The Benefits and Challenges of Downscaling a Global Reanalysis with Doubly-Periodic Large-Eddy Simulations

Bart van Stratum[1], Chiel C. van Heerwaarden[1], and Jordi Vilà-Guerau de Arellano[2]

[1]Wageningen University & Research
[2]Wageningen University Research

April 16, 2023

Abstract

Global reanalyses like ERA5 accurately capture atmospheric processes at spatial scales of O(10) km or larger. By downscaling ERA5 with large-eddy simulation (LES), LES can provide details about processes at spatio-temporal scales down to meters and seconds. Here, we present an open-source Python package named the "Large-eddy simulation and Single-column model - Large-Scale Dynamics", or (LS)2D in short, designed to simplify the downscaling of ERA5 with doubly-periodic LES. A validation with observations, for several sensitivity experiments consisting of month-long LESs over Cabauw (the Netherlands), demonstrates both its usefulness and limitations. The day-to-day variability in the weather is well captured by (LS)2D and LES, but the setup under-performs in conditions with broken or near overcast clouds. As a novel application of this modeling system, we used (LS)2D to study surface solar irradiance variability, as this quantity directly links land-surface processes, turbulent transport, and clouds, to radiation. At a horizontal resolution of 25 m, the setup reproduces satisfactorily the solar irradiance variability down to a timescale of seconds. This demonstrates that the coupled LES-ERA5 setup is a useful tool that can provide details on the physics of turbulence and clouds, but can only improve on its host reanalysis when applied to meteorological suitable conditions.

The Benefits and Challenges of Downscaling a Global Reanalysis with Doubly-Periodic Large-Eddy Simulations

B.J.H. van Stratum[1], C.C. van Heerwaarden[1], and J. Vilà-Guerau de Arellano[1]

[1]Meteorology and Air Quality Group, Wageningen University & Research
[1]P.O. Box 47, 6700 AA Wageningen, The Netherlands

Key Points:

- We developed an open-source Python package named (LS)^{2}D, designed to downscale the ERA5 reanalysis with turbulence and cloud-resolving LESs
- One month long experiments with (LS)^{2}D and MicroHH over the Netherlands demonstrate both the skill and limitations of the coupled setup
- Capturing high-frequency interactions between clouds and surface solar irradiance requires high resolution ($\mathcal{O}(10)$ m) LES

Corresponding author: Bart van Stratum, `bart.vanstratum@wur.nl`

Abstract

Global reanalyses like ERA5 accurately capture atmospheric processes at spatial scales of $\mathcal{O}(10)$ km or larger. By downscaling ERA5 with large-eddy simulation (LES), LES can provide details about processes at spatio-temporal scales down to meters and seconds. Here, we present an open-source Python package named the *"Large-eddy simulation and Single-column model - Large-Scale Dynamics"*, or $(\mathrm{LS})^2\mathrm{D}$ in short, designed to simplify the downscaling of ERA5 with doubly-periodic LES. A validation with observations, for several sensitivity experiments consisting of month-long LESs over Cabauw (the Netherlands), demonstrates both its usefulness and limitations. The day-to-day variability in the weather is well captured by $(\mathrm{LS})^2\mathrm{D}$ and LES, but the setup under-performs in conditions with broken or near overcast clouds. As a novel application of this modeling system, we used $(\mathrm{LS})^2\mathrm{D}$ to study surface solar irradiance variability, as this quantity directly links land-surface processes, turbulent transport, and clouds, to radiation. At a horizontal resolution of 25 m, the setup reproduces satisfactorily the solar irradiance variability down to a timescale of seconds. This demonstrates that the coupled LES-ERA5 setup is a useful tool that can provide details on the physics of turbulence and clouds, but can only improve on its host reanalysis when applied to meteorological suitable conditions.

Plain Language Summary

Modern global weather models are accurate in predicting atmospheric processes at scales of around 10 km or larger, but are less good at predicting smaller scale processes, like for example the interaction between solar radiation, individual clouds, and the resulting clouds shadows that are cast onto the land surface. High spatio-temporal resolution research models are able to capture these smaller scale processes, but require a coupling to a weather model to account for the day-to-day variability in our weather. In this paper, we present a framework to couple large to small scale models, and demonstrate both the benefits and challenges of using this coupled model setup. The coupled setup excels in capturing the aforementioned high frequency interactions between small clouds and surface solar radiation. However, the chaotic nature of broken to overcast clouds is proven difficult to represent. The coupled model setup is published as open-source code, and is therefore freely available to the research community.

1 Introduction

Atmospheric processes cover a large range of spatial and temporal scales, varying from less than meters and seconds, to more than thousands of kilometers and years. With improvements in models like the Integrated Forecasting System (IFS), global reanalysis such as ERA5 (Hersbach et al., 2020) have become an accurate data source to study atmospheric processes at the meso-α (Fujita, 1981) and larger scales. On smaller scales, large-eddy simulation (LES) can resolve processes on scales ranging from the order of meters up to the meso-α scale. As such, LES can potentially be a useful tool to downscale a global reanalysis, and provide details about small scale processes which are unresolved by its host model (e.g. Neggers et al., 2012; Gustafson et al., 2020). This includes – in the context of this paper – the complex interaction between turbulence, moist convection, radiative transfer, and the land-surface (e.g. Huang & Margulis, 2011; Rieck et al., 2014; Pedruzo-Bagazgoitia et al., 2017; Veerman et al., 2020; Vilà-Guerau de Arellano et al., 2023).

Solving turbulent scales simultaneously with the large-scale weather requires a coupling of LES with a large-scale weather model. Two different coupling methods are typically employed, which we will refer to as *open boundary* and *periodic boundary* LES. In an open boundary LES setup, the LES model is coupled to a large scale model via relaxation of the prognostic fields at the lateral boundaries (e.g. Talbot et al., 2012; Heinze, Dipankar, et al., 2017). The advantage of such a setup is that large- or mesoscale spatial variability – like e.g. frontal systems – is advected into the LES domain (e.g. Schemann et al., 2020). However, such as setup requires a large domain and/or grid nesting to allow turbulence to fully develop at the inflow boundaries, making the simulations expensive (Heinze, Dipankar, et al., 2017). The periodic boundary LES setup – which we will employ – circumvents this problem by using doubly-periodic lateral boundary conditions, only coupling LES to the large-scale weather by applying horizontally homogeneous but time and height varying large-scale forcings. These forcings typically contain the advective tendencies of heat, moisture, and momentum, the large-scale vertical (subsidence) velocity, and the geostrophic wind components (e.g. Neggers et al., 2012; Schalkwijk et al., 2015; Heinze, Moseley, et al., 2017; van Laar et al., 2019). Although such a setup clearly has shortcomings in complex large- or mesoscale conditions, it allows for smaller domains. This makes the simulations computationally cheaper, or allows the use of a finer computational mesh and hence resolve turbulence on smaller scales. As we will demonstrate, such high resolution LESs are needed to capture high frequency interactions between solar radiation and clouds, which is a key process required to advance our understanding of land-atmosphere interactions (e.g. Gentine et al., 2019; Vilà-Guerau de Arellano et al., 2023).

Previous work on realistic periodic boundary LESs was often done using limited area host models (e.g. Neggers et al., 2012; Schalkwijk et al., 2015; Heinze, Moseley, et al., 2017). One advantage of these models over a global model like ERA5 is the increased horizontal resolution. Nevertheless, we have chosen to use ERA5 for a number of reasons. First, ERA5 has a global coverage, and therefore does not limit the LES simulations to a certain geographical region. Second, ERA5 covers a long time period from 1950 to within 5 days from real time. And third, and perhaps most important, the ERA5 data can easily be accessed through the Copernicus Data Store (CDS[1]), bypassing the need to request data from various national weather services in case it is not openly available. Another commonality in most previous work is that the underlying code used to preprocess the host model data was closed-source. In order to improve scientific transparency and reproducibility, we released our code as an open-source Python package named the "*Large-eddy simulation and Single column model - Large-Scale Dynamics*", or (LS)^{2}D in short. This Python package contains a number of routines meant to simplify and automate all the steps required to generate the input for doubly-periodic LESs.

The goal of this paper is threefold. First, we document (LS)^{2}D by describing the coupling method between ERA5 and LES, and the procedures in the (LS)^{2}D Python package (section 2). Although (LS)^{2}D has already proven its usefulness and skill in a number of studies (Ražnjević et al., 2022; Veerman et al., 2022; Tijhuis et al., 2023; Mol, van Stratum, et al., 2023), this previous work mostly focused on clear or shallow convective conditions. Therefore, the second goal of this paper is to quantify the added value of doubly-periodic LES with respect to ERA5 for (i) dry and shallow convective conditions, and (ii) more challenging conditions. We do this by simulating August 2016 over Cabauw (the Netherlands) with (LS)^{2}D and MicroHH (van Heerwaarden et al., 2017), and validating the results with the Cabauw surface, tower, and remote sensing observations (section 4.1-4.2). The third goal is to discuss the pros and cons of using a doubly-periodic LES setup by comparing a key process – the surface solar irradiance variability – with detailed observations from the Baseline Surface Radiation Network (BSRN, Knap, 2022; Mol, Knap, & van Heerwaarden, 2023). With a sensitivity experiment consisting of six different one-month LESs, we address the impact of domain size and resolution on the ability of LES to capture different surface solar irradiance time scales (section 4.3). Finally, we discuss our findings, and conclude the paper in section 6.

[1] https://cds.climate.copernicus.eu

2 LES - ERA5 coupling with $(LS)^2D$

2.1 Methods

To account for the influence of the large scale weather acting on the local LES domain, $(LS)^2D$ creates a one-way coupling between ERA5 and LES. The coupling method that we apply is similar to the methods described by Neggers et al. (2012); Schalkwijk et al. (2015); Heinze, Moseley, et al. (2017). In this approach, the initial conditions (atmosphere and soil) and a number of time and height varying large-scale processes acting on the LES domain, are derived from routine output of (in our case) ERA5. These processes, partly in the form of a tendency of a state variable like temperature, humidity, or wind, are then added to the prognostic LES equations, as shown in Eq. 1 and Eq. 2:

$$\left.\frac{\partial \widetilde{\psi}}{\partial t}\right|^{\mathrm{LS}} = -\underbrace{\left\langle u_j^{\mathrm{LS}} \frac{\partial \psi^{\mathrm{LS}}}{\partial x_j} \right\rangle}_{\text{advection}} - \underbrace{\left\langle w^{\mathrm{LS}} \right\rangle \frac{\partial \widetilde{\psi}}{\partial z}}_{\text{subsidence}} + \underbrace{\frac{1}{\tau_{\mathrm{n}}} \left(\left\langle \psi^{\mathrm{LS}} \right\rangle - \left\langle \widetilde{\psi} \right\rangle \right)}_{\text{relaxation}} + \underbrace{\left\langle F_{\psi}^{\mathrm{LS}} \right\rangle}_{\text{source}}, \tag{1}$$

$$\left.\frac{\partial \widetilde{u}_i}{\partial t}\right|^{\mathrm{LS}} = -\underbrace{\left\langle u_j^{\mathrm{LS}} \frac{\partial u_i^{\mathrm{LS}}}{\partial x_j} \right\rangle}_{\text{advection}} - \underbrace{\left\langle w^{\mathrm{LS}} \right\rangle \frac{\partial \widetilde{u}_i}{\partial z}}_{\text{subsidence}} + \underbrace{\epsilon_{ij3} f_{\mathrm{c}} \left(\widetilde{u}_j - \left\langle u_{\mathrm{g};j}^{\mathrm{LS}} \right\rangle \right)}_{\text{coriolis}} + \underbrace{\frac{1}{\tau_{\mathrm{n}}} \left(\left\langle u_i^{\mathrm{LS}} \right\rangle - \left\langle \widetilde{u}_i \right\rangle \right)}_{\text{relaxation}}. \tag{2}$$

Here, ψ is a generic scalar, in MicroHH either the liquid water potential temperature (θ_{l}), the total specific humidity (q_{t}), or other scalars, and u_i is a vector with the horizontal wind components $(u, v, 0)$. Variables with a superscript LS are variables from ERA, and variables with a tilde denote the filtered LES fields. The angle brackets indicate a horizontal averaging operation, either over a number of $\sim 30 \times 30$ km^2 grid points in ERA5, or the entire LES domain.

The coupling between the resolved ERA5 variables and the turbulent LES variables consists of a number of terms:

- The advective tendency contains the resolved advective tendency from the host model. As this tendency is not available in the ERA5 output, it is approximated offline from the three-dimensional fields.
- The subsidence term contains the interaction between the large-scale vertical velocity from ERA5, and the turbulent LES fields.
- The source term can contain any external forcing, for example radiative heating rates for LES simulations without interactive radiation.
- The Coriolis term contains the influence of the large-scale pressure gradient on the horizontal wind components in LES, through the geostrophic wind.
- Finally, the relaxation term is a safety measure for long experiments, which prevents the LES model from deviating too far from the host model, by nudging the horizontal mean state of LES to the mean state of ERA5, on a timescale τ_{n}.

In these equations, f_{c} is the Coriolis frequency $f_{\mathrm{c}} = 2\,\Omega \sin(\phi)$, where $\Omega = 7.2921 \cdot 10^{-5}$ rad s^{-1} and ϕ is the latitude. The two horizontal geostrophic wind components are denoted by $u_{\mathrm{g;j}}^{\mathrm{LS}}$, and are defined as:

$$u_{\mathrm{g}}^{LS} = -\frac{g}{f_{\mathrm{c}}} \frac{\partial Z^{\mathrm{LS}}}{\partial y}, \tag{3}$$

$$v_{\mathrm{g}}^{LS} = \frac{g}{f_{\mathrm{c}}} \frac{\partial Z^{\mathrm{LS}}}{\partial x}, \tag{4}$$

```python
import ls2d

settings = {
    'central_lat' : 51.971,
    'central_lon' : 4.927,
    ... other settings ...
    'start_date'  : datetime(year=2016, month=8, day=15, hour=6),
    'end_date'    : datetime(year=2016, month=8, day=15, hour=18),
    'data_source' : 'CDS'
    }

# Download required ERA5 files:
ls2d.download_era5(settings)

# Read ERA5 data, and calculate derived properties:
era = ls2d.Read_era5(settings)

# Calculate large-scale forcings:
era.calculate_forcings(n_av=0, method='2nd')

# Simple linearly stretched grid, where each level increases with 1%:
grid = ls2d.grid.Grid_linear_stretched(kmax=128, dz0=20, alpha=0.01)

# Interpolate ERA5 to LES model levels.
les_input = era.get_les_input(grid.z)

# Further process data, or save as NetCDF:
les_input.to_netcdf('ls2d_output.nc')
```

Figure 1. $(\mathrm{LS})^2$D Python script: Simplified example of the $(\mathrm{LS})^2$D workflow in Python.

where $\partial Z^{\mathrm{LS}}/\partial x_i$ are the horizontal gradients of the geopotential height on constant pressure levels, and g is the gravitational acceleration.

All terms from Eq. 1 and Eq. 2 are calculated on the native (137) ERA5 model levels, as this provides the most detail in the vertical. The geostrophic wind components are an exception, which are calculated on pressure levels, and next interpolated to the model levels. All terms additionally vary in time. They are, however, applied as horizontally homogeneous quantities in the LES domain. Although it is in principle possible to calculate and apply spatially varying tendencies, this approach would be difficult to unite with the choice for periodic boundary conditions in LES.

The gradients $\partial \phi^{\mathrm{LS}}/\partial x_i$ are approximated from the ERA5 data with either second or fourth-order accurate centered finite differences:

$$\left.\frac{\partial \phi}{\partial x_i}\right|_{\mathrm{2nd}}^{j} \approx \frac{\phi^{j+1} - \phi^{j-1}}{2\delta}, \tag{5}$$

$$\left.\frac{\partial \phi}{\partial x_i}\right|_{\mathrm{4th}}^{j} \approx \frac{\phi^{j-2} - 8\phi^{j-1} + 8\phi^{j+1} - \phi^{j+2}}{12\delta}. \tag{6}$$

2.2 Python package

The main task of $(\mathrm{LS})^2$D is to automate all the steps required to generate the input for doubly periodic LES. To explain the typical workflow with $(\mathrm{LS})^2$D, we will step-by-step walk through a simplified $(\mathrm{LS})^2$D Python script, shown in Fig. 1.

After installing the Python package, $(\mathrm{LS})^2$D is available with `import ls2d`. To simplify passing the case settings to $(\mathrm{LS})^2$D, most settings are gathered in a Python dictionary. Next, the first step is to download the ERA5 data, if the required NetCDF files are not already available. $(\mathrm{LS})^2$D supports two methods, either using the Copernicus Data

Store (CDS, openly available), or the Meteorological Archival and Retrieval System (MARS, requires ECMWF supercomputer access). In both cases, and especially when using CDS, the queuing time can be substantial, as the model level data is only stored in the tape archive. Therefore, when using the CDS option, $(\mathrm{LS})^2\mathrm{D}$ will store the unique CDS request IDs once the downloads are submitted, and stop the Python script. On a subsequent launch of the Python script, `download_era5()` will check if there are active request, if so check if they are finished, and if this is the case, download the NetCDF files from the CDS server. When using MARS to download the ERA5 data, the MARS retrievals are submitted using the SLURM workload manager.

Once the required ERA5 NetCDF files are available, $(\mathrm{LS})^2\mathrm{D}$ has the `Read_era5()` routine which reads the required files, and calculates some derived properties like the model level pressure and height, and state variables in other units. Next, `calculate_forcings()` calculates the required terms from Eq. 1 and Eq. 2. The `n_av` option allows the user to average the initial conditions and large-scale forcings over ±`n_av` ERA5 grid points, and the `method` argument switches between the second and fourth-order accurate finite differences.

After specifying a vertical LES grid (in this case a stretched grid, where the k-th level has a grid spacing $\Delta z = \Delta z_0(1+\alpha)^k$), the `get_les_input()` routine interpolates all vertical profiles to the specified LES grid. The resulting `les_input` object is an xarray dataset (Hoyer & Hamman, 2017) containing all the required LES input, and additional information related to surface variables like e.g. the roughness lengths, leaf area index, and vegetation and soil types from ERA5. Each variable contains attributes describing the variable and its units. This dataset can easily be saved in NetCDF format using xarray's `to_netcdf()` method, or further processed in the $(\mathrm{LS})^2\mathrm{D}$ script, and saved into the LES model specific input format.

3 Simulation setup and post-processing

3.1 LES model

We used MicroHH for all LES simulations in this study. The core of the model is described in detail by van Heerwaarden et al. (2017). Over the last years, the model has been extended with various new physics parameterizations required for simulations of realistic weather. This includes the RTE-RRTMGP radiative transfer model (Pincus et al., 2019), an interactive land-surface model based on HTESSEL (Balsamo et al., 2009), and a single moment ice microphysics scheme (Tomita, 2008).

Describing all model settings is practically unfeasible for such complex LES setups. Therefore, we provide a limited description of the primary settings required to understand the nature of the experiments. The LES code and input are available as supplementary material, and all options can be found in our online documentation (https://microhh.readthedocs.io).

The LES domain size and horizontal resolution varied between the different experiments, as specified in Table 3.2. In the vertical direction, we used a stretched grid with a grid spacing of 20 m near the surface, stretched over 192 levels up to ∼18 km height. For advection, we used the 5^{th} order scheme from Wicker & Skamarock (2002), and a non-dynamic Smagorinsky subfilter-scale model. The radiative fluxes, calculated over the full 3D fields, were updated every 60 simulation seconds. The experiments used a spatially homogeneous surface, with short grass with a vegetation fraction of 95% and a leaf area index (LAI) of 2.6, and the medium-fine soil type from IFS (ECMWF, 2018). Finally, the nudging timescale τ_{n} (Eq. 1, 2) was set to three hours.

Each month is simulated as 31 individual days, and each simulated day is started at 22:00 UTC the previous day, with an integration time of 26 hours. In the post-processing

Table 1. Overview sensitivity experiments LES

Name	Horizontal size	Resolution
`S`	6.4 km	100 m
`M`	12.8 km	100 m
`L`	25.6 km	100 m
`XL`	51.2 km	100 m
`S-MR`	6.4 km	50 m
`S-HR`	6.4 km	25 m

of the statistics, the first two hours of each simulation are discarded as spin up, resulting in continuous one month long time series.

3.2 Post-processing

We used model output from *virtual observation sites* (individual LES grid points, without any spatial or time averaging), sampled at a 5-second frequency, to calculate most LES statistics. For these individual columns, radiative transfer was also diagnosed at a 5-second frequency. For the comparison with observations, we averaged the LES output over a 10-minute window, in line with the time averaging of the Cabauw observations. Non-averaged individual column statistics were used for the results in Section 4.3, to preserve the highest amount of variability.

4 Results

The results section starts by providing a first impression of the typical weather variability over Cabauw, and the ability of both the $(LS)^2$D-MicroHH combination (hereafter simply referred to as LES) and ERA5 itself, to capture this variability (Section 4.1). Next, we examine the skill of all models by statistically comparing LES and ERA5 with the Cabauw observations (Section 4.2). Finally, in Section 4.3 the surface solar irradiance variability is studied.

4.1 Characterization of diurnal variability

The weather over Cabauw is often highly variable, with different weather types within the time span of a few days. To give a first impression of the ability of ERA5 and LES to capture this variability, we present an overview of a short time period with varying weather conditions (11th to 16th of August) in Fig. 2.

The period started with the passage of a warm front with light rain on the 11th of August, with overcast skies which did not clear until noon on the 12th. The 13th and 14th were characterized by nearly overcast but broken cumulus and/or stratocumulus clouds, followed by a perfect diurnal cycle of shallow cumulus on the 15th of August.

In general, both LES and ERA5 perform visually similar (the actual statistics are provided in the next section). Both models capture most of the variation in the surface solar (SW$\downarrow$) and longwave (LW$\downarrow$) irradiance (Fig. 2a,b), and the surface upward longwave radiation (LW$\uparrow$, Fig. 2c). Only on clear nights (12$\rightarrow$13 and 14$\rightarrow$15 August) both models overestimate LW $\uparrow$, indicating that the modelled surface temperatures are too high. The surface sensible (H) and latent ($\mathrm{L_v E}$) heat fluxes (Fig. 2 d,e) are in line with the observations as well. The Bowen ratio $\beta \equiv \mathrm{L_v E/H} \approx 1/3$ is typical for Cabauw, and correctly reproduced by ERA5 and the land-surface model in LES. As a result, the 10-

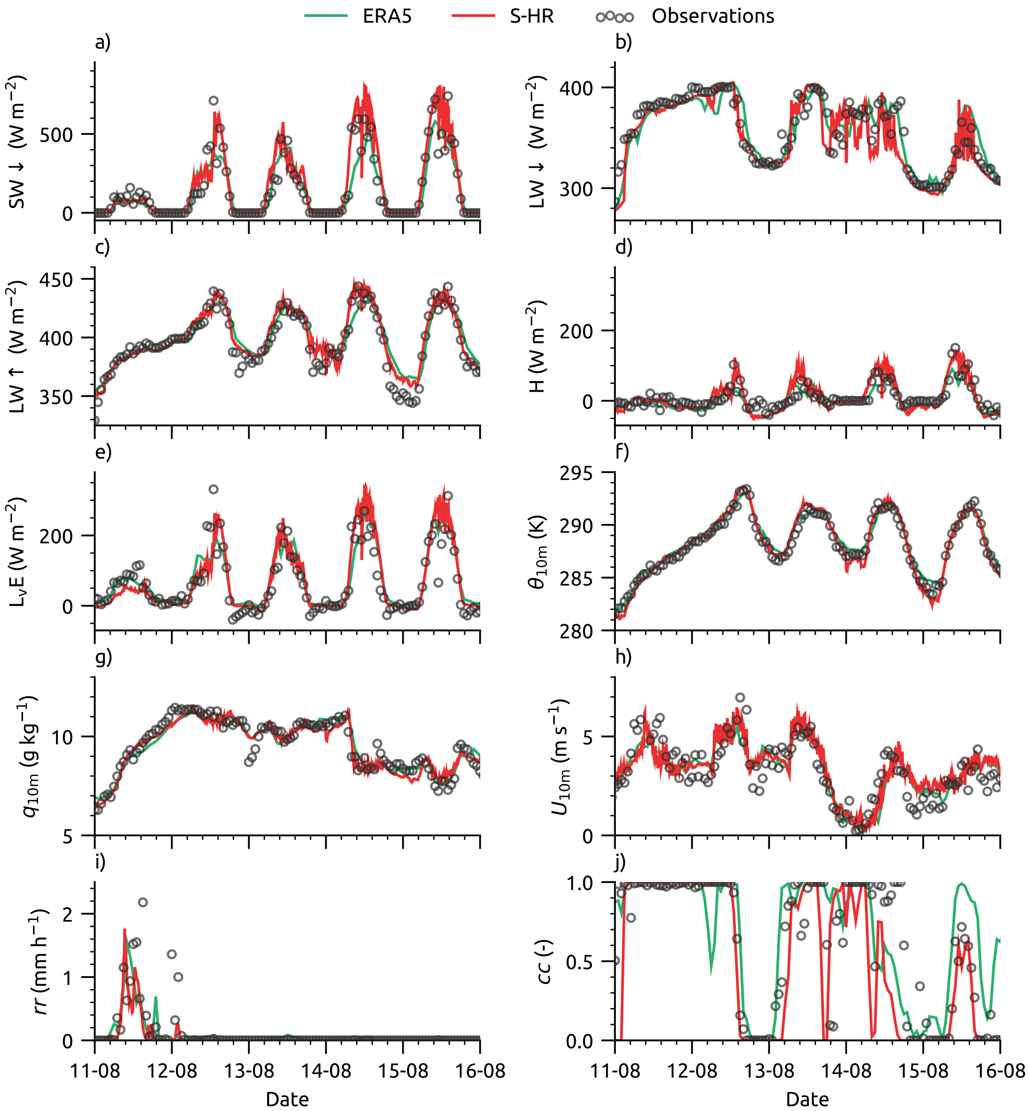

Figure 2. First impression: overview of 5 days to illustrate the day-to-day variability over Cabauw. The variables include a) the surface solar irradiance, b) surface longwave irradiance, c) surface outgoing longwave radiation, d) surface sensible heat flux, e) surface latent heat flux, f) 10 m potential temperature, g) 10 m specific humidity, h) 10 m wind speed, i) surface rain rate, and j) total cloud cover.

meter potential temperature, specific humidity, and wind speed (Fig. 2 f,g,h) closely follow the observations. However, both models fail to capture some of the fast observed fluctuations in wind speed. This is perhaps expected in ERA5, with its coarse spatial and temporal resolution, but LES should in theory capture these fluctuations. The surface rain rate (Fig. 2i) during the frontal passage was weak at around 0-2 mm per hour, which is reproduced by the microphysics schemes in both IFS and LES, although in both models precipitation stops a bit early. Most of the variability in cloud cover (Fig. 2j) is captured by both models. On the last day with shallow cumulus convection, LES is much closer to the observed cloud fraction compared to ERA5. The skill of both LES and ERA5 in predicting the cloud cover is studied in more detail in Section 4.2.2.

Overall, the $(LS)^2D$ coupling between ERA5 and LES successfully introduces most of the day-to-day variability observed in reality into LES. In the next section, the skill of LES and ERA5 is further analyzed by statistically comparing both models to the Cabauw observations.

4.2 Statistical validation LES and ERA5

4.2.1 Surface and tower observations

The statistical analysis will focus on the mean biases of LES and ERA5. To distinguish between biases in different parts of the diurnal cycle, the statistics have been calculated over three hourly periods. For the LES simulations, we focus on the highest resolution (`S-HR`) and largest domain (`XL`) experiments.

Fig. 3a and 3b show the statistics for the surface shortwave radiation. The mean bias in the incoming radiation in LES is significant at a maximum of 60 to 75 W m^{-2}. In contrast, ERA5 has a mean bias of only -1.5 W m^{-2}. The bias in the outgoing shortwave radiation in LES is almost entirely caused by the bias in the incoming radiation, indicating that the albedo used in LES (0.24%) is accurate. The ERA5 grid point for Cabauw, however, has an albedo of around 17%, resulting in a maximum negative bias in the outgoing shortwave radiation of -40 W m^{-2}. The large positive bias in the solar irradiance in LES is most likely caused by an underestimation of clouds in certain weather regimes, which is discussed in more detail in the following sections. As a result of the underestimation of clouds, the surface longwave irradiance (Fig. 3b) is underestimated in LES. However, compared to the solar irradiance, the biases are much smaller at ∼-5 to -10 W m^{-2}. The biases in the surface longwave outgoing radiation (Fig. 3c) are similar in LES and ERA5. During the night, both models overestimate the outgoing longwave radiation, indicating that the surface temperatures are too high.

The biases in the solar radiation in LES (incoming) and ERA5 (outgoing) are not compensated by the biases in longwave radiation, and therefore both models have a net excess of energy at the surface (Fig. 3e). As a result, both the sensible (Fig. 3f) and latent (Fig. 3g) heat fluxes are overestimated. The timing of the overestimation differs between the two variables: in LES, the sensible heat flux almost perfectly follows the net radiation bias, but the bias in the latent heat flux is delayed. During the night, the sensible and latent heat fluxes in LES are too negative, but these small biases are likely within the measurement uncertainty of eddy-correlation flux measurements at night (e.g. de Roode et al., 2010; Bosveld et al., 2020). ERA5 does a slightly better job at predicting the sensible heat flux, but has a positive evaporation bias throughout the diurnal cycle.

As expected, the LES biases in the 10-meter temperature (Fig. 3h) and specific humidity (Fig. 3i) follow the pattern of the surface sensible and latent heat fluxes. For the 10 m temperature, this results in an overestimation of the diurnal amplitude, as the model is too cold at night and too warm during the day. This is in contrast with ERA5, which underestimates the amplitude of $T10_m$, which is a well known problem in IFS (Sandu et al., 2013). The positive evaporation biases in both LES and ERA5 directly translate

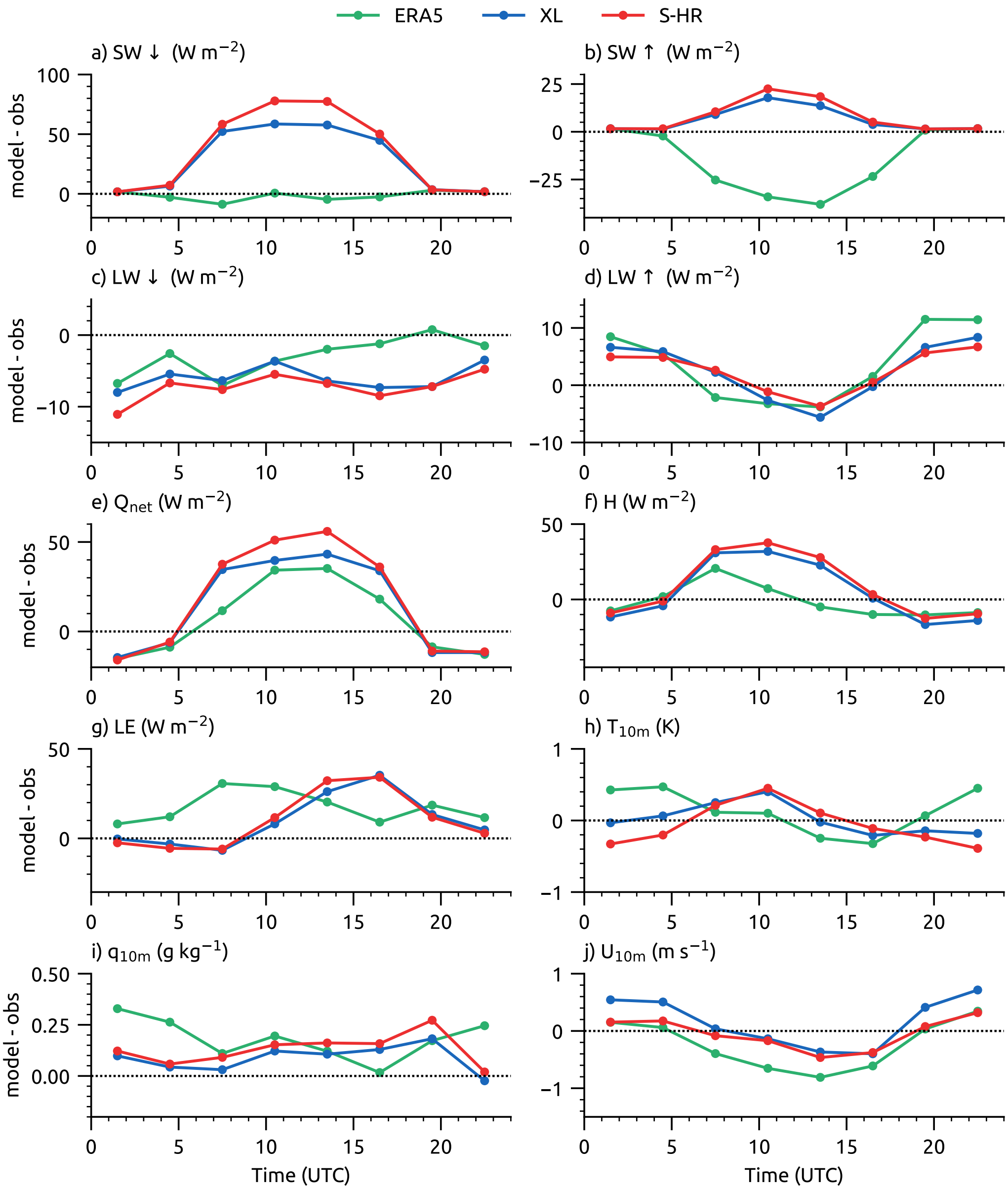

Figure 3. Statistics LES and ERA5 vs. observations: Three-hourly statistics for ERA5 and the largest (`XL`) and highest resolution (`S-HR`) LES runs, showing the mean bias (model minus observations) as a function of the time of the day. The mean bias $\bar{b}$ over the entire period is shown in text in the figures.

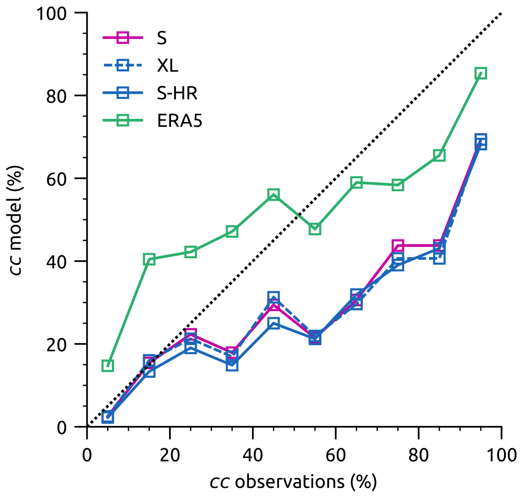

Figure 4. Statistics cloud cover: cloud cover statistics for ERA5 and three of the LES experiments, averaged in 10% bins of the observed cloud fraction.

to relatively small (+0.1-0.2 g kg^{-1}) biases in the 10 m specific humidity. Finally, the biases in the 10 m wind speed (Fig. 3j) show a similar pattern in both LES and ERA5, with a positive bias during the night, and a negative bias during the day. ERA5 has the largest (negative) bias, especially during the day, and in the LES experiments, the high-resolution experiment reduces the biases at night.

In summary, LES seems to perform neither significantly better nor worse than ERA5 for the variables shown in Fig. 3, with the surface solar irradiance being the one clear exception. To further study the irradiance bias, the next section will look at the representation of clouds in the LES experiments.

4.2.2 Clouds

The Cabauw site has multiple instruments measuring cloud properties, including an LD-40 ceilometer (Bosveld et al., 2020). The latter provides time series of both cloud presence and cloud base at a 15-second frequency, which allows for a direct (statistical) comparison with the high frequency (5 second) single column output from MicroHH. From the observations and LES, a cloud fraction is derived by first calculating a cloud mask (criteria: cloud present in the ceilometer output, positive liquid water path in LES), and secondly calculating the cloud fraction as the fraction of time that clouds are present in a one-hour period. As this method can not be applied on the ERA5 dataset, we simply used the total cloud fraction available in the ERA5 statistics.

Figure 4 shows the statistics for the LES experiments and ERA5, binned in 10% intervals as a function of the observed cloud fraction. From these results, we can conclude that the influence of the domain size (`S` vs. `XL`) and resolution (`S` vs. `S-HR`) in LES is small, as all LES experiments show similar biases. For low cloud fractions of $\leq$30% (0-2 okta), the bias in LES is close to zero, whereas ERA5 has a positive bias of $\sim$10-25%. This improvement in LES can be expected as LES explicitly resolves these small clouds. For intermediate cloud fractions of 30-50% (3-4 okta), LES underestimates the cloud fraction with $\sim$15%, where ERA5 has a positive bias of $\sim$10%. For the highest cloud fractions in the range 50-100% (5-8 okta) both LES and ERA5 underestimate the cloud fraction, although the bias in LES is much larger at $\sim$30-40%.

From these results, it is evident that in doubly periodic LESs, with the domain sizes that we used, cloudiness is underestimated for observed cloud fractions larger than 30%. The potential reasons behind these biases, and possible solutions to overcome these biases, are further discussed in Section 6.

4.3 Surface solar irradiance variability

The previous sections statistically compared the time averaged model output to observations, in order to reveal potential biases in LES or ERA5. This section will extend that analysis, but with a focus on quantities where LES can potentially add value over ERA5, as it explicitly resolves the interaction between individual clouds and radiation.

One variable which is of particular interest is the surface solar irradiance, which at Cabauw (BSRN station) is measured at a one-second frequency. As explained by Mol, van Stratum, et al. (2023), solar irradiance variability and cloud properties are closely related. Therefore, we use this quantity to analyze both the high frequency interactions between clouds and radiation, but also the low frequency interactions, which are likely influenced by the biases in cloudiness (Section 4.2.2).

4.3.1 Shadow duration and size distributions

The first part of the analysis will focus on the shadow duration and size. In both LES and the observations, we define a shadow as a continuous period where the difference between the surface solar irradiance (global horizontal irradiance, GHI) and its clear sky (CS) value is more than 50 W m^{-2}, i.e. GHI$_{\mathrm{cs}}$-GHI $>$ 50 W m^{-2}. In LES the clear sky irradiance is output by the RTE-RRTMGP model, for the observations we use the clear sky product from McClear (Lefèvre et al., 2013). The shadow duration is next translated to an approximate shadow length by multiplying it with the 200 m wind speed. This wind speed is likely less than the wind speed in the cloud layer(s), but is the highest observed wind speed available in the Cabauw observations.

Figures 5a and 5b shows the probability density functions (PDFs) of the shadow duration and size. All LES experiments are included to examine the influence of domain size and horizontal resolution. Both the observations and LESs exhibit a power-law relationship with a slope around -5/3 across a wide range of scales. The power-law slope is in line with the findings of Mol, van Stratum, et al. (2023), who studied similar shadow properties using 10 years of Cabauw BSRN data. The similarity in the PDF slope indicates that our chosen period is not anomalous in terms of the shadow (or cloud) properties.

The PDFs of the LES simulations deviate from the observations at a few distinct time and length scales, but differences between the individual LES runs and the observations are difficult to distinguish given the wide range of scales on the vertical axis. Therefore, Fig. 5c,d shows the same PDFs, but normalized by $t_{\mathrm{shadow}}^{-5/3}$ and $l_{\mathrm{shadow}}^{-5/3}$.

At the smallest time and length scales, the influence of the horizontal resolution is obvious. At the highest resolution (`S-HR`) the variability is close to the observations, but as the grid spacing increases, less variability is captured. This is most clearly visible in the shadow lengths, where the LES simulations at Δ=50 m and Δ=100 m seem to drop off at scales below $\sim 2\Delta$. At intermediate time and length scales – between 30 s $>$ t $>$ 10 min, equivalent to 200 m $<$ l $<$ 5000 m – the highest resolution experiment almost perfectly follows the observations, while the lower resolution experiments slightly overestimate the shadow occurrences. At time scales beyond $\sim$10 minutes and $\sim$5 km, all LES experiment show a sharp drop-off. This is to some extent related to the domain size, which limits the maximum cloud size and therefore the shadow durations and sizes. However, in all experiments except `S`, shadow lengths are missing which could

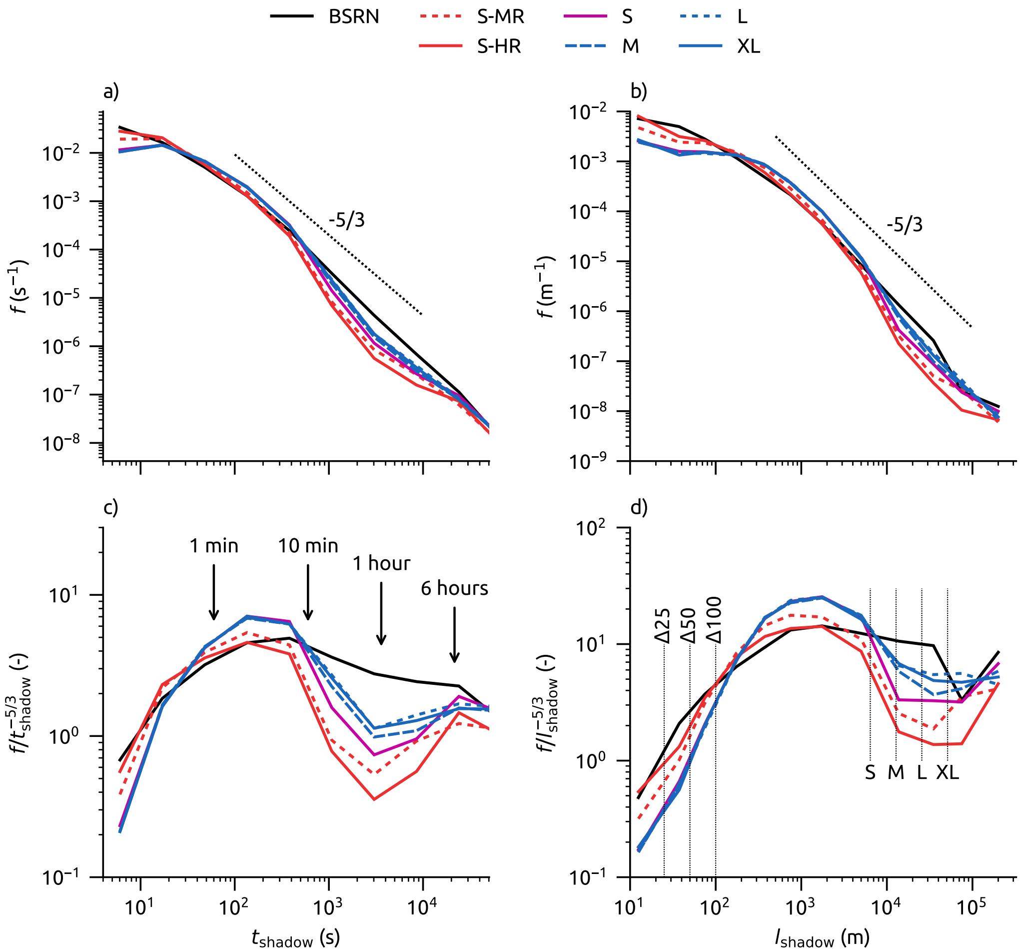

Figure 5. Cloud shadow duration and size: Normalized probability density functions (PDFs) of the cloud shadow duration (a) and size (b) from the BSRN observation and all LES experiments. In the bottom row, the mean slope – typically consisting of a $x^{-5/3}$ slope (Mol, van Stratum, et al., 2023) – has been divided out to improve readability.

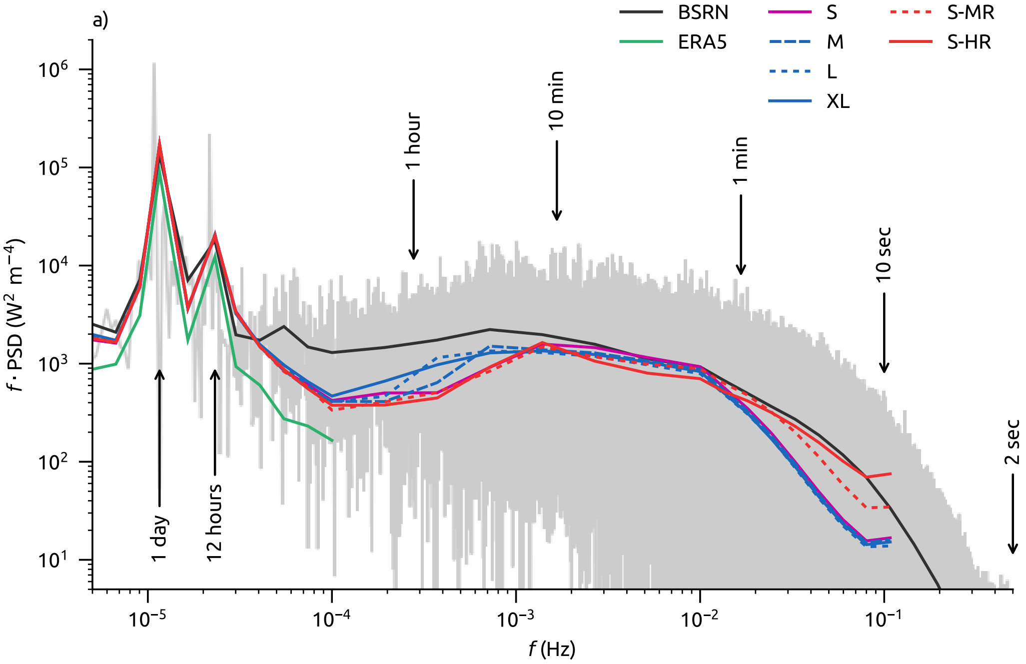

Figure 6. Solar irradiance spectra: Power spectra of the surface solar irradiance, comparing ERA5 and LES with the 1 second BSRN observations. The gray shading is the original (not averaged) PSD of the BSRN data.

fit into the LES domain. Only at the largest time and length scales do the LES experiment again converge with the observations. As these length scales are much larger than the domain sizes, this is simply the result of an overcast cloud deck which is advected multiple times through the periodic boundary conditions.

4.3.2 Solar irradiance power spectrum

The probability density functions from the previous section analyzed the surface solar irradiance as a discrete (on/off) quantity. To get more insight in the full extent of the solar variability, this section studies power spectra of the BSRN observations, ERA5, and all LES experiments.

The power spectral density (PSD) is obtained from time series using fast Fourier transforms (FFT). As these PSDs tend to be noisy on the original frequencies, we averaged the PSDs over logarithmically increasing bin sizes (e.g. Schalkwijk et al., 2015). For the LES experiments, we additionally averaged the binned PSDs over the five column locations.

Figure 6 shows the power spectral densities. The energy at time scales (τ) corresponding to the diurnal cycle and daylight period is clearly visible, and captured correctly by the LES experiments. ERA5 is close to the observations, but underestimates the variability slightly at $\tau > 12$ hours. At $\tau < 12$ hours, ERA5 clearly underestimates the variability at all time scales that can potentially be resolved with the hourly output. The LES spectra also drop off at around $\tau=$ 7 hours, although less than ERA5, and only converge with the observations around $\tau =$ 10-15 minutes. The larger LES domains (`M`, `L`, `XL`) resolve more energy at time scales between 10 minutes and 2 hours, but still do not capture the full variability seen in the observations. At time scales below 10 minutes, LES can capture the full variability down to the 10-second scale, but only in the highest resolution experiment. When the horizontal resolution is increased from 25 m to 50 m or 100 m, the resolved variability at time scales less than a minute quickly drops off.

The results from Figures 5 and 6 are in line with the findings from the previous sections. At time scales larger than 10-15 minutes – corresponding to a spatial scale of ∼10 km – variability in the surface solar irradiance is underestimated as the result of a misrepresentation of large clouds or conditions with large cloud fractions in LES. However, these results also demonstrate the added value of downscaling ERA5 with LES, as such an LES setup is capable of capturing solar irradiance variability from time scales of 10 minutes down to 10 seconds.

5 Discussion

Our validation with observations showed that (LS)^{2}D, combined with MicroHH, captures most of the day-to-day variability in the weather. However, the results also indicate that such a doubly-periodic LES is not a general purpose tool, suitable for all weather conditions.

The validation of the cloud properties and solar irradiance variability against a comprehensive set of surface and remote sensing observations, revealed that these LESs underestimate high cloud fractions and/or large cloud structures. This is most likely caused by our approach, where LES is only coupled to the large-scale model through a set of horizontally mean large-scale forcings. Consider for example a cloud layer that is correctly predicted by the host model, at a height where LES has not developed any spatial variability in temperature or humidity. If this cloud layer is advected into the LES model only through the horizontal mean state, then all grid points at that height in LES will be either saturated or not, resulting in a cloud deck in LES that essentially behaves as an on/off switch. And if the cloud deck in the host model consists of broken clouds, the relative humidity in LES will likely stay below 100%, and LES will not capture any clouds. Similar issues have recently been reported by Jansson et al. (2022) in a superparameterization setup. A setup where LES is coupled to the host model at the lateral boundaries might solve these issues, as this can result in the advection of spatial variability in e.g. temperature or humidity into the LES domain.

A second issue that we did not discuss in detail is the formal validity of our experiments in the nocturnal boundary layer (NBL). Resolving sufficient turbulence in a weak to moderately stable boundary layer requires a grid spacing of $\mathcal{O}(1)$ m (e.g. Beare et al., 2006). In addition, LES might simply not be the correct tool when studying more strongly stable conditions with little or no turbulence. These resolution requirements and other limitation make it difficult to unite a valid LES setup for nighttime conditions, with a domain large enough for (deep) convection. An LES setup with two-way grid nesting could potentially solve some of these issues, by letting the high-resolution domain feed back the correct mean thermodynamic state to its larger parent domain. Alternatively, for studies where the NBL is of secondary importance, the early morning biases could be ignored, as they are unlikely to influence processes the following day (van Stratum & Stevens, 2015, 2018).

6 Summary and conclusion

This paper presented (LS)^{2}D: an open-source Python package designed to simplify downscaling ERA5 with doubly-periodic LES. With a number of one month long simulations over Cabauw (the Netherlands), consisting of various sensitivity experiments on domain size and resolution with the MicroHH LES model, we demonstrated both the benefits and challenges of using such a realistic but doubly-periodic LES setup.

Overall, the combination of $(\mathrm{LS})^2\mathrm{D}$ and MicroHH manages to capture most of the day-to-day variability in the weather. However – as discussed in the previous section – the model setup has difficulties with capturing conditions with large cloud fractions and/or structures. The downscaling method that we presented is therefore most useful for studying processes that are internally resolved by LES, but poorly resolved or parameterized by the host model. This way, the host model sets the large-scale environment, in which LES resolves the small scale processes. A key example, important for both weather and climate, is shallow convection, which is fully parameterized by most large-scale weather and climate models, but explicitly resolved by LES. For these conditions, our comparison with the Cabauw observations showed that LES improves on its host model when predicting e.g. cloud cover.

Being able to resolve shallow convection accurately has an important implication, as it enables the use of this model setup to study the complex interplay between radiative transfer, land-surface processes, turbulent transport, and moist convection. As we have demonstrated, a sufficiently high resolution LES setup manages to capture solar irradiance variability across a wide range of temporal scales, all the way down to a timescale of seconds. This is important for correctly modeling land-atmosphere interactions, but also provides the potential to use LES for forecasting e.g. irradiance variability for solar energy applications.

In summary, the downscaling of ERA5 with $(\mathrm{LS})^2\mathrm{D}$ and LES has proven to be a useful tool to advance our understanding on the interplay between several key atmospheric processes, when applied to meteorological appropriate conditions.

Acknowledgments

MicroHH is open-source and available at https://github.com/microhh/microhh. $(\mathrm{LS})^2\mathrm{D}$ is open-source and available at https://github.com/LS2D/LS2D and https://pypi.org/project/LS2D/. The MicroHH and $(\mathrm{LS})^2\mathrm{D}$ source codes, setup of the LES cases, and the NetCDF statistics used for the analysis in this paper, are archived at Zenodo (https://doi.org/10.5281/zenodo.7797512). The Cabauw observations are available at https://dataplatform.knmi.nl/dataset/?q=Cabauw

B.v.S. acknowledges funding from the Ruisdael Observatory, co-financed by the Dutch Research Council (NWO) under grant number 184.034.015, and the European Union's Horizon 2020 research and innovation programme under grant agreement No 958927 (Prototype system for a Copernicus CO2 service). C.v.H and B.v.S. acknowledge funding from the Dutch Research Council (NWO), grant number VI.Vidi.192.068. The simulations were carried out on the Dutch national e-infrastructure with the support of SURF Cooperative (project number NWO-2021.010)

References

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., & Betts, A. K. (2009). A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. *J. Hydrometeor.*, *10*(3), 623–643. doi: 10.1175/2008JHM1068.1

Beare, R. J., Macvean, M. K., Holtslag, A. A. M., Cuxart, J., Esau, I., Golaz, J.-C., … Sullivan, P. (2006). An Intercomparison of Large-Eddy Simulations of the Stable Boundary Layer. *Bound.-Layer Meteor.*, *118*(2), 247–272. doi: 10.1007/s10546-004-2820-6

Bosveld, F. C., Baas, P., Beljaars, A. C. M., Holtslag, A. A. M., de Arellano, J. V.-G., & van de Wiel, B. J. H. (2020). Fifty Years of Atmospheric Boundary-Layer

Research at Cabauw Serving Weather, Air Quality and Climate. *Bound.-Layer Meteor.*, *177*(2), 583–612. doi: 10.1007/s10546-020-00541-w

de Roode, S. R., Bosveld, F. C., & Kroon, P. S. (2010). Dew Formation, Eddy-Correlation Latent Heat Fluxes, and the Surface Energy Imbalance at Cabauw During Stable Conditions. *Bound.-Layer Meteor.*, *135*(3), 369–383. doi: 10.1007/s10546-010-9476-1

ECMWF. (2018). *IFS Documentation CY45R1.* Author.

Fujita, T. T. (1981). Tornadoes and downbursts in the context of generalized planetary scales. *J. Atmos. Sci.*, *38*(8), 1511–1534. doi: 10.1175/1520-0469(1981)038⟨1511:TADITC⟩2.0.CO;2

Gentine, P., Massmann, A., Lintner, B. R., Hamed Alemohammad, S., Fu, R., Green, J. K., … Vilà-Guerau de Arellano, J. (2019). Land–atmosphere interactions in the tropics–a review. *Hydrology and Earth System Sciences*, *23*(10), 4171–4197. doi: 10.5194/hess-23-4171-2019

Gustafson, W. I., Vogelmann, A. M., Li, Z., Cheng, X., Dumas, K. K., Endo, S., … Xiao, H. (2020). The large-eddy simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) activity for continental shallow convection. *Bull. Amer. Meteor. Soc.*, *101*(4), E462–E479. doi: 10.1175/BAMS-D-19-0065.1

Heinze, R., Dipankar, A., Henken, C. C., Moseley, C., Sourdeval, O., Trömel, S., … others (2017). Large-eddy simulations over Germany using ICON: A comprehensive evaluation. *Quart. J. R. Meteorol. Soc.*, *143*(702), 69–100. doi: 10.1002/qj.2947

Heinze, R., Moseley, C., Böske, C., Muppa, S., Maurer, V., Raasch, S., & Stevens, B. (2017). Evaluation of large-eddy simulations forced with mesoscale model output for a multi-week period during a measurement campaign. *Atmos. Chem. Phys.*, *17*, 7083–7109. doi: 10.5194/acp-17-7083-2017

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … others (2020). The ERA5 global reanalysis. *Quart. J. R. Meteorol. Soc.*, *146*(730), 1999–2049. doi: 10.1002/qj.3803

Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled Arrays and Datasets in Python. *J. Open Res. Softw.*, *5*(1), 10. doi: 10.5334/jors.148

Huang, H.-Y., & Margulis, S. A. (2011). Investigating the impact of soil moisture and atmospheric stability on cloud development and distribution using a coupled large-eddy simulation and land surface model. *J. Hydrometeor.*, *12*(5), 787–804. doi: 10.1175/2011JHM1315.1

Jansson, F., van den Oord, G., Pelupessy, I., Chertova, M., Grönqvist, J. H., Siebesma, A. P., & Crommelin, D. (2022). Representing Cloud Mesoscale Variability in Superparameterized Climate Models. *Adv. Model. Earth Syst.*, *14*(8), e2021MS002892. doi: 10.1029/2021MS002892

Knap, W. (2022). Basic and other measurements of radiation at station Cabauw (2005-02 et seq). *Koninklijk Nederlands Meteorologisch Instituut, De Bilt*. doi: 10.1594/PANGAEA.940531

Lefèvre, M., Oumbe, A., Blanc, P., Espinar, B., Gschwind, B., Qu, Z., … Morcrette, J.-J. (2013). McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmos. Meas. Tech.*, *6*(9), 2403–2418. doi: 10.5194/amt-6-2403-2013

Mol, W. B., Knap, W. H., & van Heerwaarden, C. C. (2023, January). Ten years of 1 Hz solar irradiance observations at Cabauw, the Netherlands, with cloud observations, variability classifications, and statistics. *Earth System Science Data Discussions*, 1–19. (Publisher: Copernicus GmbH) doi: 10.5194/essd-2022-456

Mol, W. B., van Stratum, B. J. H., Knap, W. H., & van Heerwaarden, C. C. (2023). Reconciling observations of solar irradiance variability with cloud size distributions. *Journal of Geophysical Research: Atmospheres*, *n/a*(n/a), e2022JD037894.

doi: 10.1029/2022JD037894
Neggers, R., Siebesma, A. P., & Heus, T. (2012). Continuous Single-Column Model Evaluation at a Permanent Meteorological Supersite. *Bull. Amer. Meteor. Soc.*, *93*(xx), 1389–1400. doi: 10.1175/BAMS-D-11-00162.1
Pedruzo-Bagazgoitia, X., Ouwersloot, H., Sikma, M., Van Heerwaarden, C., Jacobs, C., & Vilà-Guerau de Arellano, J. (2017). Direct and diffuse radiation in the shallow cumulus–vegetation system: Enhanced and decreased evapotranspiration regimes. *J. Hydrometeor.*, *18*(6), 1731–1748. doi: 10.1175/JHM-D-16-0279.1
Pincus, R., Mlawer, E. J., & Delamere, J. S. (2019). Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models. *Adv. Model. Earth Syst.*, *11*(10), 3074–3089. doi: 10.1029/2019MS001621
Ražnjević, A., van Heerwaarden, C. C., van Stratum, B. J., Hensen, A., Velzeboer, I., van den Bulk, P., & Krol, M. C. (2022). Technical note: Interpretation of field observations of point-source methane plume using observation-driven large-eddy simulations. *Atmos. Chem. Phys.*, *22*(10), 6489–6505. doi: 10.5194/acp-22-6489-2022
Rieck, M., Hohenegger, C., & van Heerwaarden, C. C. (2014). The influence of land surface heterogeneities on cloud size development. *Mon. Wea. Rev.*, *142*(10), 3830–3846. doi: 10.1175/MWR-D-13-00354.1
Sandu, I., Beljaars, A., Bechtold, P., Mauritsen, T., & Balsamo, G. (2013). Why is it so difficult to represent stably stratified conditions in numerical weather prediction (NWP) models? *Adv. Model. Earth Syst.*, *5*(2), 117–133. doi: 10.1002/jame.20013
Schalkwijk, J., Jonker, H. J., Siebesma, A. P., & Bosveld, F. C. (2015). A year-long large-eddy simulation of the weather over Cabauw: An overview. *Mon. Wea. Rev.*, *143*(3), 828–844. doi: 10.1175/MWR-D-14-00293.1
Schemann, V., Ebell, K., Pospichal, B., Neggers, R., Moseley, C., & Stevens, B. (2020). Linking large-eddy simulations to local cloud observations. *Adv. Model. Earth Syst.*, *12*(12), e2020MS002209.
Talbot, C., Bou-Zeid, E., & Smith, J. (2012). Nested mesoscale large-eddy simulations with WRF: Performance in real test cases. *J. Hydrometeor.*, *13*(5), 1421–1441. doi: 10.1175/JHM-D-11-048.1
Tijhuis, M., van Stratum, B. J. H., Veerman, M. A., & van Heerwaarden, C. C. (2023). An Efficient Parameterization for Surface Shortwave 3D Radiative Effects in Large-Eddy Simulations of Shallow Cumulus Clouds. *Adv. Model. Earth Syst.*, *15*(1), e2022MS003262. doi: 10.1029/2022MS003262
Tomita, H. (2008). New microphysical schemes with five and six categories by diagnostic generation of cloud ice. *J. Meteor. Soc. of Japan*, *86*, 121–142. doi: 10.2151/jmsj.86A.121
van Heerwaarden, C. C., Van Stratum, B. J., Heus, T., Gibbs, J. A., Fedorovich, E., & Mellado, J.-P. (2017). MicroHH 1.0: a computational fluid dynamics code for direct numerical simulation and large-eddy simulation of atmospheric boundary layer flows. *Geosci. Model Dev.*, *10*, 3145–3165. doi: 10.5194/gmd-10-3145-2017
van Laar, T. W., Schemann, V., & Neggers, R. A. (2019). Investigating the diurnal evolution of the cloud size distribution of continental cumulus convection using multiday LES. *J. Atmos. Sci.*, *76*(3), 729–747. doi: 10.1175/JAS-D-18-0084.1
van Stratum, B. J. H., & Stevens, B. (2015). The influence of misrepresenting the nocturnal boundary layer on idealized daytime convection in large-eddy simulation. *Adv. Model. Earth Syst.*, *7*(2), 423–436. doi: 10.1002/2014MS000370
van Stratum, B. J. H., & Stevens, B. (2018). The Impact of Vertical Mixing Biases in Large-Eddy Simulation on Nocturnal Low Clouds. *Adv. Model. Earth Syst.*, *10*(6), 1290–1303. doi: 10.1029/2017MS001239
Veerman, M. A., Pedruzo-Bagazgoitia, X., Jakub, F., Vilà-Guerau de Arellano, J., & van Heerwaarden, C. (2020). Three-dimensional radiative effects by shallow

cumulus clouds on dynamic heterogeneities over a vegetated surface. *Adv. Model. Earth Syst.*, *12*(7), e2019MS001990. doi: 10.1029/2019MS001990

Veerman, M. A., van Stratum, B. J. H., & van Heerwaarden, C. C. (2022). A Case Study of Cumulus Convection Over Land in Cloud-Resolving Simulations With a Coupled Ray Tracer. *Geophys. Res. Lett.*, *49*(23), e2022GL100808. doi: 10.1029/2022GL100808

Vilà-Guerau de Arellano, J., Hartogensis, O., Benedict, I., de Boer, H., Bosman, P. J. M., Botía, S., … van Heerwaarden, C. C. (2023). Advancing understanding of land–atmosphere interactions by breaking discipline and scale barriers. *Annals of the New York Academy of Sciences*, *n/a*(n/a). doi: 10.1111/nyas.14956

Wicker, L. J., & Skamarock, W. C. (2002). Time-splitting methods for elastic models using forward time schemes. *Mon. Wea. Rev.*, *130*(8), 2088–2097. doi: 10.1175/1520-0493(2002)130⟨2088:TSMFEM⟩2.0.CO;2