition in models under stably stratified regimes. Acta Geophysica 56:88–99

Bazile E, Marquet P, Bouteloup Y, Bouyssel F (2011) The Turbulent Kinetic Energy (TKE) scheme in the NWP models at Météo-France. ECMWF Workshop proceedings on "Diurnal cycles and the stable boundary layer", 7–10 November 2011:127–136 available at www.ecmwf.int

Beare RJ (2008) The role of shear in the morning transition boundary layer. Boundary-Layer Meteorol 129:395–410

Beare RJ, Edwards JM, Lapworth AJ (2006) Simulation of the observed evening transition and nocturnal bounday layers: large-eddy simulation. Q J R Meteorol Soc 132:81–99

Bélair S, Mailhot J, Strapp JW, MacPherson JI (1999) An examination of local versus nonlocal aspects of a TKE-based boundary-layer scheme in clear convective conditions. J Appl Meteorol 38:1499–1518

Beljaars ACM (1982) The derivation of fluxes from profiles in perturbed areas. Boundary-Layer Meteorol 24:35–55

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

Beljaars ACM, Viterbo P (1998) Role of the boundary layer in a numerical weather prediction model. In: Holtslag AAM, Duynkerke PG (eds) Clear and cloudy boundary layers. Royal Netherlands Academy of Arts and Sciences, Amsterdam, p 372

Betts AK, Ball JH, Beljaars ACM (1993) Comparison between the land surface response of the European Centre model and the FIFE-1987 data. Q J R Meteorol Soc 119:975–1001

Betts AK, Ball JH, Beljaars ACM, Miller MJ, Viterbo A (1996) The land surface-atmosphere interaction, a review based on observational and global modeling perspectives. J Geophys Res 1001:7209–7225

Boers R, de Haij MJ, Wauben WMF, Baltink HK, van Ulft LH, Savenije M, Long CN (2010) Optimized fractional cloudiness determination from five ground-based remote sensing techniques. J Geophys Res 115:D24116

Bosveld FC, Baas P, van Meijgaard E, de Bruijn EIF, Steeneveld G-J, Holtslag AAM (2014) The third GABLS intercomparison case for evaluation studies of boundary layer models: Part A: Case Selection and Set-up. Boundary-Layer Meteorol. doi:10.1007/s10546-014-9917-3

Brunt D (1932) Notes on radiation in the atmosphere: I. Q J R Meteorol Soc 58:389–420

Cox PM, Betts RA, Bunton CB, Essery RLH, Rowntree PR, Smith J (1999) The impact of new land surface physics on the GCM simulation of climate and climate sensitivity. Clim Dyn 15:183–203

Cuxart J, Holtslag AAM, Beare RJ, Bazile E, Beljaars ACM, Cheng A, Conangla L, Ek MB, 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 KM (2006) Single-column model intercomparison for a stably stratified atmospheric boundary layer. Boundary-Layer Meteorol 118:273–303

de Roode SR, Bosveld FC, Kroon PS (2010) Dew formation, eddy-correlation latent heat fluxes, and the surface energy imbalance at Cabauw during stable conditions. Boundary-Layer Meteorol 135:369–383

Derbyshire S (1999) Stable boundary-layer modelling established approaches and beyond. Boundary-Layer Meteorol 90:423–446

Dürr B, Philipona R (2004) Automatic cloud amount detection by surface longwave downward radiation measurements. J Geophys Res 109:D05201

Duynkerke PG (1991) Radiation fog: a comparison of model simulation with detailed observations. Mon Weather Rev 119:324–341

ECMWF (2006) Integrated Forecast System documentation of cycle 31r1. Model documentation of the European Centre for Medium Range Weather Forecast model (www.ecmwf.int), http://www.ecmwf.int/research/ifsdocs/CY31r1/index.html

Edwards JM (2009) Radiative processes in the stable boundary layer: part I radiative aspects. Boundary-Layer Meteorol 131:105–126

Edwards JM, Beare RJ, Lapworth AJ (2006) Simulation of the observed evening transition and nocturnal boundary layers. Single column modelling. Q J R Meteorol Soc 132:61–80

Ek M, Mitchell KE, Lin Y, Rogers E, Grunmann P, Koren V, Gayno G, Tarpley JD (2003) Implementation of NOAH landsurface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model. J Geophys Res 108(D22):8851. doi:10.1029/2002JD003296

Estournel C, Guedalia D (1985) Influence of geostrophic wind on atmospheric nocturnal cooling. J Atmos Sci 42(23):2695–2698

Foken T (2008) The energy balance closure problem: an overview. Ecol Appl 18(6):1351–1367

Fu Q, Liou KN (1992) On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres. J Atmos Sci 49:2139–2156

Garratt JR (1994) The atmospheric boundary layer. Cambridge University Press, UK 316 pp

Golaz J-C, Larson VE, Cotton WR (2002) A PDF-based model for boundary layer clouds. Part I: method and model description. J Atmos Sci 59:3540–3551

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

Grachev AA, Andreas EL, Fairall CW, Guest PS, Persson POG (2007) On the turbulent Prandtl number in the stable atmospheric boundary layer. Boundary-Layer Meteorol 125:329–341

Grant ALM (1997) An observational study of the evening transition boundary layer. Q J R Meteorol Soc 123:657–677

Gultepe I, Pearson G, Milbrandt JA, Hansen B, Platnick S, Taylor P, Gordon M, Oakley JP, Cober SG (2009) The Fog remote sensing and modeling field project. Bull Am Meteorol Soc 90:341–359

Ha K-J, Mahrt L (2003) Radiative and turbulent fluxes in the nocturnal boundary layer. Tellus 55A:317–327

Holtslag AAM (2006) GEWEX Atmospheric Boundary-Layer Study (GABLS) on stable boundary layers. Boundary-Layer Meteorol 118:243–246

Holtslag AAM, Nieuwstadt FTM (1986) Scaling the atmospheric boundary layer. Boundary-Layer Meteorol 36:201–209

Holtslag AAM, Steeneveld GJ, van de Wiel BJH (2007) Role of land-surface temperature feedback on model performance for the stable boundary layer. Boundary-Layer Meteorol 125:361–376

Holtslag AAM, Svensson G, Baas P, Basu S, Beare B, Beljaars ACM, Bosveld FC, Cuxart J, Lindvall J, Steeneveld GJ, Tjernström M, Van De Wiel BJH (2013) Stable atmospheric boundary layers and diurnal cycles—challenges for weather and climate models. Bull Am Meteorol Soc. doi:10.1175/BAMS-D-11-00187.1

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

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

Jacobs AFM, Heusinkveld BG, Holtslag AAM (2008) Towards closing the surface energy budget of a mid-latitude grassland. Boundary-Layer Meteorol 126:125–136

Janjic ZI (2002) Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP mesoscale model. NCEP Office Note 437, 61 pp

Larson VE, Golaz J-C (2005) Using probability density functions to derive consistent closure relationships among higher-order moments. Mon Weather Rev 133:1023–1042

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

Mahrt L (1981) The early evening boundary layer transition. Q J R Meteorol Soc 107:329–343

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

Martin GM, Ringer MA, Pope VD, Jones A, Dearden C, Hinton TJ (2006) The physics properties of the atmosphere in the new Hadley Centre global environment model (HadGEM1). Part I: model description and global climatology. J Clim 19:1274–1301

McNider RT, Lapenta WM, Biazar AP, Jedlovec GJ, Suggs RJ, Pleim J (2005) Retrieval of model grid-scale heat capacity using geostationary satellite products. Part I: first case-study application. J Appl Meteorol 44:1346–1360

McNider RT, Steeneveld G-J, Holtslag AAM, Pielke Sr. RA, Mackaro S, Pour-Biazar A, Walters J, Nair U, Christy J (2012) Response and sensitivity of the nocturnal boundary layer over land to added long wave radiative forcing. J Geophys Res 117:D14106

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 102(D14):16663–16682

Nappo CJ, Millar DR, Hiscox AL (2008) Wave-modified flux and plume dispersion in the stable boundary layer. Boundary-Layer Meteorol 129:211–223

Neggers RAJ, Siebesma AP, Heus AT (2012) Continuous single-column model evaluation at a permanent meteorological supersite. Bull Am Meteorol Soc 93(9):1389–1400

Niemelä S, Räisänen P, Savijärvi H (2001) Comparison of surface radiative flux parameterizations Part I: longwave radiation. Atmos Res 58:1–18

Nieuwstadt FTM, Brost RA (1986) The decay of convective turbulence. J Atmos Sci 43:532–546

Noilhan J, Planton S (1989) A simple parametrization of land surface processes for meteorological models. Mon Weather Rev 117:536–549

Ohmura A (2001) Physical basis for the temperature-based melt-index method. J Appl Meteorol 40:753–761

Oncley SP, Foken T, Vogt R, Kohsiek W, DeBruin HAR, Bernhofer C, Christen A, van Gorsel E, Grantz D, Feigenwinter C, Lehner I, Liebethal C, Liu H, Mauder M, Pitacco A, Ribeiro L, Weidinger T (2007) The Energy Balance Experiment EBEX-2000. Part I: overview and energy balance. Boundary-Layer Meteorol 123:1–28

Pino D, Jonker HJJ, Vilà-Guerau de Arellano J, Dosio A (2006) Role of shear and the inversion strength during sunset turbulence over land, characteristic length scales. Boundary-Layer Meteorol 121:537–556

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 46:1383–1395

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 46:1396–1409

Porté-Agel F, Wu Y-T, Lu H, Conzemius RJ (2011) Large-eddy simulation of atmospheric boundary layer flow through wind turbines and wind farms. J Wind Eng Indus Aerodyn 99(4):154–168

Prabha T, Hoogenboom G (2008) Evaluation of the weather research and forecasting model for two frost events. Comput Electron Agric 64:234–247

Räisänen P (1996) The effect of vertical resolution on clear-sky radiation calculations: test with two schemes. Tellus 48A:403–423

Richardson D (2009) Forecast Products Users Meeting, June 2009, ECMWF. ECMWF Newsletter 120:8–9

Santanello JA, Peters-Lidard CD, Kumar SV, Alonge C, Tao W-K (2009) A modelling and observational framework for diagnosing local land–atmosphere coupling on diurnal time scales. J Hydrometeorol 10:577–599

Savijärvi H (2006) Radiative and turbulent heating rates in the clear-air boundary layer. Q J R Meteorol Soc 132:147–161

Seity Y, Brousseau P, Malardel S, Hello G, Benard P, Bouttier F, Lac C, Masson V (2011) The AROME-France convective-scale operational model. Mon Weather Rev 139:976–991

Sellers JP, Randell DA, Collatz GJ, Berry JA, Field CB, Dazlich DA, Zhang C, Collelo GD, Bounua L (1996) A revised land surface parametrization (SiB 2) for atmospheric GCMs. Part I: model formulation. J Clim 9:676–705

Skamarock WC, and co-authors (2008) A description of the advanced research WRF Version 3. NCAR Technical Note NCAR/TN-475+STR

Smedman A-S (1988) Observations of a multi-level turbulence structure in a very stable atmospheric boundary layer. Boundary-Layer Meteorol 44:231–253

Sorbjan Z (2010) Gradient-based scales and similarity laws in the stable boundary layer. Q J R Meteorol Soc 136:1243–1254

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

Steeneveld GJ, Mauritsen T, De Bruijn EIF, Svensson G, Holtslag AAM (2008) Evaluation of limited-area models for the representation of the diurnal cycle and contrasting nights in CASES-99. J Appl Meteorol 47:867–889

Steeneveld GJ, Wokke MJJ, Groot Zwaaftink CD, Pijlman S, Heusinkveld BG, Jacobs AFG, Holtslag AAM (2010) Observations of the radiation divergence in the surface layer and its implication for its parameterization in numerical weather prediction models. J Geophys Res 115:D06107

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 118:1199–1217

Svensson G, Holtslag AAM (2009) Analysis of model results for the turning of the wind and related momentum fluxes in the stable boundary layer. Boundary-Layer Meteorol 132:261–277

Svensson G, Holtslag AAM, Kumar V, Mauritsen T, Steeneveld GJ, Angevine WM, Bazile E, Beljaars A, de Bruijn EIF, Cheng A, Conangla L, Cuxart J, Ek M, Falk MJ, Larson VE, Lock A, Mailhot J, Masson V, Park S, Pleim J, Söderberg S, Zampieri M, Weng W (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

Tjemkes SA, Duynkerke PG (1989) The nocturnal boundary layer: model calculations compared with observations. J Appl Meteorol 28(3):161–175

van de Wiel BJH, Moene AF, Hartogensis OK, de Bruin HAR, Holtslag AAM (2003) Intermittent turbulence in the stable boundary layer over land. Part III: a classification for observations during CASES-99. J Atmos Sci 60:2509–2522

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 67:2679–2689

van Meijgaard E, van Ulft LH, van de Berg WJ, Bosveld FC, van den Hurk BJJM, Lenderink G, Siebesma AP (2008) The KNMI regional atmospheric climate model RACMO, version 2.1. KNMI publication: TR-302, 24/12/2008, pp43. De Bilt, The Netherlands

van Ulden AP, Wieringa J (1996) Atmospheric boundary layer research at Cabauw. Boundary-Layer Meteorol 78:39–69

Verkaik JW, Holtslag AAM (2007) Wind profiles, momentum fluxes and roughness lengths at cabauw revisited. Boundary-Layer Meteorol 122:701–719

Viterbo P, Beljaars ACM (1995) An improved land surface parametrization scheme in the ECMWF model and its validation. J Clim 8:2716–2748

Walsh JE, Chapman WL, Romanovsky V, Christensen JH, Stendel M (2008) Global climate model performance over Alaska and Greenland. J Clim 21:6156–6174

Weng W, Taylor PA (2003) On modelling the one-dimensional atmospheric boundary layer. Boundary-Layer Meteorol 107:371–400

Zilitinkevich SS, Elperin T, Kleeorin N, Rogachevskii I (2007) Energy-and flux-budget (EFB) turbulence closure model for stably stratified flows. Part I: steady-state, homogeneous regimes. Boundary-Layer Meteorol 125:167–191