Results of the GABLS3 diurnal-cycle benchmark for wind energy applications

J Sanz Rodrigo[1], D Allaerts[9], M Avila[4], J Barcons[4], D Cavar[5], RA Chávez Arroyo[1], M Churchfield[2], B Kosovic[3], JK Lundquist[2,6], J Meyers[7], D Muñoz Esparza[3], JMLM Palma[8], JM Tomaszewski[6], N Troldborg[5], MP van der Laan[5] and C Veiga Rodrigues[8]

[1]National Renewable Energy Centre, Sarriguren, Spain
[2]National Renewable Energy Laboratory, Boulder, CO, U.S.A.
[3]National Centre for Atmospheric Research, Boulder, CO, U.S.A.
[4]Barcelona Supercomputing Centre, Barcelona, Spain
[5]Technical University of Denmark, Roskilde, Denmark
[6]University of Colorado Boulder, Boulder, CO, U.S.A.
[7]University of Leuven, Leuven, Belgium
[8]University of Porto, Porto, Portugal

E-mail: jsrodrigo@cener.com

Abstract. We present results of the GABLS3 model intercomparison benchmark revisited for wind energy applications. The case consists of a diurnal cycle, measured at the 200-m tall Cabauw tower in the Netherlands, including a nocturnal low-level jet. The benchmark includes a sensitivity analysis of WRF simulations using two input meteorological databases and five planetary boundary-layer schemes. A reference set of mesoscale tendencies is used to drive microscale simulations using RANS k-ε and LES turbulence models. The validation is based on rotor-based quantities of interest. Cycle-integrated mean absolute errors are used to quantify model performance. The results of the benchmark are used to discuss input uncertainties from mesoscale modelling, different meso-micro coupling strategies (online vs offline) and consistency between RANS and LES codes when dealing with boundary-layer mean flow quantities. Overall, all the microscale simulations produce a consistent coupling with mesoscale forcings.

1. Introduction

The increasing growth of wind turbines, with rotor tip heights approaching 200 m, and wind farm clusters extending for tens of kilometers, is pushing the wind farm flow modeling community to consider effective ways of integrating forcing from large meteorological processes in the simulation of the flow at wind farm scale based on computational fluid dynamic (CFD) models. The dynamics of these forcings determine the interplay between the wind climatology, relevant for the assessment of the wind resource, and the wind conditions relevant for wind turbine siting.

A recent review of state-of-the-art methodologies for mesoscale-to-microscale modeling is provided in Sanz Rodrigo et al [1]. Outstanding challenges include: coupling of codes dealing with fundamental differences in terms of physical hypothesis and numerical methods, lack of suitable parameterization in the "terra incognita" [2] that links mesoscale and microscale turbulence processes, and lack of a systematic evaluation procedure that can identify the source of modeling errors. Additionally, for the wind energy community, design conditions are almost entirely based on idealized turbulence and boundary-layer models rather than recognizing and classifying wind conditions that account for mesoscale forcing in design standards.

With these challenges in mind, the boundary-layer meteorology community has conducted a series of model intercomparison GEWEX Atmospheric Boundary Layer Studies (GABLS) [3]. The third GABLS benchmark deals with a diurnal cycle observed at the Cabauw meteorological tower in the Netherlands, under relatively stationary synoptic conditions, that lead to the development of a strong nocturnal low-level jet (LLJ) [4][5]. Model intercomparison results for single-column models (SCM) are reported in [6].

This benchmark has been revisited by the wind energy community to help design meso-micro methodologies for resource assessment and design tools, in particular: incorporating time- and height-dependent mesoscale forcing in microscale models, turbulence modeling at varying atmospheric stability conditions, defining suitable surface boundary conditions for momentum and heat and characterization of the wind profile in (non-logarithmic) LLJ conditions [7]. Initial results, using a single-column model (SCM) as proxy for microscale 3D models, show how adding more realistic mesoscale forcing systematically leads to better performance considering cycle-aggregated rotor-based quantities of interest. Still, a large bias is observed in the hour-to-hour evolution of the vertical wind profile due to the inherent input uncertainty from the mesoscale model [8][9].

This paper summarizes the results of the Windbench/GABLS3 model intercomparison benchmark, discussing three topics: characterization of uncertainties in the input forcing, methodologies for meso-micro coupling, and consistency of turbulence models at different fidelity levels.

2. The Windbench/GABLS3 Benchmark

The GABLS3 original set-up is described in Bosveld et al. [5]. The case analyzes the period from 12:00 UTC 1 July to 12:00 UTC 2 July 2006, at the Cabauw Experimental Site for Atmospheric Research (CESAR), located in the Netherlands (51.971°N, 4.927°E). The elevation of the site is approximately -0.7 m, surrounded by relatively flat terrain characterized by grassland, fields and some scattered tree lines and villages. The roughness length for the wind direction sector of interest (60° - 120°) is 15 cm.

The CESAR measurements are carried out at a 200-m tower, free of obstacles up to a few hundred meters in all directions. The measurements include 10-min averaged vertical profiles of wind speed, wind direction, temperature and humidity at heights 10, 20, 40, 80, 140 and 200 m, as well as surface radiation and energy budgets. Turbulence fluxes are also monitored at four heights: 3, 60, 100 and 180 m. A RASS profiler measures wind speed, wind direction and virtual temperature above 200 m.

The original GABLS3 setup uses a simplified set of forcing terms that was obtained by piecewise linear approximations of the mesoscale tendencies in RACMO numerical weather prediction model to obtain a better agreement of the wind speed at 200m. In our model intercomparison we will use directly the input forcings derived from mesoscale, as described in the next section, without calibration against observations. This is in favor of testing a general methodology for offline coupling of mesoscale and microscale models.

The GABLS3 benchmark for wind energy is published as a Windbench repository in [7].

2.1. Reference WRF simulation: Input Data for Microscale Models

Input forcing for most of the microscale models (excepting VENTOS®/M) is derived from a mesoscale simulation using the Weather Research and Forecasting model WRF-ARW v3.8 [10]. Following previous work from Kleczek et al [11], the reference WRF simulation is based on a one-way nesting configuration of three concentric square domains centered at the Cabauw site, of the same size (Figure 1) based on a 181x181 points grid with 9, 3 and 1 km horizontal resolution. The vertical grid, approximately 13 km high, is based on 46 terrain-following (eta) levels with 24 levels in the first 1000 m, the first level at approximately 13 m, a uniform spacing of 25 m over the first 300 m and then stretched to a uniform resolution of 600 m in the upper part. The U.S. Geological Survey (USGS) land-use surface data, that comes by default with the WRF model, is used together with the unified Noah land-surface model to define the boundary conditions at the surface. Other physical parameterizations used are: the rapid radiative transfer model (RRTM), the Dudhia radiation scheme and the Yonsei University (YSU) first-order PBL scheme [12]. The simulation uses input data from

ERA-Interim [13] with a spin-up time of 24 hours. The input forcing of VENTOS®/M differs from the other microscale models and is discussed below.

Time and height-dependent mesoscale forcing is extracted directly from the momentum budget components (so-called "tendencies" in WRF) after filtering out fluctuations at the grid-resolution scale by spatially averaging data from a 9-km wide 3x3 grid around the site and using a 1-hour rolling window on the time series. More details about this meso-micro coupling method and a sensitivity analysis of rotor-averaged forcing on the filter set-up are discussed in [9].

2.2. Validation Data and Metrics

The performance of the models is based on quantities of interest relevant for wind energy applications. These quantities are evaluated across a reference rotor span of 160 m, between 40 and 200 m, characteristic of an 8-MW large wind turbine. Besides hub-height wind speed S_{hub} and direction WD_{hub}, it is relevant to consider the rotor-equivalent wind speed $REWS$, the turbulence intensity (not evaluated here), the wind speed shear α, and the wind direction shear or veer ψ.

The $REWS$ is especially suitable to account for wind shear in wind turbine power performance tests [14]. The $REWS$ is the wind speed corresponding to the kinetic energy flux through the swept rotor area, when accounting for the vertical shear:

$$REWS = \left[\frac{1}{A} \sum_i \left(A_i S_i^3 \cos \beta_i \right) \right]^{1/3}$$

(1)

where A is the rotor area and A_i are the horizontal segments that separate vertical measurement points of horizontal wind speed S_i across the rotor plane. The $REWS$ is here weighted by the cosine of the angle β_i of the wind direction WD_i with respect to the hub-height wind direction to account for the effect of wind veer [15].

Wind shear is defined by fitting a power-law curve across the rotor wind speed points S_i:

$$S_i = S_{hub} \left(\frac{z_i}{z_{hub}} \right)^\alpha$$

(2)

Similarly, wind veer is defined as the slope ψ of the linear fit of the wind direction difference:

$$\beta_i = \psi \left(z_i - z_{hub} \right)$$

(3)

To evaluate simulations and measurements consistently, these quantities are obtained after resampling, by linear interpolation, velocity and wind direction vertical profiles at 10 points across the rotor area and then computing the $REWS$ and the shear functional fits. While these fitting functions are commonly used in wind energy, their suitability in LLJ conditions is questionable. The regression coefficient from the fitting can be used to determine this suitability.

A rolling average with a window size of one hour is applied to simulation and observational data to remove the impact of high frequency fluctuations in the analysis.

Validation results are quantified in terms of the mean absolute error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| \chi_{pred} - \chi_{obs} \right|$$

(4)

where χ is any of the above mentioned quantities of interest, predicted (*pred*) or observed (*obs*), and *N* is the number of samples evaluated in the time series.

3. Participating Models

Table1 shows a list of the models participating in the model intercomparison benchmark. The microscale models do not include humidity, which is not relevant in this case since the selection criteria for the cycle excluded wet conditions.

Table 1: Summary of model simulations. Monin-Obukhov similarity theory (MOST) surface boundary conditions use either heat flux (H), 2-m (T_2) or skin temperature (T_{SK}) from WRF

Name	Input	Turbulence	z-Levels	Surface B.C.
WRF-YSU (ref)	ERA Interim	YSU	46	Noah
WRF	ERA Interim, GFS	MYJ, MYNN, QNSE, TEMF, YSU	46	Noah
WRF-YSU_LES	ERA Interim	LES-TKE	101	Noah
WRF-VentosM_ke	ERA Interim	YSU/k-ε	70	MOST, H
CFDWind1D_ke	WRF (ref)	k-ε	301	MOST, T_2
Alya-CFDWind1D_ke	WRF (ref)	k-ε	500	MOST, T_2
Ellipsys1D_ke	WRF (ref)	k-ε	512	MOST, T_{SK}
Ellipsys3D_ke	WRF (ref)	k-ε	192	MOST, T_{SK}
Ellipsys3D_LES	WRF (ref)	*Smagorinsky*	128	MOST, T_{SK}
SP-Wind_LES	WRF (ref)	LES-TKE	500	MOST, T_2

3.1. WRF

We use the Advanced Research Weather forecasting model (WRF) version 3.8.1 [10] configured with three nested domains with grid sizes of 27, 9, and 3 km (Figure 1). All model domains have 61 x 61 horizontal grid points and are centered at the Cabauw tower. We use the same vertical grid as in the WRF-YSU (ref) simulation (see Section 2.1), with 46 eta levels defined to increase the vertical resolution in the lowest part of the planetary boundary layer (PBL). The boundary conditions at the surface are defined by the default USGS land-use surface data and the unified Noah land-surface model. The rapid radiative transfer model (RRTM) and the Dudhia radiation scheme are also used. All simulations are spun-up for 24 hours.

We perform a sensitivity analysis varying PBL schemes and global input data, similar to Kleczek et al. [11]. The first of the selected PBL schemes is Yonsei University (YSU) [12], a non-local, first-order closure scheme. Other PBL schemes used in this study are the Mellor–Yamada–Janjic (MYJ) [16], Mellor–Yamada–Nakanishi–Niino (MYNN2.5) [17], and Quasi-Normal Scale Elimination (QNSE) [18], which are classified as local, 1.5-order turbulent kinetic energy (*TKE*) closure schemes. Total Energy Mass Flux (TEMF) is the final PBL scheme tested and is classified as a hybrid local/non-local,1.5-order closure scheme [19]. We assess the sensitivity of WRF to input data by running each of the five PBL schemes with two

Figure 1: Model domain arrangement, with the red star indicating the center of each domain representing the Cabauw tower.

different sources of input data: ERA-Interim [13] and GFS reanalysis. This ensemble of simulations is used in this study to quantify the spread of mesoscale solutions that could be input into the microscale simulations.
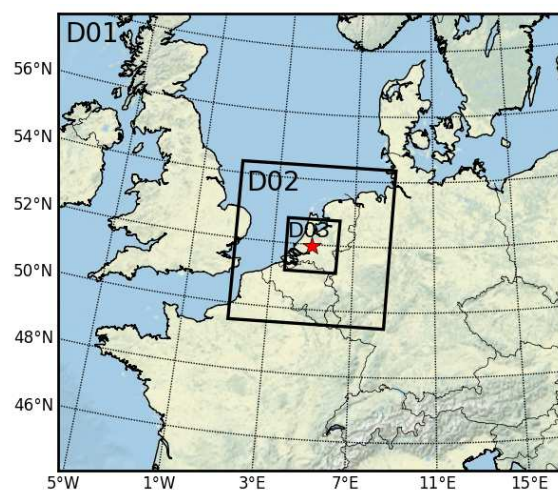
3.2. WRF-LES

Five additional telescopic nests are added to the reference mesoscale set-up of 3 nests to physically downscale, using one-way coupling, to a resolution of 12.5 m in the innermost domain using large-eddy simulation (LES). All domains use a 181x181 horizontal grid. The finest three nests are refined vertically to double the number of levels and obtain an isotropic grid of 12.5 m in the lowest 300 m. A prognostic equation for *TKE* is solved for the LES subgrid-scale model [23], using the WRF default constants [10]. The time step is reduced from 60 s in the outermost domain to 0.2 s in the innermost nest. Since the cycle starts from unstable atmospheric conditions, spin-up is deemed unnecessary. Hence, the first few hours of the cycle are transitioning to reach equilibrium with the mesoscale boundary conditions.

Mean flow profiles are obtained by spatial averaging over a 1-km wide horizontal box. Fluctuations about the mean flow within this box are used to compute the *TKE*.

3.3. VENTOS®/M

The VENTOS®/M computer code is an atmospheric flow solver employing a one-way dynamical coupling methodology, where WRF model simulation results are used as initial and boundary conditions. It solves the URANS equations assuming anelastic fluid, together with a transport equation for potential temperature and a *k-ε* turbulence model [20].

The WRF simulation follows the reference set-up described in Section 3.1 with few differences, namely the software version was WRF-ARW v3.6.1 and a fourth nesting level was added, increasing the downscaling to 1-km horizontal resolution. Physical parameterizations, grid dimensions, nesting grid ratios and the simulation spin-up time were kept equal to the reference simulation.

The VENTOS®/M domain encompasses 12x12 km^2. Centered on Cabauw, the horizontal resolution was kept constant at 160 m within a square area of 3x3 km^2, afterwards expanding towards the boundaries. At each vertical level the horizontal grid was composed of 47x47 control volumes. The vertical grid consisted of 10 km columns with 70 control volumes, expanding from 4 m above the surface to 730 m at the domain top, where a 4-km Rayleigh damping layer was set. The simulation time-step was 1.5 s and the boundary conditions were updated every 5 minutes. The simulation spin-up time was 12 hours, less than the WRF simulation. Similarly to WRF-LES, the results were smoothed through spatial averaging using 1-km wide horizontal boxes. The VENTOS®/M model surface boundary condition is based on Monin-Obukhov similarity theory (MOST) functions and a prescribed surface heat flux, as such quantity is expected to be less sensitive to height discrepancies between the mesoscale and microscale orographies. As heat-flux boundary conditions are not suitable to differentiate between intermittent and fully-turbulent regimes [21], the VENTOS®/M model always assumes the latter, which is not appropriate if stratification conditions are indeed very stable.

3.4. CFDWind

CFDWind1D is a python-based finite-difference code. The single-column model (SCM) is used as a prototype to design the CFDWind 3D model [9]. They are both based on unsteady Reynolds-Averaged Navier Stokes (URANS) equations using the *k-ε* model of Sogachev et al [23] with constants: $C_{ε1}$ = 1.52, $C_{ε2}$ = 1.833, $σ_k$ = 2.95, $σ_ε$ = 2.95 and $C_μ$ = 0.03.

The SCM is solved on a 4-km long log-linear vertical grid with 301 levels using a time step of 1 s. Pressure gradient and advection forcings are lumped together as a time and height-dependent equivalent geostrophic wind that enters momentum equations as source terms. The advection temperature tendency is also added as a source term in the potential temperature equation. No-slip conditions are defined for momentum equations at the surface. Surface boundary conditions for potential temperature are defined based on MOST, inferring the surface temperature by prescribing the diurnal 2-m temperature from the mesoscale input data and using the dynamic surface-layer friction velocity and heat flux as described in [9].

3.5. Alya-CFDWind

Alya-CFDWind1D is a fortran-based finite-elements code equivalent to CFDWindSCM. Similarly, the model is used at early-stage design and as a precursor of the in-house 3D model Alya-CFDWind for

numerical modelling of wind farms [24][25]. The same set of k-ε constants than CFDWind1D is used in this study.

The 1D mesh is 4-km tall using 500 vertical elements with a geometrically growing rate from 0.5 m up to 10 m. The time step is set to 10 s.

3.6. Ellipsys3D

The EllipSys3D code [26][27] is an in-house multiblock finite volume solver for the incompressible Navier-Stokes equations in general curvilinear coordinates. The code uses a collocated variable arrangement, where revised Rhie/Chow interpolation is used to avoid odd/even pressure coupling. The code is parallelized with Message Passing Interface (MPI) for executions on distributed/shared memory machines, using a non-overlapping domain decomposition technique.

The URANS case is based on k-ε model [22], with the same set of constants than the other URANS models of the benchmark. The problem is solved on a 6-km high vertical domain with 192 tanh stretched grid points using a time step of 1 s.

No-slip wall boundary condition at the bottom and a symmetry boundary condition at the top were applied. Periodic boundaries in both horizontal directions were used to mimic the 1D SCM basic setup. Pressure gradient and momentum advection forcings together with the temperature tendencies are included in the code in a way completely analogous to the procedure presented for the CFDWindSCM code [9].

In the LES case, the 4x4x4 km domain was discretized by 128x128x128 equidistantly spaced grid points, and Smagorinsky SGS model is used. Previously described URANS boundary conditions and time stepping are also applied in the LES case. EllipSys1D [28] is a recent "stripped down" version of EllipSys3D. The basic 3D URANS set-up was also applied in the 1D case. Only 512 tanh stretched grid points were used for discretizing the 6-km high vertical domain.

3.7. SP-Wind

SP-Wind is an in-house LES code developed at KU Leuven [29][30][31] for research on the simulation and optimization of turbulent flows. SP-Wind solves the Boussinesq form of the conservation equations for mass, momentum and potential temperature on a three-dimensional Cartesian grid. The horizontal directions are discretized with pseudo-spectral schemes, while the vertical direction uses a fourth-order energy-conservative finite difference scheme. Time integration is performed using a classic four-stage fourth-order Runge—Kutta scheme. The subgrid-scale stress and heat flux are computed with a prognostic TKE model [22], and the surface boundary conditions are imposed using classic Monin-Obukhov similarity theory [32] with standard stability correction functions [33].

The LES is solved on a horizontally periodic domain of 5x5x5 km with 400x400x500 grid points. The vertical grid has a uniform spacing of 5 m in the first 2 km, above which it is stretched to a maximum grid size of 40 m. A Rayleigh damping layer is added above 4 km. Initial profiles, mesoscale forcing and surface temperature (inferred from the 2-m temperature) are all extracted from mesoscale WRF simulations as in [9].

In view of the large computational cost, no spin-up time is used for the LES simulations. Analysis of the vertical profiles of turbulent kinetic energy shows that the positive surface heat flux at UTC 2006-07-01 12:00 leads to intense generation of turbulence, which rapidly fills up the entire boundary layer and reduces the transition time to less than one hour.

4. Results

4.1. Input uncertainties: Sensitivity Analysis at Mesoscale

We conduct a sensitivity analysis on the impacts of PBL scheme and global input data on WRF. Atmospheric profiles of potential temperature (θ), moisture (q) and winds (S, WD) over Cabauw at midnight LT on 2006-07-02 illustrate noticeable spread in the model output by the different configurations (Figure 2). Most of the configurations exhibit very similar temperature (Figure 2a) and wind (Figure 3c,d) profiles, while more spread is evident in the moisture profiles (Figure 2b). Further,

the TEMF-GFS (orange open squares) appears to deviate most from the other configurations, most noticeably in the lowest layers of the potential temperature, moisture, and wind speed profiles. The YSU-ERA (green filled dots), the reference mesoscale model used as input for the microscale simulations in this study, shows good agreement to the low-level moisture observations from the Cabauw tower (black dots), but overestimates the winds and the slope of the temperature inversion to a similar degree as the other configurations. Aside from the large difference in the TEMF-GFS and TEMF-ERA output, input data appear to have little impact on the model solution in this particular case, suggesting the input datasets agree on the larger-scale forcing, unlike other cases which show more spread between input datasets.
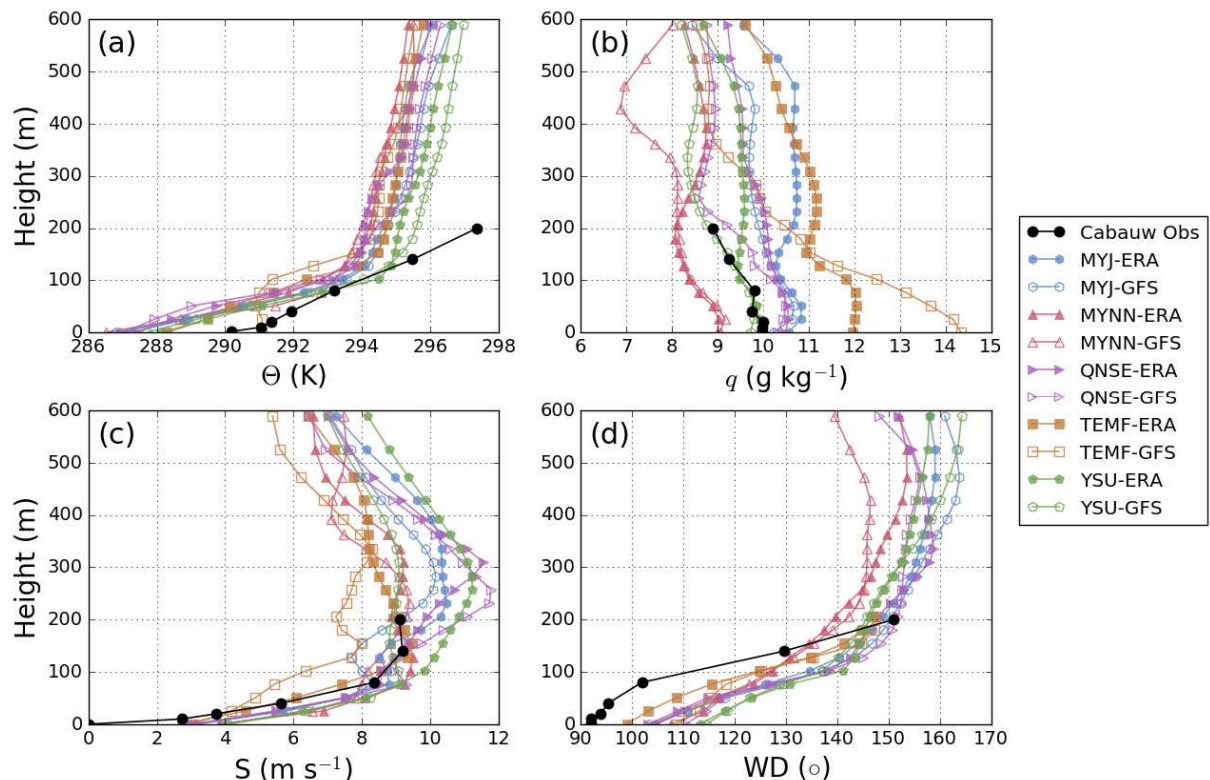
Figure 2: Atmospheric profiles of (a) potential temperature (θ), (b) water vapor mixing ratio (q), (c) wind speed (S), and (d) wind direction (WD) over Cabauw at midnight LT on 2006-07-02.

Time series of the same variables plotted over the 24-hour period between UTC 2006-07-01 1200 and 2006-07-02 1200 at 80 m show a similar spread between the PBL scheme and input data configurations (Figure 3). The TEMF-GFS again deviates most from the other simulations and the Cabauw observations, especially in the moisture (Figure 3b) and wind speed (Figure 3c) data. The other configurations form a generally narrow spread of output data. In particular, the time series of 80-m potential temperature (Figure 3a) exhibits a spread of ~2 K between the configurations, giving the whole ensemble a potential temperature bias consistently between -1 and -4 K. For the moisture and wind time series (Figure 3b-d), the ensemble of configurations frequently envelope the Cabauw observations during the 24-hour period, which may allude to minimal errors caused by choice of PBL scheme and input data on the microscale simulations. Similar conclusions may be drawn from time series of friction velocity (   ), PBL height (*PBL h*), latent heat flux (*LE*), and sensible heat flux (*H*) over the same 24-hour period at Cabauw (Figure 4). The greatest differences between the configurations is the TEMF-GFS friction velocity and the TEMF-ERA PBL height. The largest spread of output data is present in the PBL height, while the latent and sensible heat fluxes have the smallest spread. The ensemble of configurations generally overestimated latent and sensible heat fluxes during the day when compared to Cabauw observations, with fewer error during the night.
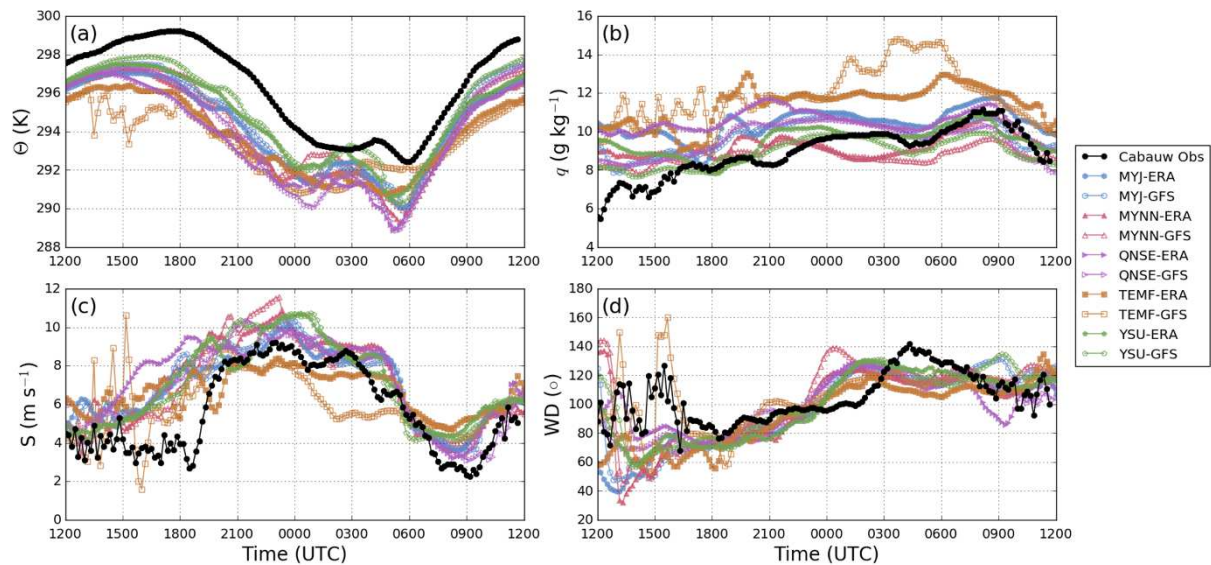
Figure 3: Time series of 80-m (a) potential temperature (Θ), (b) water vapor mixing ratio (q), (c) wind speed (S), and (d) wind direction (WD) over Cabauw during the 24-hour period between UTC 2006-07-01 1200 and 2006-07-02 1200.
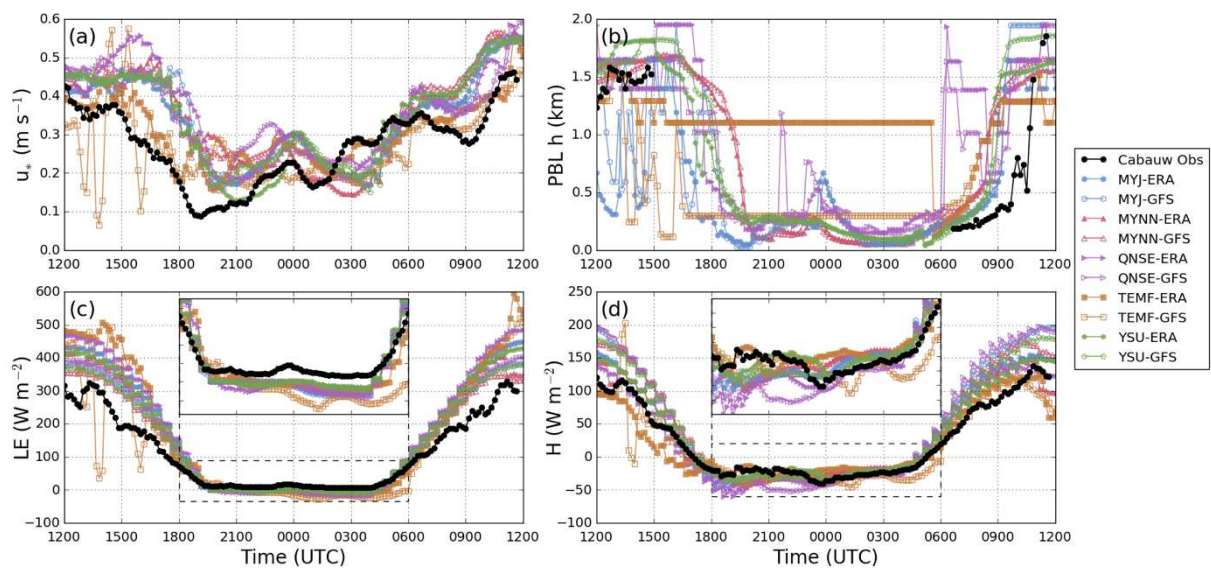

Figure 4: Time series of (a) friction velocity (u_*), (b) planetary boundary layer height ($PBL\ h$), (c) latent heat flux (LE), and (d) sensible heat flux (H) over Cabauw during the 24-hour period between UTC 2006-07-01 1200 and 2006-07-02 1200.

These results show that, excluding the TEMF, choice of input data has a near-negligible impact on the WRF output for this particular case, although input boundary condition datasets exhibit stronger influence in other cases [34]. Rather, differences in both the vertical profiles (Figure 2) and the time series (Figure 3 and Figure 4) between the tested configurations appear to be caused more so by choice of PBL scheme. Despite this dependence on PBL scheme, the ensemble of choices tends to maintain a consistent spread near the observations, with the only significant biases evident in the potential temperature and daytime latent and sensible heat fluxes. The greatest differences between PBL schemes occur in the PBL height calculations (Figure 4b). PBL heights were determined via direct output from each scheme, where each scheme uses a different approach that contributed to the scatter in the data. In the YSU and TEMF schemes, the PBL height is diagnosed from the bulk Richardson

number [12] [19]. For the local PBL schemes (MYJ, MYNN, and QNSE), PBL height is diagnosed as the height where prognostic TKE decreases to a sufficiently small value (0.2 m^2 s^{-2}, 1.0 x 10^{-6} m^2 s^{-2}, and 0.01 m^2 s^{-2}, respectively) [16] [17] [18]. Overall, the nontrivial spread in mesoscale forcings should be considered when making a choice in PBL scheme for a simulation that will be used as forcing for a microscale simulation.

4.2. Consistency among LES and RANS Microscale Models
All of the microscale models produce similar patterns of the diurnal cycle, demonstrating the effectiveness of the offline coupling methodology. Figure 5 shows contour plots of the evolution of the vertical profile of mean flow quantities computed by the participating models compared to the observations on the top row. The second row corresponds to the reference mesoscale model that is used to derive input forcings and boundary conditions to drive the microscale simulations. Since no higher-resolution terrain or land cover information is added at microscale, we can assume that this mesoscale simulation is already a good reference for microscale models to verify a correct implementation of the meso-micro methodology. The differences in the mean flow arise from the different ways each model represents turbulence, which is noticed in the *TKE* contour plots.

The ensemble mean of the WRF simulations used in the sensitivity analysis is plotted in the third row. By ensemble averaging, we obtain a better match with the observations than by using any single WRF simulation – this result was also seen for the LLJ cases discussed in [34]. The ensemble here provides a better prediction of the LLJ with a distinct velocity maximum around midnight instead of a broader double-peak as in the reference WRF simulation.

The microscale models diverge in their estimates of rotor-based quantities of interest. Time series of rotor-based quantities of interest are shown in Figure 6. The spread of the models for *REWS* is around 2 m s^{-1} and around 15º in terms of *WD$_{hub}$*. The spread in terms of wind shear and wind veer is also large specially during nighttime stable conditions. Vertical profiles of wind speed and direction at midnight (Figure 7) show how the models capture the characteristics of the LLJ. The phase error in the input data dominates the bias in the simulations; this input error cannot be corrected at microscale by simply changing the turbulence model.

Table 2 summarizes the differences between the simulations and observations in terms of (4) with respect to observations (*MAE$_{obs}$*) and with respect to the reference WRF simulation (*MAE$_{ref}$*). These are differences integrated over the whole diurnal cycle; therefore mixing all kinds of surface stability and large-scale conditions into a single quantity. We shall focus on the *MAE$_{ref}$* to quantify the impact of choosing a different turbulence model at microscale.

Table 2: MAE with respect to observations and to the reference mesoscale simulation.

|	*REWS* [m s^{-1}]		*S$_{hub}$* [m s^{-1}]		*WD$_{hub}$* [º]		*α* (shear)		*ψ* (veer)	
	MAE$_{obs}$	*MAE$_{ref}$*	*MAE$_{obs}$*	*MAE$_{ref}$*	*MAE$_{obs}$*	*MAE$_{ref}$*	*MAE$_{obs}$*	*MAE$_{ref}$*	*MAE$_{obs}$*	*MAE$_{ref}$*
WRF-YSU (ref)	1.26	0.00	1.35	0.00	10.49	0.00	0.13	0.00	0.07	0.00
WRF ensemble	1.16	0.63	1.12	0.76	14.24	5.14	0.11	0.05	0.07	0.02
WRF-YSU_LES	1.51	0.45	1.60	0.54	10.67	4.09	0.15	0.05	0.06	0.03
WRF-VentosM_ke	1.56	0.69	1.59	0.72	10.74	6.34	0.12	0.05	0.06	0.03
CFDWind1D_ke	1.56	0.63	1.62	0.68	11.49	6.22	0.15	0.09	0.06	0.05
Alya-CFDWind1D_ke	1.48	0.73	1.42	0.70	11.30	2.90	0.14	0.07	0.07	0.02
Ellipsys1D_ke	1.37	0.55	1.50	0.61	11.51	5.22	0.16	0.10	0.06	0.04
Ellipsys3D_ke	1.36	0.66	1.52	0.74	10.61	5.17	0.16	0.09	0.06	0.04
Ellipsys3D_LES	1.38	1.13	1.37	1.27	11.90	11.99	0.18	0.09	0.09	0.06
SP-Wind_LES	1.47	0.63	1.38	0.69	8.79	3.25	0.13	0.04	0.07	0.02
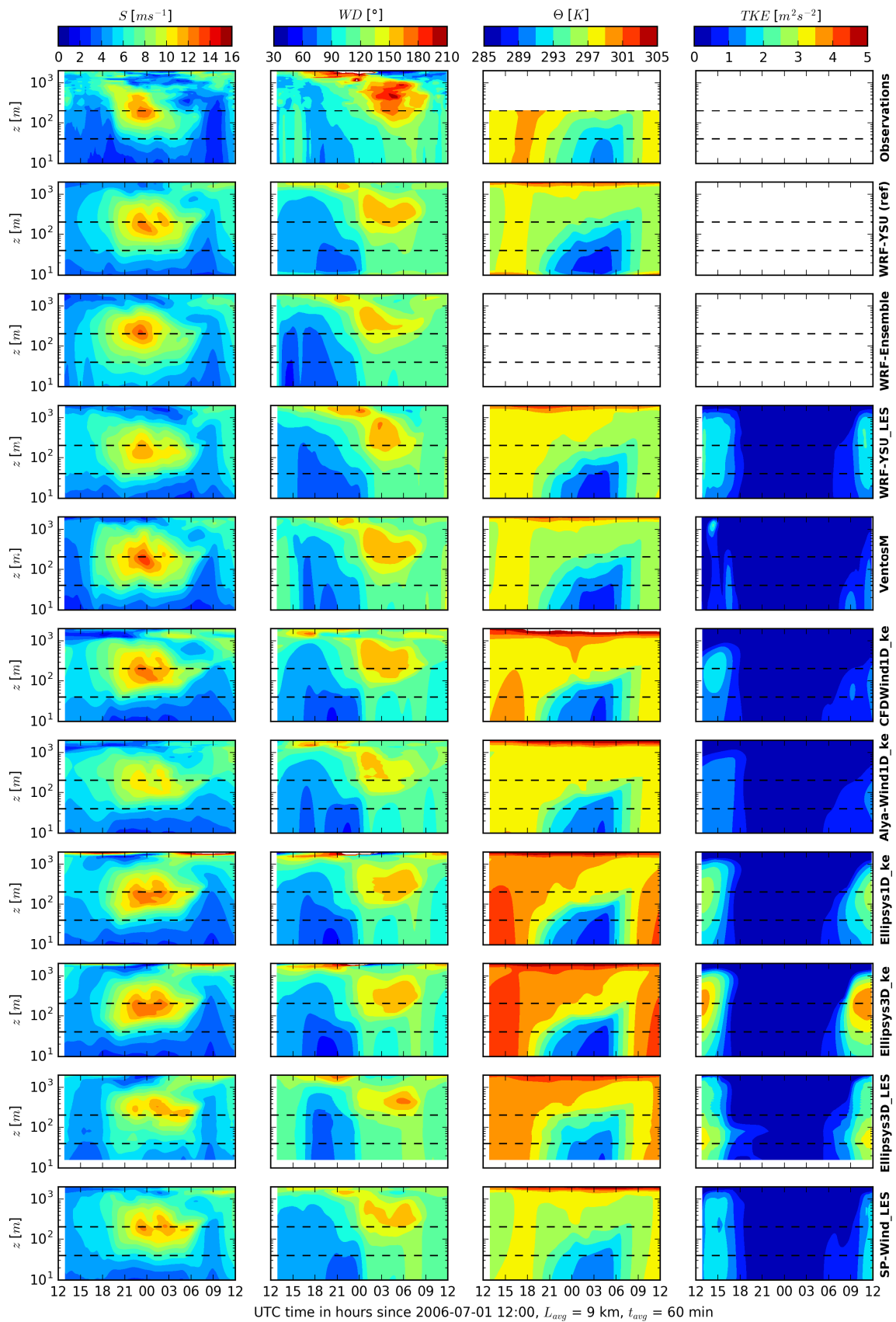
Figure 5: Time-height contours of horizontal wind speed *S*, direction *WD*, potential temperature Θ and turbulent kinetic energy *TKE*. Dotted lines denote a rotor diameter of 160 m at 120 m hub-height.
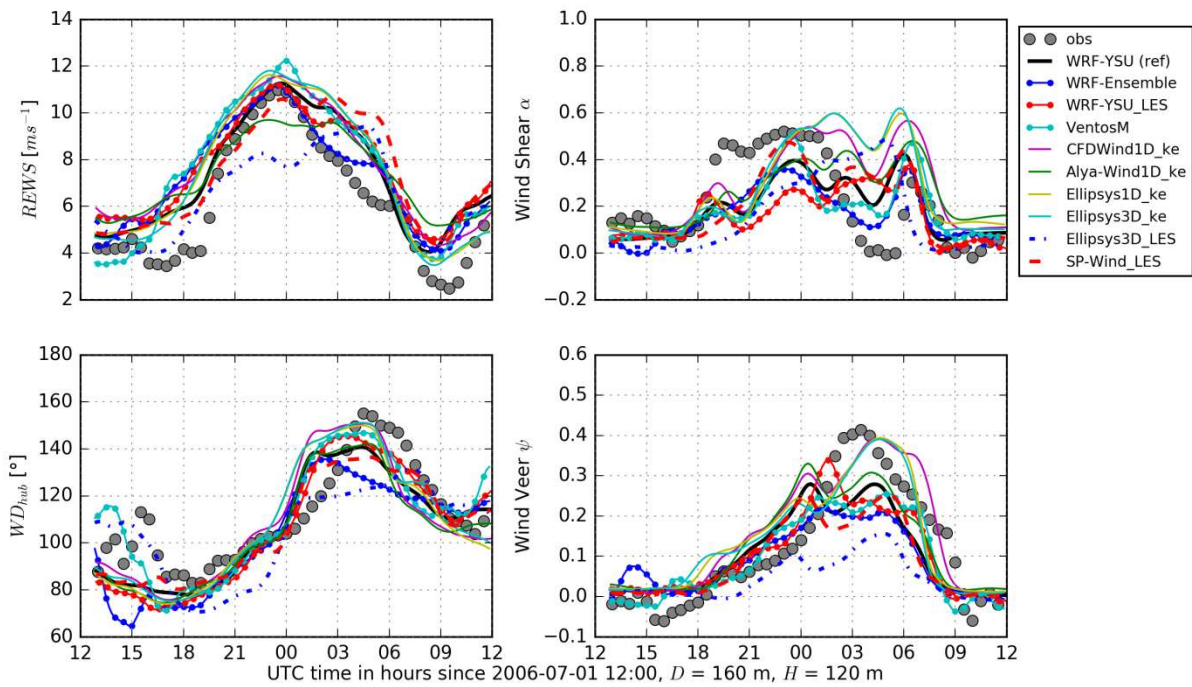
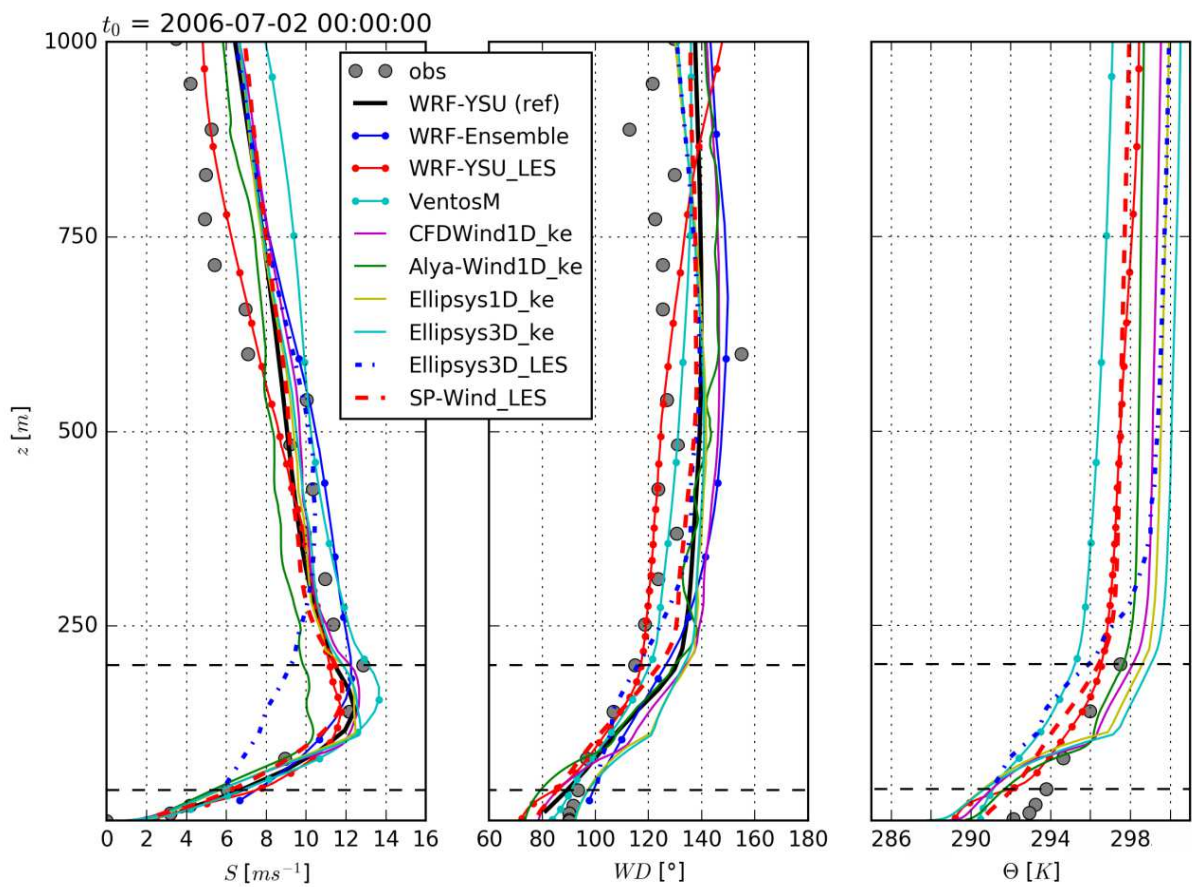**Figure 6:** Time series of rotor-based quantities of interest used for validation.

Figure 7: Vertical profiles of wind speed, wind direction and potential temperature at UTC 2006-07-02 00:00:00.

WRF-YSU_LES is the closest to the reference, which is to be expected since they are results from different nests of the same simulation. Still, differences are significant of around 0.5 m s^{-1} of wind speed and 4° of wind direction at hub-height. Microscale models increase the error by 0.2-0.7 m s^{-1} of wind speed and up to 7° of wind direction at hub-height with respect to the WRF-YSU_LES results. With respect to observations, all simulations show similar results with a *MAE* of 1.1-1.6 m s^{-1} of wind speed and up to 14° of wind direction at hub-height. Considering vertical wind speed shear and wind direction veer, SP-Wind and the WRF ensemble produce the closest results to the reference WRF simulation.

Different sets of *k-ε* constants have been tested (not shown) leading to changes that are within the spread shown in Figure 6, which is also comparable to that observed in the WRF sensitivity analysis when changing the planetary boundary-layer scheme.

Recently, capabilities of EllipSys3D have been extended to cover wall modeled LES of stratified flows and at the same time, a "striped" down 1D version of the code have been made operational [28]. Figures 5 and 6 and Tables 2 and 3 show a general good agreement between URANS based EllipSys1D and EllipSys3D computations. Some minor differences exist thought; a possible cause of them might be related to the fact that the vertical velocity (*W*) and the advection terms are implicitly assumed to be zero in the EllipSys1D. Further investigations are necessary to confirm this.

Initial LES computations based on coarse grid resolution and very basic Smagorinsky SGS model show that EllipSys3D LES is capable of reproducing the basic LLJ features observed at the Cabauw site, but Tables 2 and 3 MAE comparisons and Figures 2 and 3 indicate that significantly higher resolution and a more advanced approach to SGS turbulence modeling are needed in order to capture all main details relevant for its application in a wind energy context.

Regarding VENTOS®/M wind speed results, both Figures 5 and 7 show that the LLJ magnitude was reasonably well predicted. The diurnal cycle of the simulation wind speed shows over-predictions around 1.4 m s^{-1}, affecting the *REWS* and S_{hub} error values in Table 2 which are higher than the reference simulation. Good agreement was obtained for the wind direction, shear and veer regarding their integrated error. Analysis of Figure 6 indicates a generalized under-prediction of *α* for several nocturnal periods, analogous to a higher turbulent shape factor of the boundary-layer. These mismatches happen also with WRF and, despite the VENTOS®/M limitations of its heat-flux boundary condition, the microscale simulation predicts higher values of *α* and closer to the observations. The results further show a -2.5 K temperature bias that occurs in both WRF and VENTOS®/M results, as well as in the other microscale simulations, which originates from the ERA-Interim input data.

Table 3: MAE with respect to the reference mesoscale simulation for unstable ('*u*': z/L < -0.2), neutral ('*n*': -0.2 < z/L < 0.2) and stable ('*s*': z/L > 0.2) conditions.

	REWS [m s^{-1}]			S_{hub} [m s^{-1}]			WD_{hub} [°]			*α* (shear)			*ψ* (veer)		
	U	*n*	*s*	*u*	*n*	*s*	*u*	*n*	*s*	*u*	*n*	*s*	*u*	*n*	*s*
WRF-YSU (ref)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WRF ensemble	0.45	0.51	0.82	0.51	0.60	1.02	5.79	8.98	3.97	0.03	0.09	0.06	0.01	0.02	0.03
WRF-YSU_LES	0.48	0.44	0.42	0.58	0.72	0.48	4.59	4.15	3.62	0.03	0.05	0.08	0.02	0.02	0.03
WRF-VentosM	0.64	0.75	0.74	0.68	0.80	0.73	9.74	7.02	3.10	0.04	0.08	0.05	0.04	0.02	0.03
CFDWind1D_ke	0.74	0.45	0.56	0.73	0.53	0.65	5.62	7.87	6.53	0.07	0.10	0.11	0.05	0.09	0.05
Alya-CFDWind1D_ke	0.79	0.39	0.72	0.78	0.51	0.66	4.66	3.16	1.24	0.09	0.03	0.06	0.02	0.02	0.03
Ellipsys1D_ke	0.58	0.40	0.55	0.57	0.40	0.69	4.30	7.47	5.75	0.06	0.12	0.13	0.03	0.10	0.05
Ellipsys3D_ke	0.67	0.47	0.68	0.67	0.60	0.83	4.13	8.03	5.70	0.04	0.14	0.13	0.02	0.10	0.04
Ellipsys3D_LES	0.56	1.27	1.64	0.59	1.06	1.93	11.69	8.36	12.81	0.07	0.10	0.11	0.02	0.05	0.11
SP-Wind_LES	0.48	1.36	0.66	0.50	1.38	0.76	3.00	1.72	3.71	0.02	0.02	0.06	0.01	0.02	0.03

Finally, Table 3 shows the MAE_{ref} is computed for different stability classes filtering with the observed stability parameter z/L, where *L* is the Obukhov length and *z* = 10 m: unstable

(u': $z/L < -0.2$), neutral (n': $-0.2 < z/L < 0.2$) and stable (s': $z/L > 0.2$). Not all the models behave similarly depending on stability. The WRF ensemble *REWS* is more sensitive in stable conditions. This is also the case for Ellipsys3D_LES probably due to the coarser resolution of the simulation compared to the other LES simulations that do not show this high sensitivity in stable conditions.

In general, it is difficult to extract more meaningful conclusions from Table 2 and Table 3 due to the limited statistical significance of the samples. The overall assessment would be richer if several diurnal cycles from uncorrelated synoptic conditions would have been tested. The ensemble WRF simulations for the same cycle already show significant improvement on mean flow quantities.

5. Conclusions

Results of the GABLS3 diurnal cycle benchmark with an emphasis on rotor-relevant values are presented. The main challenge for microscale models was to produce consistent flow fields with respect to the mesoscale model that was used to derive their input forcings. This consistency has been achieved by both LES and URANS models. The spread of the models is significant but of similar magnitude as that shown by WRF using different boundary-layer parameterizations. The input uncertainty coming from the mesoscale, even in relatively ideal conditions, is large and results in *MAE* of wind speed at hub-height of the order of 1.1-1.6 m s^{-1} over the whole cycle and hourly errors of up to 3 m s^{-1}. This is partly mitigated when using an ensemble average of several simulations which also lead to better results in terms of wind shear and wind veer.

By ensuring consistency of the microscale models at introducing input forcings we can proceed with further analysis on how RANS and LES models interpret the structure of turbulence in different stability conditions.

Acknowledgments

References
[1]   Sanz Rodrigo J, Chávez Arroyo R-A, Moriarty P, Churchfield M, Kosović B, Réthoré P-E, Hansen KS, Hahmann A, Mirocha JD and Rife D 2016 Mesoscale to microscale wind farm flow modelling and evaluation. *WIREs Energy Environ*, doi:10.1002/wene.214

[2]   Wyngaard JC 2004 Toward Numerical Modeling in the "Terra Incognita". *J. Atmos. Sci*. **61**: 1816–1826, doi: 10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2

[3]   Holtslag AAM, et al. (2013) Stable atmospheric boundary layers and diurnal cycles—challenges for weather and climate models. *Bull. Am. Meteorol. Soc*. **94**: 1691–1706, doi:10.1175/bams-d-11-00187.1

[4]   Baas, P., Bosveld, F.C., Baltink, K. and Holtslag, A.A.M. 2009 A Climatology of Nocturnal Low-Level Jets at Cabauw, *J. Appl. Meteorol. Climatol.* **48**: 1627-1642, doi: 10.1175/2009JAMC1965.1, 2009

[5]   Bosveld FC, Baas P, vanMeijgaard E, de Bruijn EIF, Steeneveld GJ and 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*. **152**: 133-156, doi:

10.1007/s10546-014-9917-3

[6] Bosveld FC et al. 2014 The third GABLS Intercomparison case for evaluation studies of Boundary-Layer Models Part B: Results and Process Understanding, *Boundary–Layer Meteorol*. **152**: 157-187, doi: 10.1007/s10546-014-9919-1

[7] Sanz Rodrigo J 2017 Windbench/GABLS3 benchmark. http://windbench.net/gabls-3, last accessed in March 2017

[8] Sanz Rodrigo, Churchfield M and Kosovic B 2016 A wind energy benchmark for ABL modeling of a diurnal cycle with a nocturnal low-level jet: GABLS3 revisited. *J. Phys. Conf. Ser*. **753**: 032024, doi:10.1088/1742-6596/753/3/032024

[9] Sanz Rodrigo J, Churchfield M and Kosović B 2017 A methodology for the design and testing of atmospheric boundary layer models for wind energy applications. *Wind Energ. Sci*. **2**: 1-20, doi:10.5194/wes-2-1-2017

[10] Skamarock WC, Klemp JB, Dudhia J, Gill DO, Barker DM, Duda MG, Huang X-Y, Wang W and Powers JG 2008 A description of the advanced research WRF version 3, Technical Note NCAR/TN-475+STR, NCAR, Boulder, CO, June 2008.

[11] Kleczek MA, Steeveneveld GL and Holtslag AAM 2014 Evaluation of the Weather Research and Forecasting Mesoscale Model for GABLS3: Impact on Boundary-Layer Schemes, Boundary Conditions and Spin-Up. *Boundary-Layer Meteorol*. **152**: 213-243, doi: 10.1007/s10546-014-9925-3

[12] Hong S-Y and Noh Y 2006 A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Mon. Wea. Rev.*, **134**: 2318–2341, doi: 10.1175/MWR3199.1

[13] Dee DP et al. 2011 The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. R. Meteorol. Soc*. **137**: 553-597, doi: 10.1002/qj.828

[14] Wagner R, Cañadillas B, Clifton A, Feeney S, Nygaard N, Martin CSt, Tüxen E and Wagenaar JW 2014 Rotor equivalent wind speed for power curve measurement – comparative exercise for IEA Wind Annex 32. *J. Phys. Conf. Ser*. **524**: 012108, doi: 10.1088/1742-6596/524/1/012108

[15] Choukulkar A, Pichugina Y, Clack CTM, Calhoun R, Banta R, Brewer A and Hardesty M 2015 A new formulation for rotor equivalent wind speed for wind resource assessment and wind power forecasting. *Wind Energy* **19**: 1439–1452, doi:10.1002/we.1929

[16] Janjić ZI 1994 The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Mon. Wea. Rev.* **122**: 927–945, doi: 10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

[17] Nakanishi M and H Niino 2006 An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorol* **119**: 397-407, doi:10.1007/s10546-005-9030-8.

[18] Sukoriansky S, Galperin B and Pero V 2005 Application of a New Spectral Theory of Stably Stratified Turbulence to the Atmospheric Boundary Layer over Sea Ice. *Boundary-Layer Meteorol*. **117**: 231-257. doi:10.1007/s10546-004-6848-4.

[19] Angevine WM 2005 An Integrated Turbulence Scheme for Boundary Layers with Shallow Cumulus Applied to Pollutant Transport. *J. Appl. Meteor*. **44**: 1436–1452, doi: 10.1175/JAM2284.1.

[20] Veiga Rodrigues C, Palma JMLM and Rodrigues ÁH 2016 Atmospheric Flow over a Mountainous Region by a One-Way Coupled Approach Based on Reynolds-Averaged Turbulence Modelling. *Boundary-Layer Meteorol*. **159**: 407–437, doi: 10.1007/s10546-015-0116-7

[21] van de Wiel BJH, Moene AF, Steeneveld GJ, Hartogensis OK and Holtslag AAM. 2007 Predicting the Collapse of Turbulence in Stably Stratified Boundary Layers. *Flow Turbulence Combust.*, **79**:251–274, doi: 10.1007/s10494-007-9094-2

[22] Sogachev A, Kelly M and Leclerc MY 2012 Consistent Two-Equation Closure Modelling for Atmospheric Research: Buoyancy and Vegetation Implementations. *Boundary-Layer Meteorol.* **145**: 307–327, doi:10.1007/s10546-012-9726-5

[23] Deardorff J W 1980 Stratocumulus-capped mixed layers derived from a three dimensional

model. *Boundary-Layer Meteorol*. **18**: 495-527, doi:10.1007/BF00119502

[24] Avila M, Folch A, Houzeaux G, Eguzkitza B, Prieto L and Cabezon D 2013 A Parallel CFD Model for Wind Farms. *Procedia Comput. Sci*. **18**: 2157 – 2166, doi:10.1016/j.procs.2013.05.386

[25] Vázquez M et al. 2015 Alya: Multiphysics engineering simulation toward exascale. *J. Comput. Scie*. **14**: 15–27, doi:10.1016/j.jocs.2015.12.007

[26] Michelsen JA 1994 Block structured Multigrid solution of 2D and 3D elliptic PDE's. Technical Report AFM 94-06, Technical University of Denmark, Department of Fluid Mechanics, May 1994.

[27] Sørensen NN 1995 General Purpose Flow Solver Applied to Flow over Hills. Risø-R-827-(EN), Risø National Laboratory, Roskilde, Denmark, June 1995

[28] van der Laan MP and Sørensen NN 2017 A 1D version of EllipSys. Technical Report, DTU Wind Energy E-0141, March 2017

[29] Meyers J, Meneveau C 2010 Large eddy simulations of large wind-turbine arrays in the atmospheric boundary layer. AIAA Paper No. 2010-827, doi:10.2514/6.2010-827

[30] Munters W, Meneveau C and Meyers J 2016 Turbulent inflow precursor method with time-varying direction for large-eddy simulations and applications to wind farms. *Boundary-Layer Meteorol*. **159**: 305-328, doi:10.1007/s10546-016-0127-z

[31] Allaerts D and Meyers J 2017 Boundary-layer development and gravity waves in conventionally neutral wind farms. *J. Fluid Mech*. **814**: 95-130, doi:10.1017/jfm.2017.11

[32] Moeng C-H 1984 A large-eddy-simulation model for the study of planetary boundary-layer turbulence. *J. Atmos. Sci*. **41**: 2052-2062, doi:10.1175/1520-0469(1984)041<2052:ALESMF>2.0.CO;2

[33] Dyer A J 1974 A review of flux-profile relationships. *Boundary-Layer Meteorol*. **7**: 363-372, doi:10.1007/BF00240838

[34] Vanderwende BJ, Lundquist JK, Rhodes ME, Takle ES and Irvin SL 2015 Observing and Simulating the Summertime Low-Level Jet in Central Iowa. *Mon. Weater Rev*. 143(6): 2319–2336, doi:10.1175/mwr-d-14-00325.1