# A Method to Improve Land Use Representation for Weather Simulations Based on High-Resolution Data Sets-Application to Corine Land Cover Data in the WRF Model

M. de Bode, T. Hedde, P. Roubin, P. Durand

AGU ADVANCING EARTH AND SPACE SCIENCE

Earth and Space Science

RESEARCH ARTICLE
10.1029/2021EA002123

**Special Section:**
Advances in scaling and modeling of land-atmosphere interactions

# A Method to Improve Land Use Representation for Weather Simulations Based on High-Resolution Data Sets—Application to Corine Land Cover Data in the WRF Model

**M. de Bode[1,2], T. Hedde[1], P. Roubin[1], and P. Durand[2]**

[1]CEA, DES, IRESNE, DTN, Laboratory for Environmental Transfer Modeling, Saint-Paul-Lez-Durance, France, [2]Laboratoire d'Aérologie, Université de Toulouse, CNRS, UPS, Toulouse, France

**Key Points:**

- A new grid aggregation method (MLM) for land use (LU) categories is described, which addresses ambiguities in assigning the LU to cells
- Simulations with a physical parameter table specific for the Corine Land Cover (CLC) classes show limited improvements over the conversion to the USGS table
- Using CLC as an LU map improves the Weather Research Forecasting compared to USGS maps

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

M. de Bode,
Michiel.de.bode@rivm.nl

**Citation:**

de Bode, M., Hedde, T., Roubin, P., & Durand, P. (2023). A method to improve land use representation for weather simulations based on high-resolution data sets—Application to Corine Land Cover data in the WRF model. *Earth and Space Science*, *10*, e2021EA002123. https://doi.org/10.1029/2021EA002123

Received 9 NOV 2021
Accepted 14 NOV 2022

**Author Contributions:**

**Conceptualization:** M. de Bode, T. Hedde, P. Roubin, P. Durand
**Data curation:** T. Hedde, P. Roubin, P. Durand



**Abstract** Land cover (LC) data incorporation, for weather modeling purposes, highlights many problems. The straightforward Single Level Mode (SLM) aggregation is not adapted for high-resolution LC maps, with a high number of classes, because it could generate false classifications. We propose a Multi-Level Mode (MLM) aggregation method that includes a hierarchical structure. This study focuses on the Corine Land Cover (CLC) data set. Differences between MLM and SLM methods are small at the finest horizontal resolution and increase to a value of around 16% at 9-km horizontal resolution. To further integrate CLC data into WRF (Weather Research Forecasting model), we included a dedicated table of physical parameters next to using the classical conversion toward the USGS one. To evaluate the LC impact on the modeled boundary layer, we used WRF at 1 km to simulate a 4-day period with diurnal cycles of valley winds in the heterogeneous western pre-Alps area. We tested three LCs (USGS, MODIS, and CLC), where CLC has two physical parameter tables and two aggregation methods. CLC performs better than other tested LCs. The CLC aggregation methods revealed limited differences between the simulated variables, although the MLM method gives better results. Since these comparisons are restricted to a single location with the same LC type, the differences between various simulations regarding atmospheric parameters probably result from horizontal advection from upwind areas where surface conditions differ. Extended time series with multiple locations and diverse meteorological conditions need to be considered to reinforce the results presented here.

## 1. Introduction

In meteorological simulations, it is crucial to accurately represent the exchange of momentum, energy, and matter between the surface and the atmosphere. These exchanges ("surface fluxes") are computed through parameterizations, relying on both a theoretical framework (i.e., a set of equations) and a set of physically measured coefficients required in the equations. For example, the friction velocity, which reflects the slowing down of the flow on the surface, is related to the wind in the surface layer and the value of the roughness length ($z_0$), which is a physical characteristic of the surface. Surface energy balance, in which surface albedo and emissivity are driving surface characteristics for the distribution between the radiative terms of the surface energy balance, of which the sensible heat flux is a key component. Another critical component in this balance is the ground heat flux, dependent on the physical characteristics of the vegetation, soil, and soil water content. All these coefficients depend on the surface land cover (LC).

Consequently, the representation of LC in models through the appropriate parameters not only has a direct influence on the simulated processes close to the surface (Oke, 2002), but it also affects mesoscale circulation (Hartmann, 2015; Weaver & Avissar, 2001; Yang, 2004), and LC is crucial for correct simulations by numerical weather prediction (NWP) models (Jach et al., 2020). Most models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al., 2019), often accommodate simulations of both small and large scales in a so-called grid-nesting mode for resource-saving and numerical error reduction purposes (Daniels et al., 2016; Wang & Gill, 2012). This means that the fluxes are computed at different resolutions according to the different domains, and the question arises as to how to define an appropriate way to aggregate either the surface parameters or the fluxes to the considered cell size.

Two main approaches exist for using the LC information available in finer detail than the resolution of the grid (subgrid variability). The first approach selects a unique representative LC from the available categories for each cell, often through the dominant approach (most common value a.k.a. SLM; Single Level Mode) or the nearest

**Formal analysis:** M. de Bode, T. Hedde, P. Durand
**Funding acquisition:** T. Hedde
**Investigation:** M. de Bode, T. Hedde, P. Roubin, P. Durand
**Methodology:** M. de Bode, T. Hedde, P. Durand
**Project Administration:** T. Hedde, P. Roubin, P. Durand
**Resources:** T. Hedde, P. Roubin, P. Durand
**Software:** M. de Bode, T. Hedde, P. Roubin
**Supervision:** T. Hedde, P. Roubin, P. Durand
**Validation:** M. de Bode, T. Hedde, P. Roubin, P. Durand
**Visualization:** M. de Bode, T. Hedde, P. Roubin, P. Durand
**Writing – original draft:** M. de Bode
**Writing – review & editing:** M. de Bode, T. Hedde, P. Roubin, P. Durand

neighbor value. This LC type will represent the entire cell area; hence, other LC types do not contribute to the surface parameters. This is the default option in WRF (Skamarock et al., 2019). The second approach includes all the fractions of LC and calculates fluxes based on the surface of each LC category. For example, WRF has the Mosaic option to include subgrid variation (Li et al., 2013). Both methods highlight the need for good representative data sets.

Information on the LC type originates from data sets such as the commonly used USGS (Loveland et al., 2000) and MODIS (Broxton et al., 2014). Other data sets with more local coverage are Corine Land Cover (CLC; EEA, 2020) for Europe and NLCD for the United States (Homer et al., 2020). A map like CLC offers a well-detailed description of LC, and therefore appears to be well adapted for meteorological models which resolve fine scales but since it is in a vector type, it requires a rasterization step (De Meij & Vinuesa, 2014; Golzio et al., 2021; Prósper et al., 2019; Román-Cascón et al., 2021).

Meteorological models often offer an operational configuration in which a given number of classes represent the LC. For example, default WRF runs with surface parameters originating from 20 classes defined in the world-wide, 30″ resolution (∼1 km) MODIS LC database. Each class corresponds to a specific value for each surface parameter through look-up tables (LUTs; with a seasonal variation of certain parameters like Leaf Area Index (LAI), albedo, or roughness length). NWP models are generally equipped with one or more LC databases with their matching physical parameter LUT. For example, WRF may use native USGS or MODIS LC maps. If one wants to run the model with a new LC database containing its own class definitions, a new LUT has to either be defined or the new LC classes have to find equivalent classes in an existing LUT. The creation of a new LUT minimizes the loss of information but is an intensive job as one has to determine the values for each physical parameter and each LC class index. Finding an equivalence with an existing LUT is easier but could lead to a loss of information if the existing LUT possesses fewer classes than the new one. Examples of both exist; the 44 classes of the European CLC data set were condensed to a 13-class subset of the 24 USGS LUT to run the MM5 model by Pineda et al. (2004). The use of a new LC is therefore easy, but at the expense of the initial richness of the LC description. On the other hand, Golzio et al. (2021) built an LUT for the CLC data set so as not to lose the diversity of the 44 classes.

In literature, one finds numerous studies which aim to evaluate the impact of LC representation on the simulation of atmospheric parameters: Schicker et al. (2016) analyzed WRF simulations based on MODIS or CLC instead of the default USGS LC data. Santos-Alamillos et al. (2015) compared simulations with USGS and CLC input maps. Li et al. (2018) investigated the urban heat island with the USGS, CLC, and Urban Atlas (UA). Golzio et al. (2021) analyzed the effect of CLC with 44 classes, MODIS and USGS on heat fluxes. Most studies agree that over European areas, CLC represents the surface better than USGS, which improves weather simulation performance. Most studies convert CLC data to an USGS raster format however, preventing full benefit of the spatial and class detail that CLC offers. In this study, we will test the benefit of the CLC 44-classes LUT over the CLC converted to USGS, compared to MODIS and USGS.

One might assume that the higher the number of classes used to represent the LC, the better the description of the surface. Nevertheless, aggregation errors could be introduced by having a large number of classes as suggested by Román-Cascón et al. (2021). Figure 1 illustrates how the SLM aggregation option could generate a different output over an area according to the initial number of classes available and could lead to a wrong class choice when the class number increases. In this example, the surface is composed of 44% water, 37% broad-leaved forest, and 19% mixed forest. Assuming the SLM approach, water prevails, even though the two forest types closely resemble each other and represent 56% of the area. A convenient spatial aggregation method is therefore required to prevent misclassification when using a large class amount. To address the issue, this paper describes an approach for the performing of aggregation on an LC data set with many classes more correctly than with the SLM approach.

Section 2 describes several LC data sets, presents the proposed aggregation method, and analyses the aggregation results at various resolutions. This section also explains how to keep the detail of the original database through an appropriate LUT. Section 3 contains the description of the case study on which the WRF numerical simulations are run, in a well-instrumented valley of the southern French pre-Alps. Sections 4 describes the performance of the simulations, realized with different LC representations and it is evaluated against the observations. Section 5 consists of the discussion of our results, and finally, Section 6 concludes the paper.

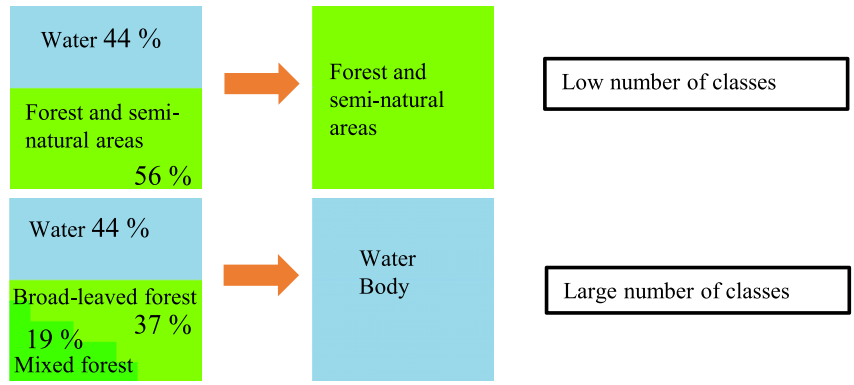


**Figure 1.** Example of Single Level Mode (SLM) aggregation of an area and how the number of available classes affects the results.

## 2. Land Cover Data Sets and Aggregation

### 2.1. Land Cover Data Sets

To test the different means of aggregation, we start by selecting LC maps. The choice of the data set and its associated resolution will depend on the intended purpose (Neumann et al., 2007). Table 1 shows a selection of LC maps suited for meteorological use that are available for research without cost. For our WRF applications (de Bode et al., 2021), we require from a data set a spatial resolution of $O(100\ m)$, many classes for a detailed surface representation, and regular updates. USGS (Anderson et al., 1976) has 24 classes (WRF extends it to 27, but not over Europe), with a 1 km resolution MODIS shares the same characteristics but with a yearly update and 20 classes. Alternatively, CLC has 44 classes in a 3-level hierarchy organization with updates every six years (2012 update is described by Büttner (2014)). CLC is described below. Theia has two 22-class data sets (Inglada et al., 2017, 2018), covering the whole earth and France at 300 m and 10 m resolution, respectively. Lastly, Globeland30 (Gong et al., 2013) has a map with 30 m horizontal resolution, covering the entire world, but it distinguishes only 10 classes. Considering all these options, we will focus on CLC because it covers Europe, has a fine enough spatial resolution, and distinguishes many classes.

### 2.2. Corine Land Cover

CLC contains more detail than most other LC maps and even uses three levels of detail. The first and most basic level describes 5 level-1 classes, namely, "Artificial surfaces," "Agricultural lands," "Forest and seminatural areas," "Wetlands," and "Water bodies." The second detail-level distinguishes 15 level-2 classes, which in turn can be further divided into a third and finest level of 44 classes, as described in https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png and the Supporting Information S1. Furthermore, a few attempts exist to create classifications at a more detailed level by adding another CLC classification level or taking into account,

**Table 1**
*Details of Several LC Data Sets Covering Europe That Are Freely Available for Research Purposes*

| | Name | | | | | |
|---|---|---|---|---|---|---|
| Property | USGS | MODIS | Corine Land Cover | Theia-France | Theia-World | Globeland30 |
| Coverage | World | World | Europe | France | World | World |
| Format | Raster | Raster | Vector | Raster | Raster | Raster |
| Most recent map | 2001–2010 | 2019 | 2018 | 2019 | 2009 | 2020 |
| Update period | Not specified | Yearly | 6 years | Mostly yearly | Not specified | 3 years |
| Resolution | 1 km | 1 km | Minimal width: 100 m | 10 m | 300 m | 30 m |
| | | | Minimal surface: 25 ha polygons | | | |
| Categories | 24 | 20 | 44 | 22 | 22 | 10 |

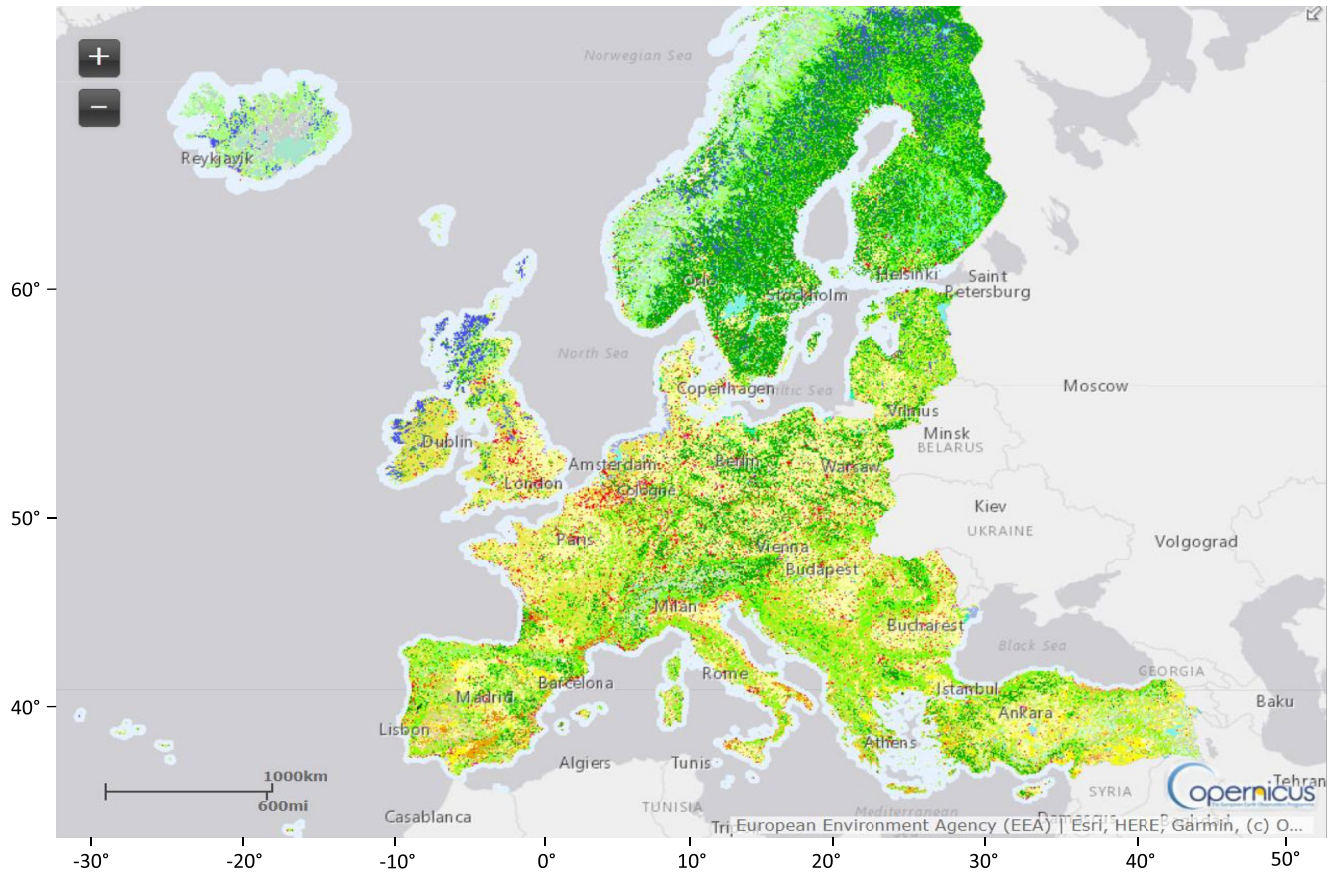


**Figure 2.** 2018 Corine Land Cover (CLC) coverage at detail level 3. For nomenclature see https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png/. Map from: https://land.copernicus.eu/pan-european/corine-land-cover/clc2018.

e.g., the differences in local climate, as is done in the ECOCLIMAP database (Faroux et al., 2013; Gudowicz & Zwolinski, 2016; Ikİel et al., 2012).

The CLC data set is a vector map consisting of polygon shapes, which entails spatial specifications different from those of a raster map. Their delineation is drawn from satellite observations, the latter having a horizontal resolution of 25 m in the older data set (2012) and 10 m in the newest one (Copernicus, 2020), defining the spatial accuracy of the polygons delineation. Every polygon spans an area of at least 25 ha, has a minimum width of 100 m, and is allocated to one of the 44 existing classes (Kosztra et al., 2019). Every 6 years, when an update of the CLC data occurs, an additional map highlights all surface changes larger than 5 ha. An overview of the latest coverage (2018) is displayed in Figure 2. Note that CLC only describes the 39 participating countries (since 2012) but it covers the whole of Western Europe. Sea surfaces are only included when they are within 25 km off the nearest coast.

### 2.3. New Multi-Level Mode Aggregation

#### 2.3.1. Description of the Method

As mentioned in Section 1, using the SLM aggregation technique can result in discrepancies between maps with many classes and maps with a limited amount of classes (Figure 1), with possible artifacts caused by a too high level of detail in the LC database. For this reason, we propose the use of a new multistep aggregation to avoid such drawbacks while taking into account the richness of the initial data set. In order to do this, we use a method we call "Multi-Level Mode" (MLM) aggregation. This method is described below and illustrated in Figure 3 which shows the aggregation of the CLC data set on a real example restrained to 6 × 6 arcmin (~11 × 8 km) domain containing nine cells, each of them with 2 × 2 arcmin (~3.7 × 2.6 km). The figure details the complete process

10.1029/2021EA002123



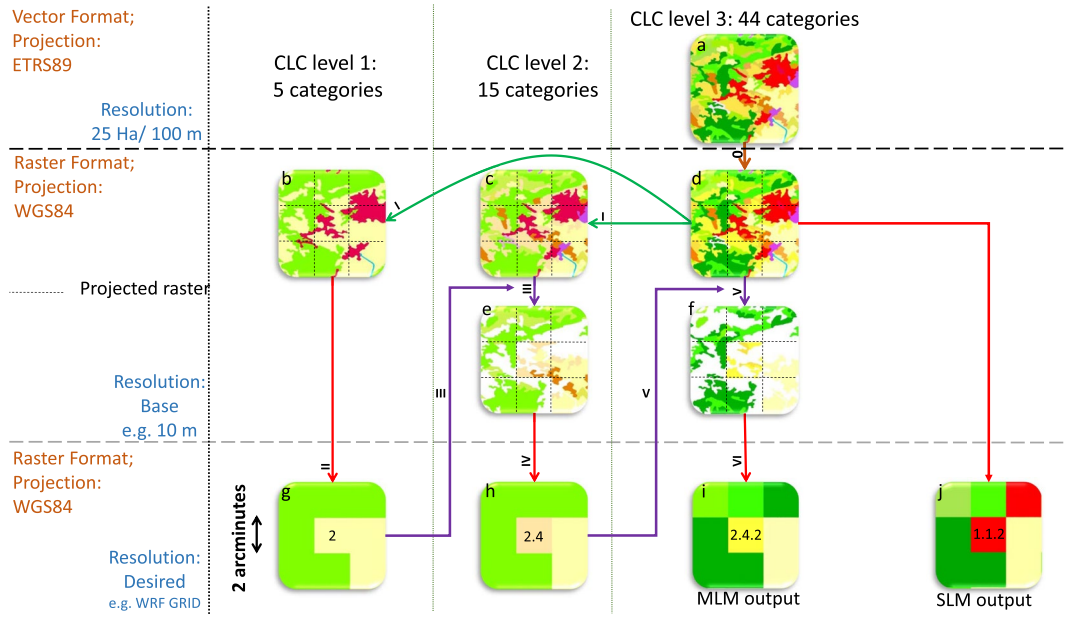


**Figure 3.** Flowchart for the new Multi-Level Mode (MLM) aggregation for categorical data sets. The diagram depicts the steps performed in order to generate a map at a required horizontal resolution (here 2 arcmin or 3.7 × 2.6 km, represented by the dashed lines on finer resolution maps) from the Corine Land Cover (CLC) vector base map (a). Color-coding originates from the Copernicus color structure and is described in more detail at https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png and in Supporting Information S1. Each plot represents the same ∼11 × 8 km area. The arrows indicate the process of each step described in the text. The brown arrow means rasterization and projection (step 0). Green arrows indicate reducing child classes to their parent class (step I). Red arrows denote SLM (steps II, IV, and VI). Purple arrows indicate increasing detailed level of dominant classes (steps III and V). Plots b to f are (⅓ arcsec) resolution maps, g to j are at target resolution (here around 3 km). i is the final map obtained through MLM, j is the one obtained through SLM.

of MLM, starting with the vector map (Figure 3a) until the creation of an aggregated grid (Figure 3i) at the end of the process. The aggregation consists of the following operations:

1. We preprocess the CLC vector data set by rasterizing it (Figure 3a) to a mesh size of ⅓ arcsec (∼10 m). With such a fine raster resolution, there is practically no information loss during this step as it is equal to or better than the satellite resolution from which CLC is computed (25 m for the 2012 data set, 10 m for CLC 2018 data set). The rasterization is made on a regular grid (⅓ arcsec) in the longitude/latitude WGS84 geodesic system (EPSG:4326; Figure 3d).
2. Based on this raster map at ⅓ arcsec resolution, we create CLC level-1 (Figure 3b) and level-2 (Figure 3c) fine-resolution maps by reducing all level-3 child classes to their corresponding parent class level, i.e., contracting 44 level-3 classes to 15 level-2 and 5 level-1 classes, respectively.
3. Here, we apply the SLM on the level-1 map (Figure 3b) in order to generate the first map with the targeted size of the mesh (Figure 3g; 2′ horizontal resolution in our example). Each cell matches the dominant level-1 class (e.g., the central cell belongs to level-1 class **2** "Agricultural areas").
4. Crossing this level-1 map (Figure 3g) with the fine-resolution (⅓ arcsec) level-2 map (Figure 3c), we then build a fine-resolution map with only child classes of the selected dominant level-1 (Figure 3e). The level-2 data which are not children of a selected level-1 class are discarded and represented as white areas in Figure 3e (e.g., at the central cell, all level-1 areas with a value other than 2). The white areas are excluded from the aggregation process.
5. We apply the SLM to the remaining level-2 classes (i.e., nonwhite areas from Figure 3e) so as to extract the level-2 dominant class at the final targeted resolution (2′), represented in Figure 3h. To give an example, the central mesh now belongs to level-2 class **2.4** "Heterogeneous agricultural areas" which is a child class from level-1 class **2** "Agricultural areas."
6. Similarly to step **III**, we cross the previous level-2 map (Figure 3h) with the fine-resolution level-3 map (Figure 3d), and build a fine-resolution map where only child classes areas of dominant level-2 remain (Figure 3f). The level-3 areas which are not child classes of level-2 as selected in Figure 3d are thus discarded

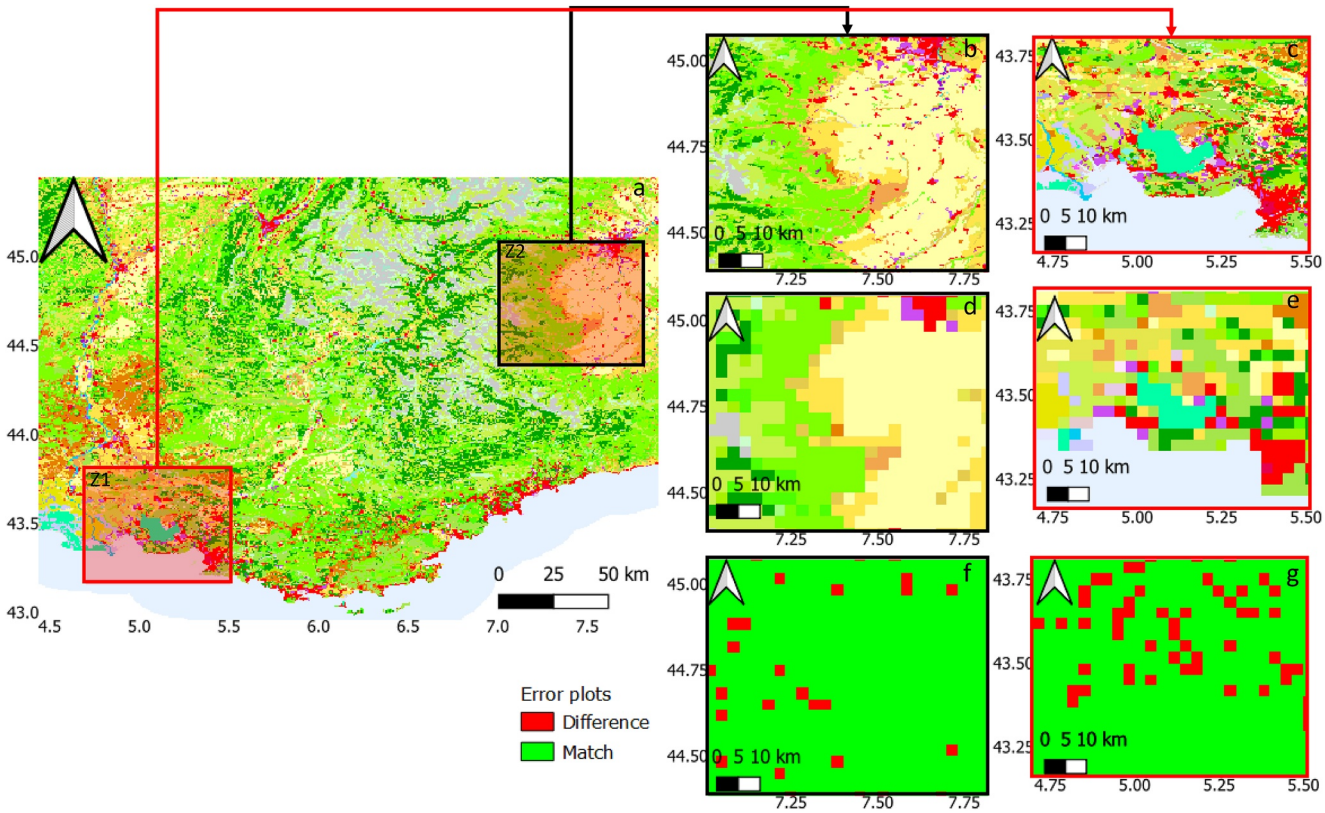


**Figure 4.** Land covers in SE France and westernmost Italy at fine resolution. A zoom-in area Z1 encompasses the shoreline (c, e, and g), and a second zoom-in area Z2 encompasses the French-Italian border (b, d, and f). (b) and (c) show fine-resolution LC map. (d) and (e) show the 2 arcmin MLM-aggregated map for the same areas. (f) and (g) show the cells with a difference between the two aggregation methods in red. A legend for all 44 LC classes is in https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png and Supporting Information S1. The latitude is positive to the north and longitude to the east.

and represented as white areas in Figure 3f. This latter presents even more white areas than in Figure 3e, as, e.g., on the central cell. All level-2 areas with value other than **2.4** were removed.

7. Similarly to step **IV**, we apply the SLM on the remaining level-3 classes (built as explained in **V**, Figure 3f) for selecting level-3 classes at the final targeted resolution (2′), represented in Figure 3i. To give an example, the central cell is now associated to level-3 class **2.4.2** "Complex cultivation patterns."

Figure 3i thus represents the final output of the MLM method. By comparison, Figure 3j shows the results of the traditional SLM applied directly to Figure 3d map, which illustrates how the SLM output can differ from MLM. Figure 3j suggests more "Urban fabric" (red color with class index **1.1.2**) as a dominant type, which could be a consequence of "Urban fabric" being less spread in different level-3 types than the forested areas. On the other hand, in areas where land use is highly variable, the final outcome of the aggregation can be sensitive to the geographic division of the cells. If the 3-km grid borders had been shifted to the north, the "Urban fabric" on the right side of Figure 3d would have occupied the midright cell rather than that of the top right in Figure 3j, and would possibly also appear in Figure 3i a script is available see the data availability statement (de Bode, 2022a).

### 2.3.2. Difference With a Conventional Aggregation Method

We analyze the differences between the two aggregation methods, as illustrated between the "i" and "j" plots of Figure 3, over an area covering the SE of France and NW of Italy. The land cover of this area is represented in Figure 4a, according to the CLC classification. Two zones are highlighted, the first one (Z1) includes the Mediterranean shoreline, as well as the Marseilles conurbation at its eastern side, and the second one (Z2) is purely continental over the French-Italian border (encompassing a large homogeneous cultivated area in the Po valley). From the enlarged plots of Z1 and Z2, we can identify at a glance that the inland part of Z1 presents a much higher diversity than Z2, both at the initial resolution (Figures 4b and 4c), and once reduced at a 2′ scale with the MLM method (Figures 4d and 4e). If we now examine the difference between the two aggregation methods (Figures 4f

**Table 2**
*Difference Overview Between SLM and MLM Aggregation Techniques for Different Grid Resolutions*

| Final horizontal resolution (arc angle) | Number of grid cells | Northwest corner | Southeast corner | Land fraction | Difference overland (%) |
|---|---|---|---|---|---|
| 3″ | 34,560,000 | 3°E, 46°N | 9°E, 42°N | 0.64 | 0.03 |
| 10″ | 3,110,400 | | | 0.64 | 0.5 |
| 30″ | 345,600 | | | 0.63 | 8.3 |
| 2′ | 21,600 | | | 0.63 | 12.0 |
| 5′ | 3,456 | | | 0.63 | 16.6 |
| 5′ | 42,336 | 4°W, 52°N | 16°E, 38°N | 0.80 | 11.6 |

*Note*. The difference metrics exclude all cells with sea from the data. The domain is D03 as shown in Figure 5 except for the last line which corresponds to D02 in Figure 5.

and 4g), we can see that locations with little LC spatial variability, such as sea, large lakes, and large homogeneous agricultural areas (as in Z2), are less affected by the aggregation method than highly variable areas (basically, the heterogeneous part of Z2 and the inland of Z1).

In order to establish a quantitative comparison between the two methods, we computed the difference (in % of the land surface) after aggregation for various horizontal resolutions; the results are presented in Table 2. All the domains cover the same area regardless of the mesh size. Additionally, the coarsest domain (5′ resolution) has a second entry, covering a larger area. The sea cells were excluded for difference computation, since aggregation is straightforward over sea. On a side note, we did not specifically look at coast representation but the MLM maps have more land cells than the SLM maps. When comparing the resolutions over the same domain, we move from a difference of 16.6% at 5′ resolution to a difference of 0.03% at 3″. So, as we expected, for the finest resolution, the two aggregation methods behave quite similarly. Since a cell in a 3″ resolution grid almost covers 1 ha, and the vector shapes have at least 25 ha surface, it leaves little chance for multiple classes to be present in a single cell. On the other hand, increasing the aggregation surface size leaves more classes to be present in a single cell, and the probability of difference between classification increases. In short, the coarser the resolution, the greater the contrast between aggregation methods, with the land cover variability as a strong influencer.

## 3. Description of WRF Simulations

### 3.1. How WRF Uses LC

#### 3.1.1. Physical Parameter Tables

The LC classes are associated to physical parameters through LUTs so they can be used by the land surface model parameterizations. The WRF model proposes such LUTs for USGS, MODIS, and NCLD LC data sets: the "LANDUSE.TBL" file contains summer and winter values of land use parameters, whereas, also depending on LC, the "VEGPARM.TBL" file allows maximum and minimum values of some vegetation parameters (other tables may be required for different land surface schemes, see Section 3.2.1 for our settings). The use of CLC in a model such as WRF requires either the converting of CLC data to an already existing LUT or the creation of a dedicated LUT. Before the study of Golzio et al. (2021), CLC did not have a dedicated LUT of its own.

A method developed by Pineda et al. (2004) converts each CLC class into the closest USGS class. This method matches the 44 CLC classes with 13 of the 24 USGS classes with a surjective but noninjective function thus leading to a loss of LC information due to the aggregation of several CLC classes into the same USGS class. Alternatively, creating a new complete LUT with values for all 44 CLC classes requires the knowledge or the computation of the associated parameter values (obtained, e.g., from satellite and in situ measurements). Since the computation of these values is well beyond the scope of this study, we have endeavored to test a CLC LUT with 44 classes, based however on existing parameter values.

Using methods similar to Golzio et al. (2021), we used albedo, roughness length, emissivity, thermal inertia, and moisture availability data provided by Pineda et al. (2004) in order to build a 44-class LUT. We consider Pineda et al. (2004)'s values valid for our study because they were estimated for northern Spain areas, which share, in terms of LC and climate, a great similarity with our study area: both are at similar latitudes, in Mediterranean climate, and include a part of a major mountain chain (the Pyrenees versus the Alps). Afterward, once we included the satellite-based data from Pineda et al. (2004), some parameters were still missing, such as LAI and rooting depth. These gaps were filled with the values available from the USGS database. Since USGS has fewer classes than CLC, an USGS class can sometimes supply a value to several CLC classes. For example, LAI values of USGS class "Dryland Cropland and Pastures" will be attributed to the two CLC classes "Nonirrigated arable land" and "Pastures." The final result is a 44-entry LUT, called $CLC_{44}$, which can now be used for WRF simulations. Since data from Pineda et al. (2004) is not as differentiated as the CLC classes, therefore some classes remain clustered. To give an illustration, all the artificial surfaces classes share the same parameter values (11

third level classes in CLC). All these nondifferentiated classes of the 44 CLC options lead to an effective number of 28 unique classes among the total of 44. This table is available in the data statement a link to the data set is provided (de Bode, 2022b). In the future, this LUT might be easily enriched if more data would become available (e.g., parameter values reflecting the diversity of artificial surfaces). Having made available this 44-class LUT, we encourage people applying WRF over Europe to improve it.

#### 3.1.2. Assigning Values to Cells

The physical parameters associated with each of the WRF surface cells are determined during the WRF preprocessing. Since the WRF V3.8, the default method for representing LC class is based on the largest LC fraction present in a cell from the given input LC map. It works similarly to an SLM approach, and it is therefore important to give the model an input map with a resolution as close as possible to the actual WRF resolution, otherwise WRF would apply an SLM aggregation and we would lose the benefit of our MLM method. Over midlatitudes, WRF is usually run on a rectangular horizontal grid on a Lambert Conformal Conic (LCC) projection. This means that even with a similar resolution, a longitude/latitude square grid LC map needs aggregation in order to implement the values in the LCC grid.

As an alternative to the conventional largest fraction method, WRF provides the Mosaic option to account for the subgrid LC variability that the input maps may contain in a cell (Li et al., 2013). It is a form of resolution dependency where physical parameters do not originate directly from one representative value in the LUT but are computed for each cell through a weighted average of the LUT parameters; the weights are computed from the N main LC classes area fractions (N is tunable). The computed value of each cell parameter reflects the main LC fractions within WRF domain cells. To assess the effect of this subgrid option in the WRF model, we will also use this option among the various simulation tests described below.

### 3.2. Description of the Simulations

#### 3.2.1. WRF General Settings

WRF is a nonhydrostatic NWP model, the configuration of which can be adapted through a set of parameterizations available for the different modules, such as boundary layer, convection, microphysics, and radiation. Altogether, our settings follow the previous parameterizations and domain settings established by Kalverla et al. (2016) for simulations of mesoscale flows over our area of interest, except for the radiative schemes which match the more recent RRTMG version. The WRF experiment common tunings are summarized in Table 3.

The simulations are initialized and later constrained at a large scale every hour with ERA5 reanalysis from ECMWF. According to the WRF guidelines, recommending a ratio of 3 between the horizontal resolution of a domain and its parent, the resolutions of the nested domains had thus been chosen by Kalverla et al. (2016) as 27, 9, 3, and 1 km, keeping roughly the same ratio of 3 between reanalysis and coarsest domain resolutions. Since ERA5 offers a grid spacing roughly similar to the coarsest resolution (27 and 30 km), we wondered whether the outermost domain could be omitted in the simulations, starting with the coarsest resolution at 9 km. After some tests and comparisons between the model outputs (not shown here) with 3 and 4 nested domains and our measurements, we were able to conclude that four domains simulations starting with a 27 km parent domain were of higher quality. This conclusion was drawn from metrics computed on the differences between local observations and simulations in a style similar to the one used in the discussion section. The most decisive metrics for this decision are those related to the wind fields, which were better with the 27-km resolution parent domain.

Regarding the finest resolution, we relied on Kalverla et al. (2016) stating that 1 km was fine enough to correctly simulate the flows and their diurnal cycle at the scale of the large topographic systems in the area. Furthermore, our goal is to evaluate the impact of LC on the overall atmospheric structure, and not to resolve very local circulations, as done by de Bode et al. (2021). To investigate the influence of land cover, we outlined the two experiments described below.

#### 3.2.2. Experiment Plan

The first experiment (Experiment #1) was to test the influence of the chosen LC data set (USGS, MODIS, and CLC), the influence of the aggregation method used to build up the CLC data sets and the influence of the CLC LUT. Six runs were performed: $USGS_{24}$, $MODIS_{20}$, $CLC\text{-}SLM_{24}$, $CLC\text{-}MLM_{24}$, $CLC\text{-}SLM_{44}$, and $CLC\text{-}MLM_{44}$.

**Table 3**
*Summary of WRF Settings*

| Setting | Input | | | | Reference |
|---|---|---|---|---|---|
| WRF model version | V4.3.2 | | | | Skamarock et al. (2019) |
| Global data forcing (initial and boundary conditions) | ECMWF ERA5 1 hr time step, 0.25° horizontal resolution, 38 vertical levels | | | | Hersbach et al. (2018a, 2018b) |
| Simulation time 4-day period | 108 hr (starting on 2013-02-18 at 00:00 UTC) | | | | |
| Nudging 4-day period | Spectral | | | | Miguez-Macho et al. (2004) |
| Model top | 50 hPa | | | | |
| Two-way nested domain | D01 | D02 | D03 | D04 | |
| Horizontal resolution | 27 km | 9 km | 3 km | 1 km | |
| Horizontal cell number (W-E × S-N) | 89 × 89 | 106 × 100 | 100 × 100 | 100 × 121 | |
| Time step | 162 s | 54 s | 18 s | 6 s | |
| Output interval | 180 min | 180 min | 60 min | 10 min | |
| Topography maps | 10′ | 5′ | 2′ | 30″ | |
| Vertical levels | 46 | | | | |
| Microphysics | WRF single-moment 6-class scheme | | | | Hong et al. (2006) |
| Planetary boundary-layer scheme | Quasi-Normal Scale Elimination (QNSE) scheme | | | | Sukoriansky et al. (2005) |
| Surface layer scheme | QNSE surface layer unified | | | | Sukoriansky (2008) |
| Land surface option | NOAH land surface model | | | | Tewari et al. (2004) |
| Longwave radiation | RRTMG | | | | Iacono et al. (2008) |
| Shortwave radiation | | | | | |
| Radiation time step | 10 min | | | | |
| Cumulus parameterization | Kain Fritsch | | No parameterization | | Kain (2004) |

*Note*. The last column contains the relevant references.

The subscript indicates the number of available classes in the LUT. USGS and MODIS are thus used in the WRF default configuration, with 24 and 20 classes, respectively. In the label of the CLC runs, "SLM" and "MLM" refer to the corresponding aggregation method. For the latter, two LUTs are used corresponding to 24 (USGS) or 44 classes ($CLC_{44}$) as indicated by the subscript. Table 4 shows the settings in more detail: note that the MODIS data set is only available at 30″ (~1 km), so consequently it does not have adapted resolutions for coarser domains resolution from 3 to 27 km.

The second experiment (Experiment #2) tries to quantify the effects of the different LC input maps resolutions and how subgrid availability influences the skill of the WRF model. We use the CLC-$SLM_{24}$ LC data set for this experiment, since it is a data set used in other comparison studies with CLC (Golzio et al., 2021; Li et al., 2018; Santos-Alamillos et al., 2015; Schicker et al., 2016) and has been shown to be similar or superior to MODIS and USGS. Since we focus here on input map resolution influence, the same aggregation method (SLM) is used for all the runs. The map resolutions tested were 3″ (named "res_111m") and 30″ ("res_1km") and the same map resolution is given whatever the WRF domain. One case ("res_dep") from Experiment #1 is used as a reference; it is using a resolution map similar to the cell size. For these three runs, the SLM method and the largest fraction option were adopted. A fourth simulation was run, identical to "res_111m" except that the Mosaic option was used instead of the largest fraction. In total we compared three different largest factor approaches complemented with one Mosaic approach simulation. Table 4 denotes the exact resolutions used for the different simulations.

The results of the local comparisons in Section 4.2 have no horizontal interpolation but have the values of the cell that encompasses the location. On the vertical, we apply interpolation to the altitude above mean sea level.

**Table 4**
*WRF Runs of Experiments #1 and #2*

| | Case name | Map provider | WRF domains (hor. res.) | Input map resolution | Map aggregation method | LUT | WRF cell representation |
|---|---|---|---|---|---|---|---|
| Land Cover, Experiment #1 | $USGS_{24}$ | USGS | D01 (27 km) | 10′ | — | USGS | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′ | | | |
| | | | D04 (1 km) | 30″ | | | |
| | $MODIS_{20}$ | MODIS | D01 (27 km) | **30″** | — | **MODIS** | Largest fraction |
| | | | D02 (9 km) | | | | |
| | | | D03 (3 km) | | | | |
| | | | D04 (1 km) | | | | |
| | CLC-$MLM_{24}$ | CLC | D01 (27 km) | 10′ | **MLM** | USGS | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′ | | | |
| | | | D04 (1 km) | 30″ | | | |
| | CLC-$SLM_{24}$ | CLC | D01 (27 km) | 10′ | **SLM** | USGS | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′n | | | |
| | | | D04 (1 km) | 30″ | | | |
| | CLC-$SLM_{44}$ | CLC | D01 (27 km) | 10′ | **SLM** | **$CLC_{44}$** | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′ | | | |
| | | | D04 (1 km) | 30″ | | | |
| | CLC-$MLM_{44}$ | CLC | D01 (27 km) | 10′ | **MLM** | **$CLC_{44}$** | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′ | | | |
| | | | D04 (1 km) | 30″ | | | |
| Resolution, Experiment #2 | res_dep same as CLC-$SLM_{24}$ | CLC | D01 (27 km) | 10′ | **SLM** | USGS | Largest fraction |
| | | | D02 (9 km) | 5′ | | | |
| | | | D03 (3 km) | 2′ | | | |
| | | | D04 (1 km) | 30″ | | | |
| | res_1km | CLC | D01 (27 km) | **30″** | **SLM** | USGS | Largest fraction |
| | | | D02 (9 km) | | | | |
| | | | D03 (3 km) | | | | |
| | | | D04 (1 km) | | | | |
| | res_111m | CLC | D01 (27 km) | **3″** | **SLM** | USGS | Largest fraction |
| | | | D02 (9 km) | | | | |
| | | | D03 (3 km) | | | | |
| | | | D04 (1 km) | | | | |
| | res_mos | CLC | D01 (27 km) | **3″** | **SLM** | USGS | **Mosaic on 13 dominant classes** |
| | | | D02 (9 km) | | | | |
| | | | D03 (3 km) | | | | |
| | | | D04 (1 km) | | | | |

*Note*. The cells in bold indicate where the configurations differ from to the $USGS_{24}$ simulation. SLM is the classic Single Level Mode aggregation. MLM is the new Multi-Level Mode aggregation. $CLC_{44}$ is the LUT with 44 entries.

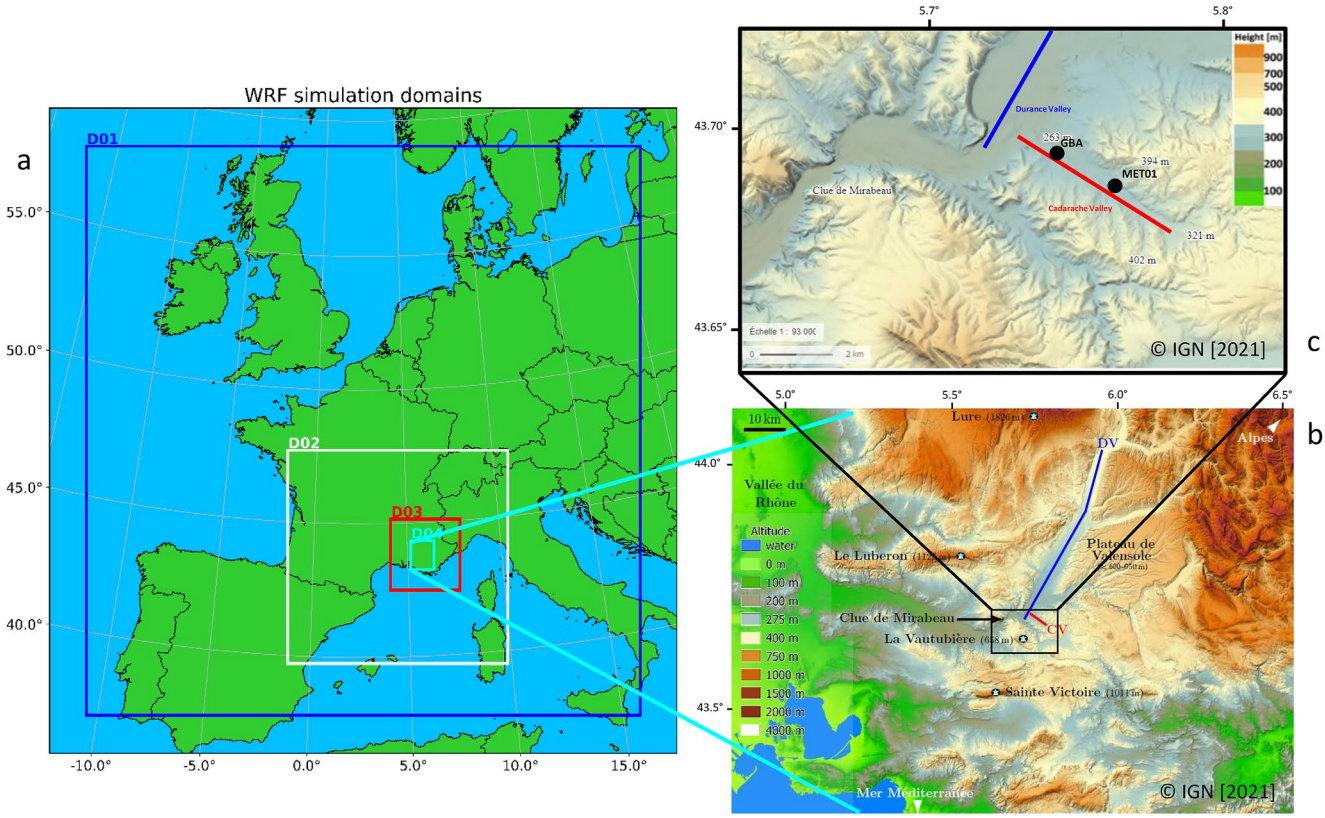


**Figure 5.** Map (a) shows the four nested domains (D01 to D04) used in Weather Research Forecasting (WRF) simulations. Map (b) shows D04 orography, around the pre-Alps, the middle Durance valley (DV; blue line) and the tributary Cadarache valley (CV; red line). Map (c) is a zoom-in around the measurement site of KASCADE-2013 campaign inside CV. M30 30 m mast and GBA 110 m tower positions are 43.6855°N, 5.7617°E and GBA 110 m 43.6944°N, 5.7461°E, respectively.

### 3.3. Case Study

#### 3.3.1. Location

Our case study takes place in the pre-Alps of Southern France (Figure 5a), bordered by the Rhône valley on the west and the Alps to the north and east. In the south and west lie two mountain ridges, culminating around 1 km elevations (Luberon and Sainte-Victoire in Figure 5b). We focus on the middle segment of the Durance valley (DV), separated from its lower and upper parts by narrow paths ("clues" in French: The "Clue de Mirabeau" at the downstream end and the "Clue de Sisteron" at the upstream end). This section of the DV is a relatively straight and wide valley, with several small tributary valleys such as the Cadarache valley (CV; Figure 5c), surrounded by high-raised reliefs preventing synoptic winds from accessing the valley bottom easily. Typical winds in the area are the local "Mistral" (an NW wind which has been channeled down the Rhône valley and then up the DV), SE winds which often come with precipitating perturbations, sea breezes and anabatic winds from the W, and katabatic valley winds from the NE. Most types of wind regimes have previously been realistically reproduced by simulations with a 1-km grid (Duine, 2015; Kalverla et al., 2016).

#### 3.3.2. Observations

Observations come mainly from the "La Grande Bastide" (GBA) site, a permanent 110 m tower that stands within the confluent zone of the DV and the CV (Figure 5c), which means that funneled winds from the CV can influence low-level observations of the tower whereas its top is under the influence of DV winds. As a complement to these installations, a 30 m mast (M30) has been installed in the CV for the duration of the KASCADE-2013 campaign (December 2012 to March 2013; Duine et al., 2017b). M30 focused on observing nocturnal down-valley winds in the CV, a recurring phenomenon in calm-wind, clear-sky periods. M30 measured sensible and latent heat fluxes at 30 m, and the four components of the net radiation (longwave and shortwave, incoming and outgoing) at 2 and 20 m. M30 was also the location for observations with radiosondes on both tethered and free ascending balloons.

A Sodar was placed to the north and outside the CV to continuously measure the wind profile between 75-m and 500-m AGL (above ground level).

#### 3.3.3. Meteorological Conditions of the Selected Case

We select cases from the KASCADE-2013 campaign because we assume that under low synoptic forcing and especially stable conditions the effects of the different LC data sets are more pronounced. We selected a period with several Intensive Observation Period (IOP) from KASCADE-2013, including two IOPs used as reference case for determination of WRF settings by Kalverla et al. (2016). Our case took place in 2013 from noon on 18 February to noon on 22 February, covering four IOPs (15–18). In this period, we use the period of noon 19 February to noon 20 February (IOP16) as example for close-up views. The sky was clear, no fronts were present in the area surrounding the DV, and high-pressure conditions prevailed. At the geopotential height of 500 hPa, a weak pressure gradient existed, with pressures decreasing to the east (not shown). This pressure gradient, though adequately oriented, was too weak to ensure Mistral winds. Accordingly, weather briefings from Météo-France did not report any Mistral wind, which was not detected in our observations either.

Figure 6 shows a 5-day period embracing the 4IOPs. The potential temperature exhibits a diurnal range around 18 K at 2 m and 12 K at 110 m, which are typical ranges for clear-sky, moderate wind events enabling stable stratification to develop at night. The specific humidity increases in the morning because of evaporation, whereas in the afternoon, it diminishes under the drying effect of entrainment at the top of the boundary layer. At night, the decrease continues probably caused by dew deposition (after sunset relative humidity at 2 m reaches 95% and M30 measurements show that specific humidity at 2 m is lower than at 30 m (not shown)). The wind speed at 110-m peaks in the afternoon, coming from the west (pre-Alps anabatic wind favored by weak synoptic forcing) or south, while at night the wind continuously slows down and orientates to the NE around midnight, following the DV downward. The wind direction shows dramatic changes between daytime and nighttime regimes. While the night-to-day change occurs 2 hr after sunrise however, there is a 6-hr delay after sunset for the day-to-night change. This delayed setting results from the fact that DV down-valley wind starts in upslope parts of the DV (up to 80 km upstream) and progressively reaches lower regions (Duine et al., 2017a). The arrival of the NE winds after 00:00 UTC on 20 February coincides with the more intense cooling observed on 110-m potential temperature observations.

Additional observations of these IOPs show that the radiative balance (detailed analyzed in Section 4.2.3) displays near-perfect clear-sky days and nights, with net radiation maxima around 400 W $m^{-2}$ just before 12 UTC. Furthermore, radiosondes were released during each noon-to-noon IOP time period so as to observe the tropospheric vertical structure (Section 4.2.2). The illustration given for IOP16 shows a neutral profile in the atmospheric boundary layer during the daytime (mixed layer), which persists after sunset as a residual layer, whereas after sunset a stable layer begins to grow from the surface upwards (more details will be given about these profiles in Section 4.2.2).

## 4. WRF Simulation Results

### 4.1. Domain-Wide Averages

#### 4.1.1. Domain-Wide Differences Among Static Variables

To evaluate the overall effect of using different LC data sets, aggregation methods, and LUTs, we compare a few surface parameters averaged over the whole 1-km resolution domain ("D04" in Figure 5), an area of 100 × 121 km, encompassing the middle DV, the pre-Alps and a small part of the Mediterranean sea shore around Marseille's conurbation, i.e., comprising a wide variety of LC classes ($USGS_{24}$ uses 13 of 24 possible classes, $MODIS_{20}$ 15 of 20, $CLC\text{-}MLM_{44}$ 30 of 44, $CLC\text{-}SLM_{44}$ 31 of 44, and both $CLC_{24}$ 10 of 24). The results are presented in Table 5, for albedo, emissivity, and roughness length, which are crucial parameters for surface flux estimates. These three parameters have limited variations on short periods of time and thus can be compared independently of the atmospheric simulations.

Differences in albedo are noticeable since $CLC_{44}$ values are roughly 20% lower than the others. Such differences can significantly impact the surface radiation budget during daytime, with direct consequences on surface temperature and heat fluxes, hence on the boundary-layer structure. As $CLC_{24}$ values are close to those of USGS and MODIS, we may conclude that $CLC_{44}$ values highlight the effect of more LC classes. Indeed, LUTs with 20

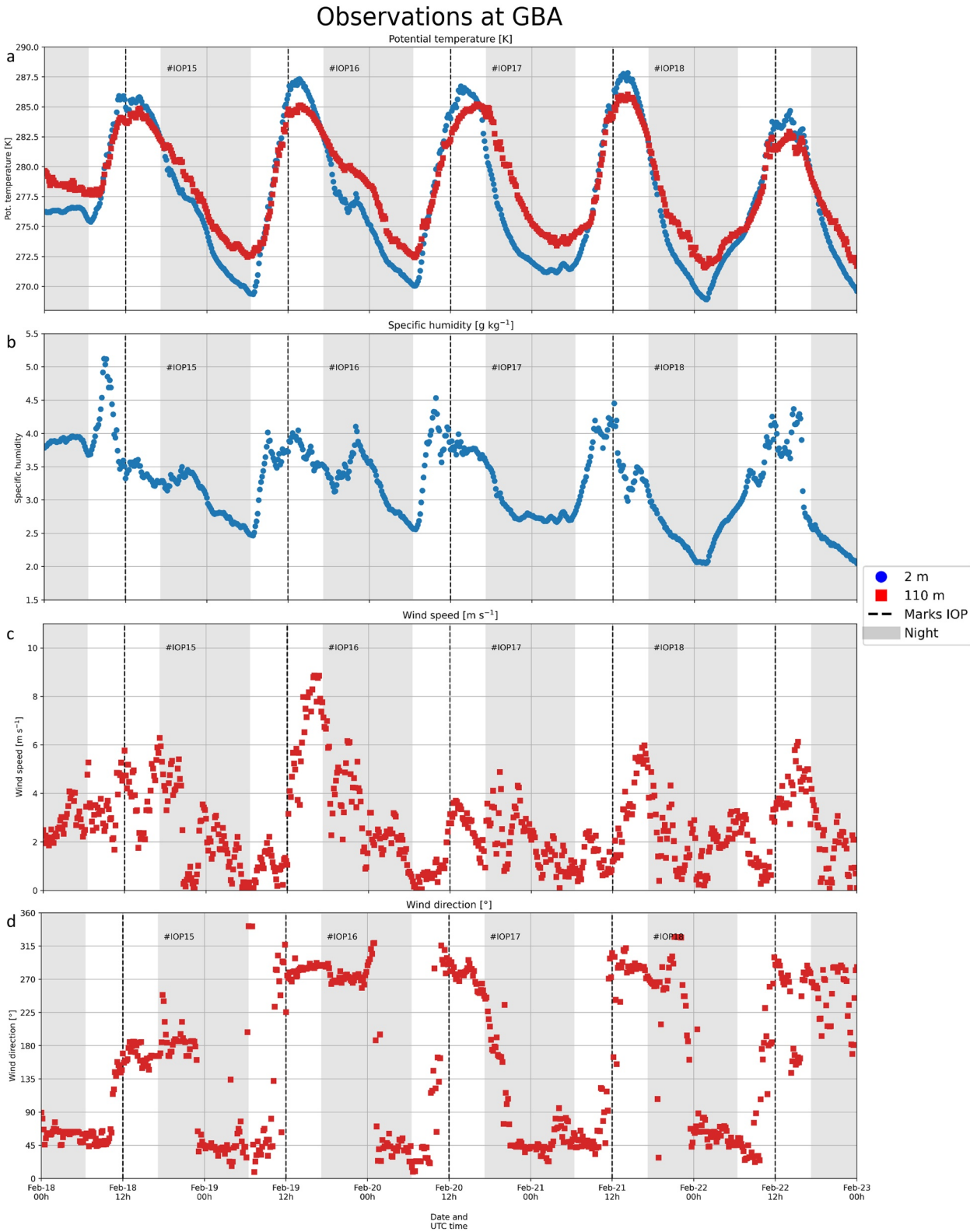


**Figure 6.** Observations at GBA station from 18 February to 23 February 2013. (a) Potential temperature at both 2-m and 110-m heights. (b) Specific humidity at 2 m. (c) Wind speed and (d) direction at 110 m. The nighttime periods are shaded. Each IOP starts at 12:00 UTC and ends 24 hr later.

**Table 5**
*One-Kilometer Resolution Domain-Wide (Over a 100 × 121 km Area) Averaged Values of Albedo, Emissivity, and Roughness Length for All Six Different LCs Corresponding to the Case Study Described in Section 3.3*

| | Case | | | | | |
|---|---|---|---|---|---|---|
| Parameter | $USGS_{24}$ | $MODIS_{20}$ | $CLC\text{-}SLM_{24}$ | $CLC\text{-}MLM_{24}$ | $CLC\text{-}SLM_{44}$ | $CLC\text{-}MLM_{44}$ |
| Albedo (–) | 0.221 | 0.217 | 0.211 | 0.212 | 0.171 | 0.171 |
| Emissivity (–) | 0.925 | 0.924 | 0.928 | 0.929 | 0.964 | 0.964 |
| $z_0$ (cm) | 17 | 29 | 27 | 27 | 32 | 32 |

or 24 classes lead to similar albedo values (∼0.22) whatever the LC source (MODIS, USGS or CLC), whereas 44 classes result in significantly lower albedo (∼0.17).

The same conclusion applies to the average emissivity, which does not deviate much from 0.93 except for $CLC_{44}$, for which it increases to ∼0.96. However, the impact on meteorological parameters is not expected to be great, since the increase in outgoing longwave radiation caused by a higher surface emissivity is in part compensated for by a higher absorption of incoming longwave radiation, and vice versa.

Looking at the roughness length $z_0$, the $USGS_{24}$ gives a small value (17 cm), whereas all other LCs are around 30 cm. We may notice a light influence of using 44 classes, as the values increase from 27 cm for $CLC_{24}$ and up to 32 cm for $CLC_{44}$. Such differences have a direct impact on the wind speed and friction velocity computed in the simulation. Since Table 5 presents values averaged over the domain, it may hide a local variability. Using MAE to measure domain variability for each variable and each case shows that variability is the smallest in the $USGS_{24}$ case (0.02, 0.006, and 11 cm for albedo, emissivity, and $z_0$, respectively), while the other cases reveal a variability that increases to 0.05, 0.014, and 19 cm, respectively.

#### 4.1.2. Domain-Wide Averages of Simulated Variables

In the rest of this paper, we will focus on the 4-day simulations with WRF v4.3.2, this 4-day period (from IOP15 to 18) allows us to have robust statistics of the simulation performances compared to the observations. As illustrated above, the weather conditions and trends among each IOP appear similar.

Here, a domain-wide comparison between LC data sets was undertaken for the 2-m potential temperature ($\theta_{2m}$) and the 10-m wind speed (WS10) to evaluate the LC data sets influence over a large area, as these parameters give a first idea of how different LCs affect the meteorological variables near the ground. Verifications with observation at local stations will follow in Section 4.2. The studied domain is the innermost one (D04, the same as in Section 4.1.1) with 1-km horizontal resolution (Figure 5b). For the averages, we excluded the cells in the five rows and columns on the outer edges to avoid the influence of D03.

The results are presented in Table 6. Detailed information are given by splitting out the spatial and temporal components of variations, looking at their minimum and maximum. We determine the temporal variation by looking at the time series of the domain-wide average and we determine the spatial part by looking at the domain after time averaging every cell. The table contains the whole 4-day simulation with WRF v4.3.2 because the QNSE parameterization before v4.2 incorrectly recalculates the WS10 values (Cao & Fovell, 2016). The v4.3.2 run thus allows us to avoid the influence of this error.

$\theta_{2m}$ presents little difference between data sets, which is striking because this result seems in contradiction with the different albedo and emissivity associated with CLC at least when 44 classes are considered (Table 5). Physical processes related to soil composition (in particular soil water evaporation) are probably of larger importance in controlling the temperature close to the ground than the restricted modification of the radiation balance through an albedo variation of 0.05. The variability in time and space is comparable between the data sets.

Regarding wind speed, $USGS_{24}$ has the highest value, and $MODIS_{20}$ is ∼0.3 m $s^{-1}$ slower with a comparable decrease in value range. The $CLC_{44}$ winds are slower on average, but with a similar temporal variability to MODIS, which is valid for all CLCs, irrespective of the aggregation method or the LUT used. As expected, the higher average winds are generally related to lower roughness lengths and vice versa (see Table 5): a $z_0$ value of 17 cm corresponds to a WS10 of 3.3 m $s^{-1}$, whereas a $z_0$ value of 32 cm corresponds to 2.8 m $s^{-1}$.

**Table 6**
*Mean and Range of Values Averaged Over Domain 04 for the 2-m Potential Temperature ($\theta_{2m}$) and 10-m Wind Speed (WS10)*

| Quantity | Case | | | | | |
|---|---|---|---|---|---|---|
| | $USGS_{24}$ | $MODIS_{20}$ | $CLC\text{-}SLM_{24}$ | $CLC\text{-}MLM_{24}$ | $CLC\text{-}SLM_{44}$ | $CLC\text{-}MLM_{44}$ |
| 4-day simulation WRF v4.3.2 | | | | | | |
| $\theta_2$ (K) temporal min, max | 280.6 | 280.7 | 280.6 | 280.6 | 280.5 | 280.5 |
| spatial min, max | 277.5, 286.0 | 277.6, 286.2 | 277.5, 286.2 | 277.4, 286.2 | 277.2, 286.3 | 277.2, 286.3 |
| | 276.3, 286.7 | 276.9, 286.8 | 276.1, 286.8 | 276.0, 286.8 | 276.7, 286.7 | 276.0, 286.7 |
| WS10 (m $s^{-1}$) temporal min, | 3.3 | 3.0 | 2.9 | 2.9 | 2.8 | 2.8 |
| max spatial min, max | 1.5, 4.7 | 1.4, 4.1 | 1.4, 4.1 | 1.4, 4.1 | 1.3, 3.9 | 1.3, 3.9 |
| | 1.0, 8.3 | 0.8, 8.2 | 0.9, 8.3 | 0.9, 8.7 | 0.8, 7.6 | 0.9, 7.6 |

*Note*. In each table cell, the top value is the average over both the whole D04 domain (121 × 100 values) and the 4-day period; temporal min (max) is the 4-day period minimum (maximum) from the domain-averaged values; spatial min (max) is the domain minimum (maximum) determined among the 4-day averaged cells values.

We checked the significance of the differences between the data sets through a multivariable analysis of variance (MANOVA; Bock & Haggard, 1968; Tabachnick & Fidell, 2011) and a post-hoc analysis. The input data are the temporal series of the domain-wide average including both $\theta_{2m}$ and WS10. The MANOVA showed that some significant differences existed within the simulations with a Pillai's trace $p$-value (Pr > $F$; with Pr the Pillai's trace and $F$ the $F$-statistic) of $<4 \times 10^{-4}$. Furthermore, for each pair of cases presented in Table 6, we computed the values of two-sided two-sample $t$-tests with a sample size of 577. Given the six LC cases, this resulted in 15 $t$-test $p$-values in both potential temperature and wind speed. Using a confidence interval of 95%, these analyses showed no significant differences in the potential temperature fields, with the lowest $p$-value at 0.290. In the wind field, the all runs were significantly different from other data sets with a $p$-value of $<4 \times 10^{-4}$.

### 4.2. Local Evaluation Against Observations

We will examine in the following subsections the LC effects on some variables at two localized points in the CV, namely the GBA and M30 stations (see Figures 5 and 7). We extracted from the simulation outputs the values at the cells which encompass the observation sites. Additionally, we linearly interpolated the model output on

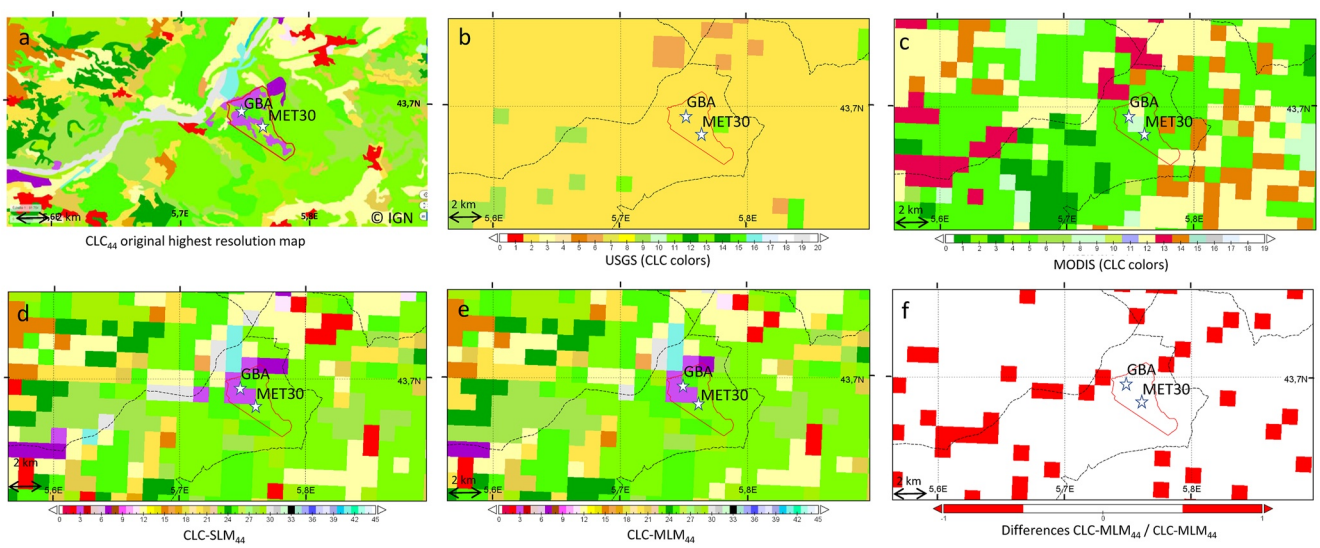


**Figure 7.** Land cover (LC) maps around the observation site (the site is delimited with a red line); the map frame resembles Figure 5c. GBA and M30 are the measurements stations used for the case study. (a) shows as a reference the original $CLC_{44}$ highest resolution map. (b)–(e) are the 1-km resolution maps as used in the WRF domain D04. Map (b) is the LC from USGS data set. (c) is from MODIS data set. (d) is the map from CLC SLM aggregation with 44 classes. (e) is as (d) but from MLM aggregation. CLC standard colors (https://land.copernicus.eu/Corinelandcoverclasses.eps.75dpi.png/) are used for all LCs. (f) shows the difference (in red) between the MLM and SLM aggregations.

the vertical to match 110-m AGL (or 375 m AMSL; above mean sea level) the height of GBA observations. The closest model levels are at 370 m (Z1) and 395 m (Z2) AMSL, thus 0.8 and 0.2 are weights given to Z1 and Z2 levels, respectively.

#### 4.2.1. Local Land Cover Maps

The LC around the case study site is quite heterogeneous, as seen from the full resolution CLC map, with 16 classes in the area (see Figure 7a): it is a mix of seven forest and seminatural covers (green areas), four agricultural (yellow to brown), and four urban (purple and red) with some water surfaces (blue). The USGS map (Figure 7b) is quite homogeneous and far from terrain reality, as it mainly shows uniform dry land and crop pasture (brown) despite a theoretical resolution of 1 km. MODIS map (Figure 7c) is closer to terrain reality showing a mix of forest, agricultural and urban areas with two noticeable mistakes: the water surface does not appear though its surface is ∼2 $km^2$ (i.e., 2 MODIS 1 $km^2$ pixels) and urban areas have a ∼2 km shift to the North. The CLC maps (Figures 7d and 7e) even at 1-km resolution show a high variety of LC with both having 15 classes present among the 16 seen on the high-resolution map (Figure 7a). When we compare the CLC maps obtained through SLM and MLM aggregation methods (Figure 7f) we count roughly 15% of cells with a different class. It is worth noticing that inside the observation site area (red line) there is no difference between the SLM and MLM maps at 1 km. We might highlight that the GBA and M30 stations are located at the limit between industrial and forested cells.

#### 4.2.2. Radiosondes

The comparison of the radiosonde profiles on IOP16 with the different simulation cases (Figure 8) shows that the influence of the LC maps on WRF simulations is weak and restricted to the boundary layer. We show the IOP16 case because it is representative for general clear-sky daily cycle, the other three IOPs can be found in the Supporting Information S1. The effect is more visible on the specific humidity plots. The observations show a boundary-layer depth up to 1,400 m during the daytime convective conditions (12 UTC plots). At 18 UTC, the cooling has begun and has created a thin stable layer below the residual layer, whereas at 06 UTC, just before sunrise, the vertical profile has become fully stable with a few sublayers close to neutral. The simulations show similar patterns for the other days. At noon, the model is roughly 1 K too cool within the boundary layer and ∼3 K too cool above. At 18 UTC, the model and the observations nearly overlap, except for the fossil capping inversion, where the jump is smoothed in the simulations. The night cooling near the surface is underpredicted as seen at 06 UTC.

The observed humidity shows a very clear mixed layer profile up to 1,400-m AGL at 12 and 18 UTC, whereas the 06 UTC profile shows a peak around 250-m AGL. This development is well simulated although the model seems to dry faster than the observations.

IOP15 and IOP17 do resemble IOP16 in the potential temperature and humidity structure, while IOP18 differs in the humidity structure. The humidity of all simulations is larger than that of observations, with all simulations close together. The difference is roughly 1 g $kg^{-1}$ at 300–1,000-m AGL at 12 UTC and becomes smaller toward 06 UTC. The high humidity and too low potential temperatures in simulations are likely the reason for a snowfall present in all simulations but that was not observed. Seeing that this paper focuses on the land cover effects and not the snow effects we stop further analyses at 22 February 00:00 instead of 12 hr later, thus the below statistics originate from a 3.5-day period.

#### 4.2.3. Radiation

Figure 9 shows the radiation balance components observed at 20-m AGL on M30 and the simulations with the six different LC maps and four different input settings (nine simulations in total).

For downward shortwave radiation (SWd; Figure 9a), all simulations show similar clear-sky days; the values are fairly identical for all the simulations and coincide well with the observations, except the evening of IOP15 and morning of IOP18. Deviations of daily maxima of 19–21 February are always below 25 W $m^{-2}$, and 3.5-day biases range from 13 to 14 W $m^{-2}$. In cloud-free conditions, this resemblance is not surprising because all the simulations use the same radiation transfer schemes and there is no direct influence of the LC on this quantity.

On the other hand, upward shortwave radiation (SWu; Figure 9b) is directly dependent on LC through albedo. The observed value during the IOP stands at 0.17 (close to the value of 0.16 computed from the whole KASCADE-2013 period, not shown), whereas the values at the corresponding cells are 0.23, 0.25, 0.17, and

10.1029/2021EA002123



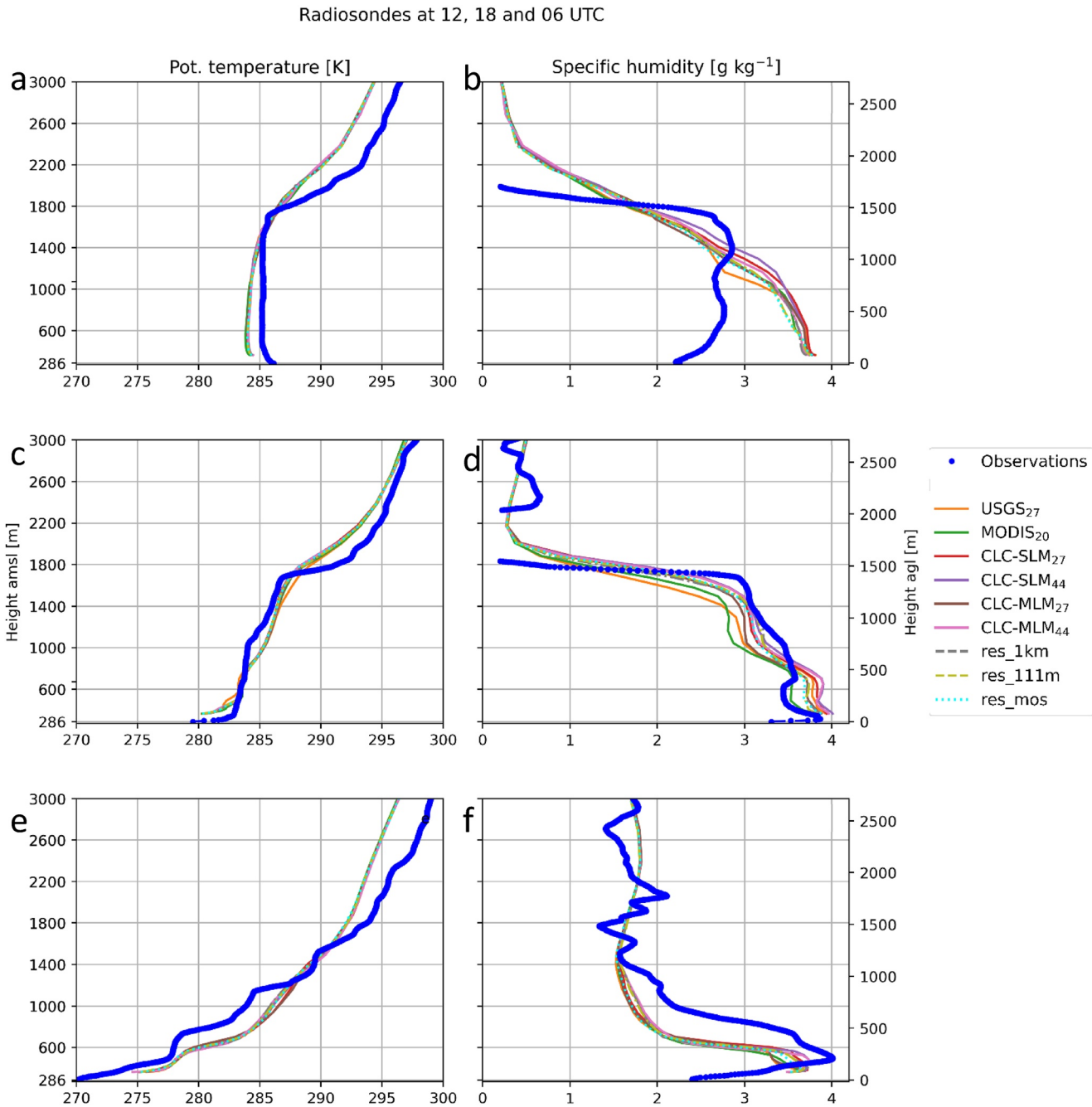


**Figure 8.** Radiosondes and simulation results at M30 location during IOP16 at 12 UTC (a and b), 18 UTC (c and d), and 06 UTC (e and f) with potential temperature on the left and specific humidity on the right. Blue dots are observations, lines are the WRF simulations cases. 286 m is the actual terrain height above sea level, WRF ground level is slightly above due to discretization.

0.16 for USGS$_{24}$, MODIS$_{20}$, CLC$_{24}$, and CLC$_{44}$, respectively. In IOP18 (Figure 9b), we see clear deviations from observations in the morning, which results from a snow layer present in all simulations, whereas this snow is not observed. As mentioned before, we will not include 22 February 2022 00:00 to 12:00 in statistics, because these parts would show the effect of the snow and not the LC effects themselves. In USGS$_{24}$ and MODIS$_{20}$, SWu values at noon are higher than observations by +40 and +52 W m$^{-2}$, respectively and consistently in 19–20 February. This indicates an excessive reflection at this location. In contrast, CLC$_{24}$ and CLC$_{44}$ simulations are closer to the observed values, with errors of 2 and −4 W m$^{-2}$ of the noon maximum for CLC$_{24}$ and CLC$_{44}$ respectively. 3.5-day biases range from −0.5 to 1.1 W m$^{-2}$ in the CLC runs, whereas the bias reaches 11 and 15 W m$^{-2}$ for USGS$_{24}$ and

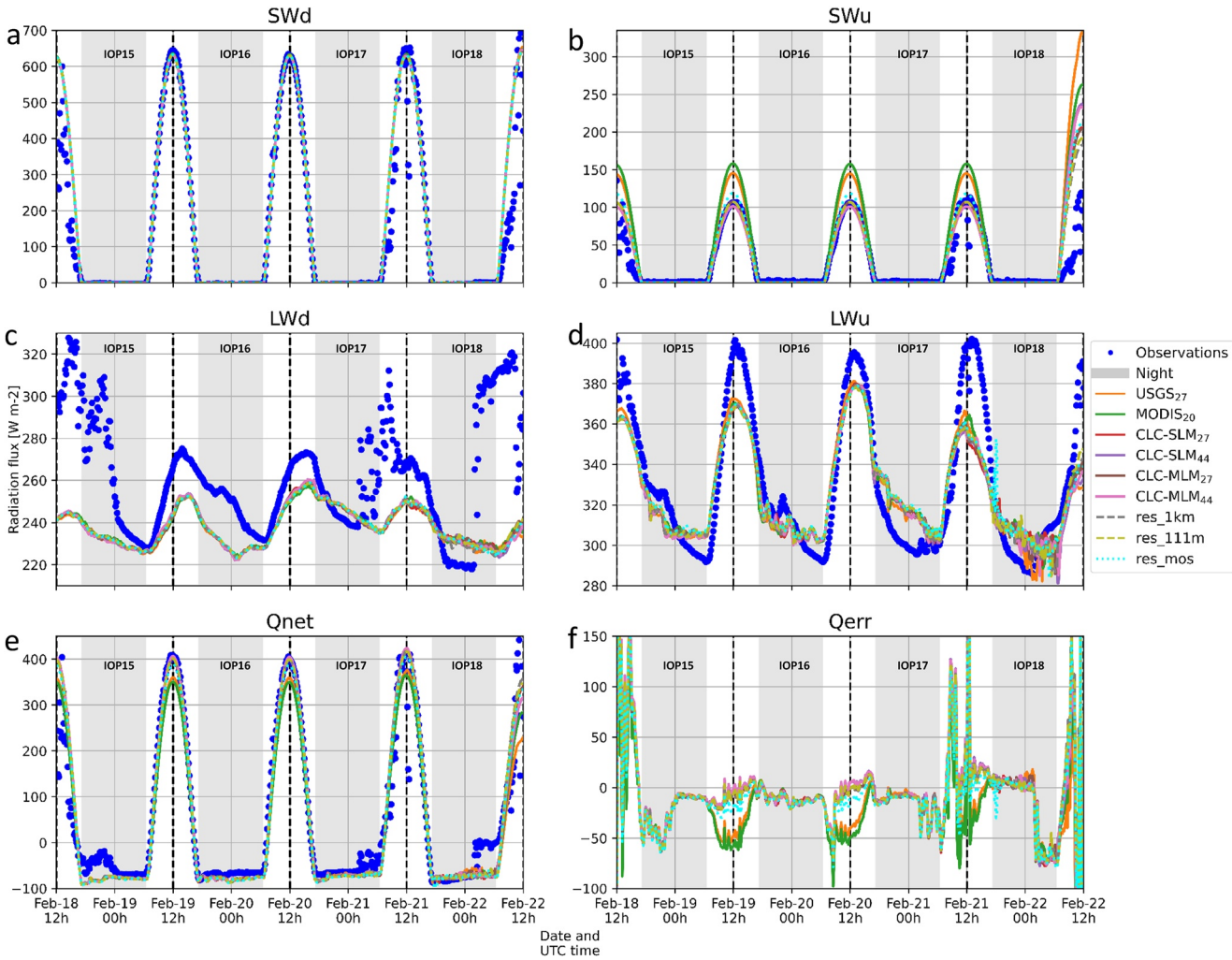


**Figure 9.** Modeled radiation components (lines) and 20-m observation at M30 (blue dots). SWd stands for shortwave downward; SWu, shortwave upward; LWd, longwave downward; LWu, longwave upward; Qnet is the net radiation and Qerr is the difference of Qnet between the simulations and the observations. All quantities are in W $m^{-2}$. Shaded areas are nighttime periods. Dotted lines indicate the split between different IOPs, starting with IOP15 going to IOP18.

$MODIS_{20}$. The differences resulting from the aggregation method (SLM versus MLM) are not significant. This is due to the fact that the LC class of the cell at the M30 location is that of the closest grid cell, which was found to be the same CLC class ("Broad-leaved forest," index **3.1.1**, albedo = 0.16; see Figure 7) for either MLM or SLM as can be seen in Figures 7d and 7e. Since upward shortwave radiation depends only on local characteristics, the values are thus identical for the two methods.

Downward longwave radiation (LWd; Figure 9c) is underpredicted regardless of the simulation with maximum deficits of 26–27 W $m^{-2}$ in the clear sky afternoons of 19 and 20 February, while 18 and 21 have deficits of 82–84 W $m^{-2}$ related to the partial cloudiness that was observed but not simulated. This underestimation is not directly related to LC representation but could reflect a too dry and/or too cold overlying atmosphere, although radiosondes (Figure 8 and Figures S2–S4 in Supporting Information S1) at 12, 18, and 06 UTC do not indicate differences between model and observation in the first three kilometers that might explain a 25 W $m^{-2}$ error. This underestimation of LWd is a known issue with WRF (Cerenzia, 2017; Kleczek et al., 2014; Steeneveld & de Bode, 2018; Sterk et al., 2013; Velde et al., 2010). Nonetheless, the model improves at dawn, and the minima of LWd coincide well with the observation, both in timing and magnitude, and the resulting 3.5-day average biases range from −18 to −19 W $m^{-2}$, with $CLC\text{-}SLM_{24}$ performing better. Even though this improvement at dawn is intriguing we shall not investigate it further, since it is not directly related to the LC representation.

For upward longwave radiation (LWu; Figure 9d), directly related to the ground skin temperature ($T_{SK}$), all the simulations seem to be close to one another. The 3.5-day average biases are quite low with values not exceeding $-3.5$ W m$^{-2}$. However, this good bias performance results from a compensation of the nighttime overestimation having maximum values of (20–21 W m$^{-2}$ in the night of 20 unto 21 February) by the daytime underestimation (47–54 W m$^{-2}$ 21 February and 29–33 W m$^{-2}$ on 19 and 20 February). Thus, all the simulations underestimate the daytime peak value, and overestimate the temperature of the cooling surface in the second half of the night. Such a diurnal course has direct feedback on air temperature, as we will see later on. Furthermore, the simulated time series show fluctuating behavior during the night, the origin of which is not yet determined but could lie in transient advection of air parcels with different thermodynamic characteristics.

Combining all four radiative fluxes (SWd − SWu + LWd − LWu; Figure 9e) gives us the net radiation (Qnet). In short, all the model simulations using the CLC maps are close to observations. MODIS$_{20}$ and USGS$_{24}$ runs have biases of $-15$ and $-12$ W m$^{-2}$, respectively, mainly since they use a too high albedo value. In comparison, the CLC runs have biases from $-2$ to 0 W m$^{-2}$. CLC-SLM$_{44}$ performs the best with a bias of $-0.2$ W m$^{-2}$. Evaluating the net radiation error over time (Qerr; Figure 9f) shows that the net radiation is generally underestimated, especially in early morning, while in the afternoon CLC provides the best results with errors around zero. At night, the simulated net radiation is always too low. The MODIS$_{20}$ and USGS$_{24}$ have large underestimations during daytime, originating from overestimated albedo values, while at night they resemble the other runs. An error peak appears slightly before 9 UTC on 21 February: At least part of this is due to an artifact in the observations related to a reflection on the mast, as revealed by corresponding SWd and SWu values that deviate from the overall trend. Other peaks, such as those in IOP15 and IOP18 are caused by observed clouds absent in simulation.

For Experiment #2, no large differences exist between the res_dep, res_1km, and res_111m simulations. Their biases range within 0.5 W m$^{-2}$ of each other. The last simulation of res_mos (Mosaic approach) has larger differences with net radiation bias of $-6$ W m$^{-2}$ compared to the $-2$ W m$^{-2}$ of the other three simulations. This difference primarily results from a corresponding increase in the outgoing shortwave radiation.

Net radiation is a primary driver of the surface energy balance, where it is partitioned into soil het flux, sensible heat flux ($H$) and latent heat flux (LE). Although an attempt was made to measure $H$ and LE during the KASCADE-2013 campaign, the recorded data were not of sufficient quality to ensure comparisons with simulations. As an illustration, the calculated value of $H$ + LE at noon was less than half of the net radiation, which is not realistic. It was therefore impossible to perform reliable comparisons between observed and simulated turbulent fluxes.

#### 4.2.4. Potential Temperature

The finest resolution of our simulation is 1 km, which cannot adequately simulate the winds channeled by the small (1–2-km wide) tributary CV (see Figure 5). We therefore switch to observations at the GBA tower (110-m AGL), which are better suited for comparisons (Kalverla et al., 2016), as 100 m is the height of the surrounding ridges of the CV with respect to its floor and the tower top is therefore above the flow channeled by this valley and is under the influence of the DV wind. As mentioned above the simulated values are vertically interpolated to the 110-m AGL from the two closest model levels.

The time series of the observed and simulated temperatures and the corresponding statistics are presented in Figure 10 and Table 7, respectively. The daily course of the 2-m potential temperature ($\theta_{2m}$) has a larger amplitude than that of the 110 m ($\theta_{110m}$). In WRF simulations, the potential temperature at 2 m ($\theta_{2m\_diag}$) is computed from the diagnostic temperature at 2 m ($T_{2m\_diag}$), which in turn is dependent on LC parameters such as skin temperature, roughness length, and friction velocity (see WRF file module_surface_driver.F).

$$T_{2m\_diag} = T_{skin} - \frac{H}{\frac{P_{surf}}{R_d * T_{skin}} * C_p * CHS_2}$$

where $T_{skin}$ is the surface temperature in K, $H$ is the upward sensible heat flux at the surface in W m$^{-2}$, $P_{surf}$ is the surface pressure in Pa, $R_d$ (287.06 J kg$^{-1}$ K$^{-1}$) is the gas constant for dry air, $C_p$ (1,004 J kg$^{-1}$ K$^{-1}$) is the heat capacity of dry air at constant pressure, and $CHS_2$ is a coefficient for the heat exchange formula at 2 m in m s$^{-1}$ depending on both the roughness length and the friction velocity.

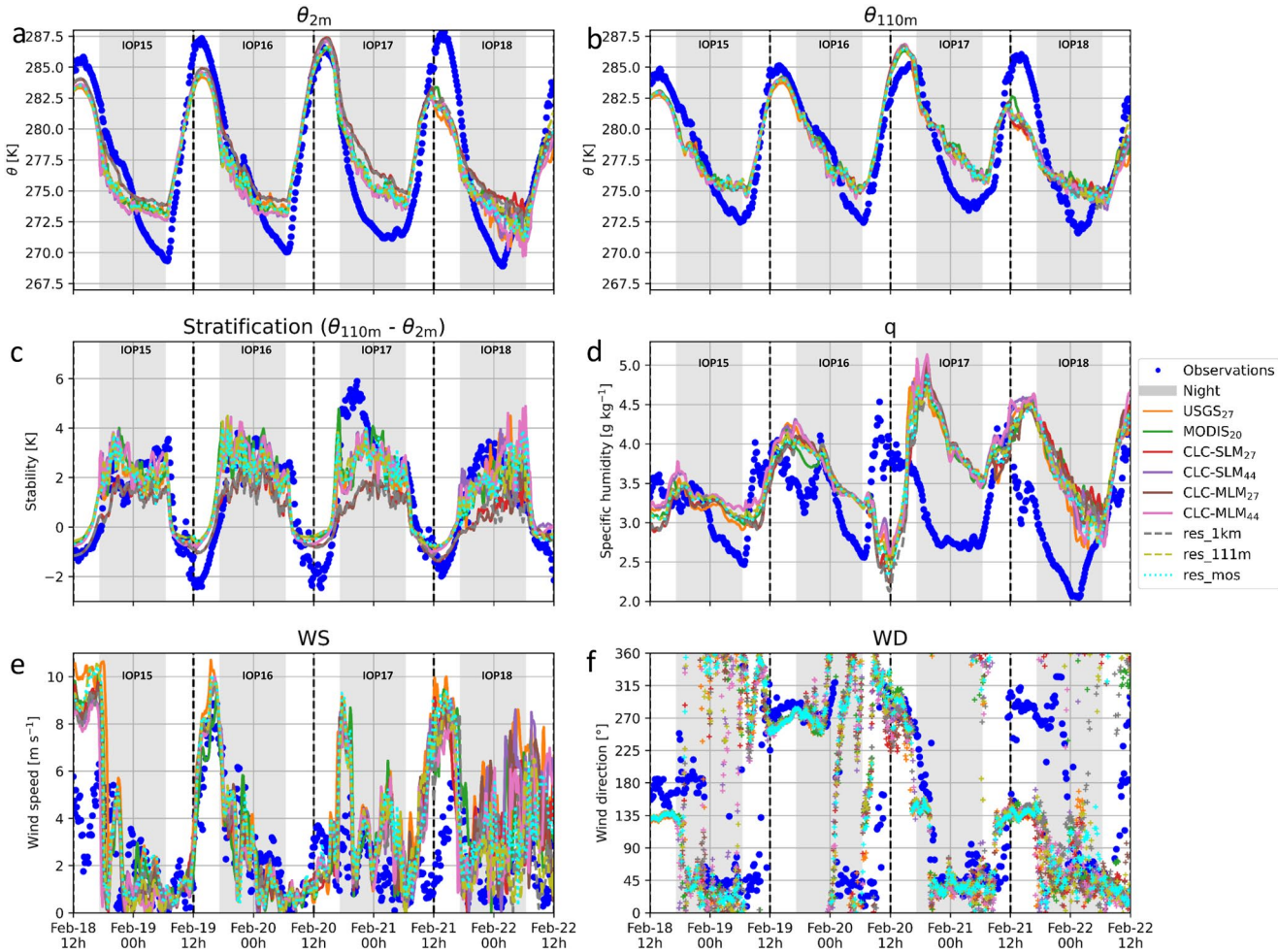


**Figure 10.** Time series at GBA of potential temperature at 2 m ($\theta_{2m}$, top left), 110 m ($\theta_{110m}$, top right) and the difference between these two levels ($\theta_{110m} - \theta_{2m}$, midleft), specific humidity at 2 m (q, midright), wind speed (WS, bottom left), and wind direction (WD, bottom right) at 110-m AGL. The simulations are the lines and the small dots for WD and the observations are the large blue dots, as in Figure 9. Shaded areas are nighttime periods.

Overall, 3.5-day biases of 2-m potential temperature range from −0.04 to 1.0 K (Table 7), with $CLC_{44}$ simulations performing better. During the daytime, however, the simulated potential temperatures are generally lower than the observations, whereas soon after midnight they become higher. After sunrise, the simulations underpredict the observed potential temperature rise at 2 m. At nighttime, the simulations display two different regimes, the smooth cooling of $CLC_{24}$ and res_1km against the somewhat quick fluctuant behaviors of the other simulations. We think that fluctuant regimes reflect more diverse temperatures in air parcels passing causing the fluctuations in potential temperature. Further, a daily temperature range (DTR) of 16.7, 17.3, 15.6, and 18.9 K is observed at 2 m for IOP15 to 18, respectively. While simulations of IOP15 and 16 underestimate the DTR by 5.0–6.5 K, IOP18 has underestimations of 6.6–7.0 K and CLC-$SLM_{24}$ and res_1km even have an 9.2 K underestimation, and the total simulated DTRs range from 9.7 to 12.4 K during these IOPs. These three IOPs have errors of the same order that was already mentioned by Kalverla et al. (2016). This DTR error consists of a 2–4 K underestimation of the daytime peak and an overestimation of 1–5 K of the nighttime low. Errors in DTR are a general problem in many models and not only WRF (Lindvall & Svensson, 2015; Wyszogrodzki et al., 2013). Kalverla et al. (2016) attribute the DTR to unresolved orography, but in the fine-resolution simulation of de Bode et al. (2021) in spite of the added two domains that refined the horizontal resolution to 111 m, the expected improvement in DTR simulated values was not observed. To fix this problem de Bode et al. (2021) suggest to explore the effect of the moisture in the soil, which could be overestimated in the simulations. More generally, in wintertime, temperature forecasting might have larger errors (Duan et al., 2018), because freezing processes add to the challenge of proper forecasts. In summary, improving LC representation is not the main lever for reducing the DTR error.

**Table 7**
*Bias and RMSE for 2- Potential Temperature ($\theta_{2m}$), 110-m Potential Temperature ($\theta_{110m}$), Stratification ($\Delta\theta$), and Specific Humidity ($q$) for Simulations Against Observations at the GBA Tower Site*

| | Variable | $\theta_{2m}$ | | $\theta_{110m}$ | | $\Delta\theta$ | | $q$ | |
|---|---|---|---|---|---|---|---|---|---|
| | Metric Case | Bias (K) | RMSE (K) | Bias (K) | RMSE (K) | Bias (K) | RMSE (K) | Bias (g kg$^{-1}$) | RMSE (g kg$^{-1}$) |
| Experiment #1 | $USGS_{24}$ | 0.13 | 2.8 | **0.08** | 2.2 | −0.04 | **1.1** | 0.4 | **0.8** |
| | $MODIS_{20}$ | 0.20 | **2.7** | 0.35 | **2.1** | 0.16 | 1.2 | 0.4 | **0.8** |
| | $CLC\text{-}MLM_{24}$ | 1.0 | 2.9 | 0.19 | **2.1** | −0.8 | 1.5 | 0.4 | 0.9 |
| | $CLC\text{-}SLM_{24}$ | 1.0 | 2.9 | 0.21 | 2.2 | −0.8 | 1.5 | 0.4 | 0.9 |
| | $CLC\text{-}MLM_{44}$ | **0.04** | 2.6 | 0.13 | **2.1** | 0.09 | 1.2 | 0.5 | **0.8** |
| | $CLC\text{-}SLM_{44}$ | **0.04** | **2.7** | 0.13 | **2.1** | 0.10 | 1.2 | 0.5 | **0.8** |
| Experiment #2 | res_dep $CLC\text{-}SLM_{24}$ | 1.0 | 2.9 | 0.21 | **2.1** | −0.8 | 1.5 | 0.4 | 0.9 |
| | res_1km | 1.0 | 2.9 | 0.17 | **2.1** | −0.8 | 1.5 | **0.3** | 0.9 |
| | res_111m | 0.1 | **2.7** | 0.20 | **2.1** | 0.07 | 1.2 | 0.4 | **0.8** |
| | res_mos | 0.2 | **2.7** | 0.20 | **2.1** | **−0.01** | 1.2 | 0.4 | **0.8** |

*Note*. Statistics are calculated from 18 February 2013 12:00 to 22 February 2013 00:00. Best performing cases are bold.

On the contrary, IOP17 shows exceptional low errors with 1.9–3.0 K underestimations of simulated DTRs, which range between 12.5 and 13.7 K, with $CLC_{44}$ performing best. This IOP could serve as a lead for investigations for improvement of the DTR.

Concerning $\theta_{110m}$, Figure 10b shows that all simulations perform well during the morning warming and around a similar peak in potential temperature. In the afternoon of 21 February when heating stops, a nonobserved cooling starts in all simulations. In comparison, the cooling at night is generally not continued long enough or a cooling system is missed. For example, around 20 February 01:00 UTC, the simulations miss a cooling that coincides with a change of wind direction visible in Figure 10f. The simulations thus miss at that time the influence of the DV winds. This valley wind is buoyantly forced with an NNE origin (see Figure 5b) thus a source of cold air. Consequently, the simulations remain too warm thereafter. The metrics of the differences between simulations and observations (Table 7) show that $USGS_{24}$ performs better, with a bias of 0.1 K and while all runs have comparable RMSEs of 2.1–2.2 K. Considering the aggregation methods, they result in a minimal difference with a benefit (0.01–0.06 K) for the MLM runs. This MLM benefit is thus slightly larger at $\theta_{110m}$ than at $\theta_{2m}$ (0.00–0.01 K).

Regarding stratification, Figure 10c shows the potential temperature difference between 110 and 2 m ($\Delta\theta$). $USGS_{24}$ run has best bias at 0.0 K and RMSE of 1.4 K, while $CLC_{24}$ has a bias of −0.8 K and RMSE of 1.5 K. $MODIS_{20}$ and $CLC_{44}$ have biases ranging from 0.1 to 0.2 K and RMSE of 1.2 K. In fact, the biases of $\theta_{2m}$ and $\theta_{110m}$ combine to form the biases of stratification, therefore, it is not unexpected that $USGS_{24}$ runs given their low bias at both levels.

Experiment #2 shows a variability in simulation skill comparable to Experiment #1 for the temperature and stratification metrics, indicating that if the LC source and its aggregation are crucial, the way in which it is presented to WRF is just as important. Res_dep and res_1km are relatively close in skill, and in the same way the res_111m and res_mos resemble each other. Therefore, we cannot extract the best configuration for Experiment #2.

#### 4.2.5. Humidity

Figure 10d shows that all simulations resemble each other in humidity development, with generally an overestimation except in the mornings of IOP15–IOP17. MODIS and USGS perform best, but there is no serious difference on specific humidity prediction as shown in Table 7 either through Experiment #1 or #2: the bias and RMSE remain for all cases in the range 0.3–0.5 and 0.8–0.9 g kg$^{-1}$, respectively.

Most noteworthy is the sudden jump in humidity seen 20 February 2013 16:00 UTC, as it resembles a moist air flow system described in de Bode et al. (2021) with most winds coming from the south east, that looks like winds matching the CV orientation, while the 110-m AGL GBA tower measures above the CV sidewall tops. In the morning, the humidity quantities match again. de Bode et al. (2021) shows that this is due to a lack of proper

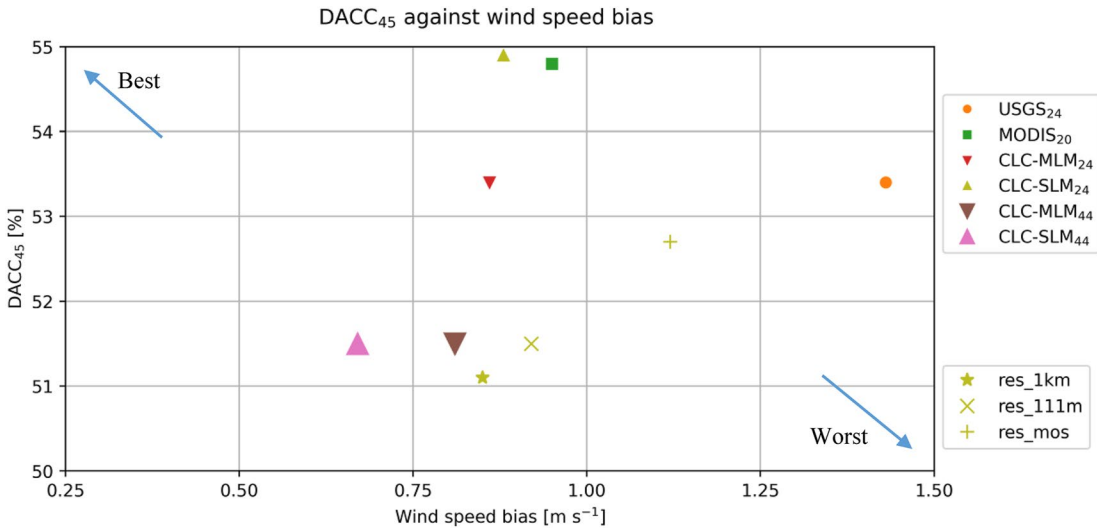


**Figure 11.** Directional accuracy ($DACC_{45}$) as a function of wind speed bias. The colors and shape indicate what type of land cover (LC) data set is used, large symbols have 44 classes while smaller symbols use 20 or 24 classes. Experiment #2 runs are indicated with bronze color symbols.

stability development in the valley resulting in the moist winds from the south east reaching the surface whereas in the observations they remain above valley. However, no radiosonde profiles are available during the night to check the correctness of a moist air layer.

#### 4.2.6. Wind

Figures 10e and 10f show the wind speed and wind direction at 110-m AGL. In general, the simulations manage to reproduce the moderate winds and have the typical daytime direction. The simulations of wind speeds generally perform relatively well, although the peaks are too large, with the exception of IOP18 when all simulations start to diverge. The error in wind speed is characterized by the bias and RMSE, whereas the method to determine the model wind direction skill is the directional accuracy ($DACC_X$; Santos-Alamillos et al., 2013). $DACC_X$ is the percentage of forecasted wind within a specified angular deviation $X$ (in degrees) from the observations

$$DACC_X = \frac{\sum_N \begin{matrix} 1(\text{if } \Delta\varnothing \le X) \\ 0\,(\text{else}) \end{matrix}}{N} x\,100$$

$$\Delta\varnothing = \min[\,\text{abs}(WD_{obs} - WD_{mod}), 360 - \text{abs}(WD_{obs} - WD_{mod})\,]$$

where $WD_{obs}$ is the observed direction and $X$ is the threshold of the acceptable error in simulated wind direction ($WD_{mod}$). All the angles are in degrees, in the range [0–360]. For numerical calculations with angles outside that range, $\Delta\varnothing$ should be calculated with a trigonometry approach. Figure 11 shows the plot of $DACC_{45}$ (i.e., $X = 45°$) against wind speed bias for all the runs. Simulations nearer to the top left corner perform the best in both wind speed and direction metrics whereas runs closer to the right bottom corner perform the worst in both metrics. There is no clear best compromise. The influence of LC data set on wind speed in WRF is more profound than that of resolution settings (Experiment #2). Variation due to LC ranges from 0.7 m $s^{-1}$ bias for $CLC\text{-}SLM_{44}$ to 1.4 m $s^{-1}$ bias for $USGS_{24}$, with all others in between, around 0.9 m $s^{-1}$. Regarding the RMSE (not shown), the order of performance is similar with $CLC\text{-}SLM_{44}$ performing best with 2.4 m $s^{-1}$ and $USGS_{24}$ worst with 3.1 m $s^{-1}$ and the CLC runs in between with values of 2.6–2.7 m $s^{-1}$ and res_1km also at 2.4 m $s^{-1}$. Looking at the $DACC_{45}$, Experiment #2 covers both $DACC_{45}$ extremes of the graph, res_dep (aka $CLC\text{-}SLM_{24}$) has a $DACC_{45}$ of ~55% while the res_1km has ~51%. Generally, the MLM method seems to improve wind speed skill in simulations: it improves the $DACC_{45}$ in $CLC_{44}$. This is not the case for the $CLC_{24}$ data set however.

**Table 8**
*Normalized Biases and RMSE of Humidity, Potential Temperature, Stratification ($\Delta\theta$), Wind Speed, and Wind Direction to Determine the Best Performing Run at the GBA Location*

| | Variable Metric | $q$ bias | $q$ RMSE | $\theta_{2m}$ bias | $\theta_{2m}$ RMSE | $\Delta\theta$ bias | $\Delta\theta$ RMSE | WS bias | WS RMSE | $DACC_{45}$ | Weighted average |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Weight LC data set | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | |
| | | | | | | Ranking in [0, 1], 0 the worst, 1 the best, no units | | | | | |
| Experiment #1 | $USGS_{24}$ | 0.18 | 0.09 | 0.87 | 0.06 | 0.94 | 0.25 | 0.00 | 0.00 | 0.53 | 0.35 |
| | $MODIS_{20}$ | 0.21 | 0.7 | 0.80 | 0.09 | 0.80 | 0.20 | 0.33 | 0.17 | 0.55 | 0.38 |
| | $CLC\text{-}SLM_{24}$ | 0.23 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.40 | 0.15 | 0.53 | 0.19 |
| | $CLC\text{-}MLM_{24}$ | 0.24 | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.38 | 0.17 | 0.55 | 0.19 |
| | $CLC\text{-}SLM_{44}$ | 0.00 | 0.03 | 0.96 | 0.09 | 0.87 | 0.23 | 0.43 | 0.17 | 0.51 | 0.38 |
| | $CLC\text{-}MLM_{44}$ | **0.02** | **0.04** | **0.96** | **0.10** | **0.88** | **0.23** | **0.53** | **0.22** | **0.51** | **0.40** |
| Experiment #2 | res_dep | 0.23 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.40 | 0.15 | 0.53 | 0.19 |
| | res_1km | 0.26 | 0.01 | 0.04 | 0.03 | 0.02 | 0.01 | 0.40 | 0.21 | 0.51 | 0.20 |
| | res_111m | **0.17** | **0.07** | **0.86** | **0.09** | **0.91** | **0.25** | **0.36** | **0.13** | **0.51** | **0.39** |
| | res_mos | 0.21 | 0.06 | 0.78 | 0.08 | 0.99 | 0.23 | 0.21 | 0.11 | 0.53 | 0.37 |

*Note*. 1 is the best, 0 the worst result for the weighted rank. The best options from either Experiment #1 or #2 are in bold.

## 5. Discussion

### 5.1. Performance Evaluation

Here, we try to determine the best overall simulation based on the bias and RMSE performances for the variables $\theta_{2m}$, stratification, humidity, wind speed, and wind direction at GBA. Determination of the skill of each simulation comes from the performance per metric. Each variable will get a score between 0 and 1 (1 being the best). We define $B_i$ as the bias of a simulation and $B_w$ as the worst bias among the simulations for that metric. The run score is defined as $1 - |B_i|/|B_w|$. Using the $\theta_{2m}$ bias as an example, a hypothetical bias of 0.00 K will give a score of 1, and the worst score will get 0. The same is done for all other RMSEs and biases. The DACC is already a normalized value of performance but it differs from other scores because both limits are defined whereas for other variables the worst performance is zero.

We ranked the simulations using a weighted average. We attributed all metrics a weight of 1, except for the DACC that has a weight of 2 because all other meteorological variables contribute with two metrics to the weighted average. Altogether, our method results in nine scores between 0 and 1. Note that the weight attribution is subjective and can be linked to the finality of the forecasts (i.e., dispersion modeling could favor wind direction score with a higher weight); a change of weight could lead to a change of ranking.

Table 8 shows the normalized performance of all the previously discussed variables, with the relative best simulations in bold. The best performing Experiment #1 simulation at the GBA location is the $CLC\text{-}MLM_{44}$. More generally, $MODIS_{20}$ and $CLC_{44}$ are the best LC data set to represent the surface with scores of [0.38, 0.40] whereas $CLC_{24}$, and $USGS_{24}$ LC remain below, ranking in ranges of [0.19, 0.35], with $USGS_{24}$ at 0.35 and $CLC_{24}$ at 0.19. Furthermore, from Experiment #1, MLM has a benefit over SLM aggregation, but the results may not settle the question of best aggregation method as they rank closely to one another. This similarity could be due to the absence of difference in LC close to the observation site (Figure 7f). Note that the 44-class LUT is not yet complete and requires additional work to specify all parameters for each of the 44 classes.

An additional point for attention is the introduction of the local climate zone (LCZ) classes between 31 and 41 for the functioning of the urban option WUDAPT. This claim on these classes brought some troubles to LUTs that have more than 31 classes, such as NCLD40. The release of WRF v4.4.1 fixed issues of NLCD40, therefore we advise to take into account that the introduction of $CLC_{44}$ in recent version of WRF could face similar problems as NCLD40 did.

In Experiment #2, res_111m outperforms the other simulations by scoring 0.39, the weighted metric average being comparable to the best Experiment #1 simulations. Interestingly, res_1km remains below indicating that

using the LC high-resolution map (3″) can lead to representable LC in the 1 × 1-km WRF cell. More astonishingly, Mosaic method "res_mos" does not lead to the best ranking: this should prompt further investigation.

The MLM method does not substantially influence the weighted average of the simulation, while the $CLC_{44}$ option do resemble the $CLC_{24}$ in these simulations. Even if the MLM method produces comparable results to the SLM in fine resolution at a given location, we expect more differences with simulations at coarser resolutions: as described in Table 2, the coarser the resolution; the higher the difference between aggregation methods. Further, we may expect a higher discrepancy between simulations at locations of interest that present more differences between the methods; this was the case at neither M30 nor GBA. In the end, MLM simulations produced the better simulation of Experiment #1, regardless of the LUT used.

Our newly introduced MLM method seems to perform well. However, in a few cases, the clustering of classes from CLC might not be always the most appropriate from a meteorological point of view. For example, do orchards and olive groves fit best in agricultural land, or do they correspond more closely to forest types? The Pineda et al. (2004)'s conversion places them in an agricultural USGS class. To determine the most appropriate structure, we recommend tests with other class structures. For example, Varga et al. (2020) tried two additional class clustering methods next to the conventional CLC class clustering. One method was based on physical behavior (Aldwaik et al., 2015) and the other method used thresholds of LC changes with the last edition of the CLC data set. Furthermore, MLM could be extended to other 1-level LC data sets by using hierarchical class system developed through meteorological parameters clustering.

Note that this research is too limited for a definite answer on the performance of the different LC data sets and aggregation methods since our case covers mostly specific weather conditions (clear sky, calm wind, and stable night) at one location. Even though, our results mismatch with most findings of other researchers. In Schicker et al. (2016)'s study, using $MODIS_{20}$ or $CLC_{24}$ instead of the default $USGS_{24}$ improves the temperature representation in simulations. Note that they did not specify whether they aggregate the 100-m and 250-m resolution CLC data sets to match domain resolutions, nor whether they used the Mosaic option. Their study is related to two areas in Austria and the simulations are evaluated against satellite data, surface observations, and radiosondes. We agree with their conclusion that $MODIS_{20}$ generally outperform $USGS_{24}$ but CLC is more dependent on the LUT in use. Their prospect that a complete LUT would improve the forecasts of WRF does not seem to come true with the 44-class LUT.

Santos-Alamillos et al. (2015) compare the wind from simulations based on $USGS_{24}$ and $CLC_{24}$ LC maps. No Mosaic option is mentioned for the 100 m $CLC_{24}$ data set. They find that while CLC represents the area better, $USGS_{24}$ creates better wind fields. They suggest that the nearest neighbor aggregation approach could introduce this paradox. Their CLC input data set has a 100-m horizontal resolution for 1-km WRF resolution using the nearest neighbor interpolation to aggregate. They mention that more studies are needed with different spatial interpolation procedures. Our study did not investigate the nearest neighbor approach as we expected it to be worse than SLM or MLM.

Li et al. (2018) simulate the Urban Heat Island of Berlin with the $USGS_{24}$, $CLC_{24}$, and Urban Atlas (UA) LC data sets, the latter being a product from Copernicus as is CLC. They applied the Mosaic approach on all three data sets. They found that the UA or CLC maps represent surface and air temperature better than the USGS does. In our study, $CLC_{44}$ similarly improves the representation of 2-m potential temperatures, while $CLC_{24}$ worsens the representation. In Li et al. (2018), the change is mainly attributed to a better representation of the urban fraction as the USGS map is outdated (survey from 1994) compared to CLC (data from 2012) and UA (from 2012, linked to CLC). We assume that the better surface representation of CLC is also the reason for the improvement in our study.

Golzio et al. (2021) test $MODIS_{20}$, $MODIS_{21}$ ($MODIS_{20}$ with additional lake class at 15″ resolution), $USGS_{24}$, and $CLC_{44}$. They determine that the $CLC_{44}$ performs better than the other LC data sets. Their results indicate that LC maps have comparable influences on the results as the parameterization changes do. We find that the resolution fed to WRF has a similar impact. Their findings do not match our results, although we note that in some single IOP cases the $CLC_{44}$ LUT does indeed perform well. Nonetheless, they have not checked $CLC_{24}$ to verify if the $CLC_{44}$ is indeed the best LC available at this location. Similarly, to our results, their $CLC_{44}$ LUT is not yet complete and needs further improvement.

### 5.2. Further Improvements

This work shows that how the LC maps are presented to WRF has an influence, as great as changing the LC data set (cf. the range of weighted average between Experiments #1 and #2). More research on setting options and at different locations is needed to determine optimal ways to present surface data sets to WRF. The inclusion of the MLM aggregation method as an option of WRF (or another NWP) is needed to fully evaluate its impact. The implementation of this method for categorical data will require a section that determines the number of levels considered in MLM and the grouping of the classes.

The comparison study only covered a period of 4-day, which is enough to get a first estimate of the performances with different LC data sets but not enough to confirm them unambiguously. As we computed the metric on a 3.5-days period (i.e., with 505 points from 10-min outputs), the confidence interval might be relatively low. A longer time period (from 1 month to 1 year hourly time series) will enable us to estimate these confidence intervals (e.g., through a bootstrap method). Similarly, we checked the results of the simulations against two observation points that were not subjected to changes caused by the aggregation method. Case studies at different locations would show in a broader sense how well MLM performs in simulations. Such simulations would benefit from weather stations observations at locations where the model shows a difference between SLM and MLM, in order to compare the performance of the two methods. In this study, the only differences observed at the location are advected from other places where the MLM and SLM methods differ.

The 44-class LUT does not show a great improvement with respect to the $USGS_{24}$. A large number of parameters, however, are taken from the original USGS tables. The CLC table needs to have 44 unique classes for more accurate comparisons between the two LUTs. Therefore, more fieldwork and satellite data analyses are required to update the last parameters not available in Pineda et al. (2004). The LANDUSE.TBL parameters that have not yet been updated are the heat capacity, and the snow cover effect on albedo (SCFX). In the VEGPARM.TBL variables related to vegetation such as stomata resistance, heat stress, LAI range, and vegetation height range also need an update for the 44-classes LUT.

The coarse maps of soil moisture content also influence the finer local simulations. For example, if the water storage is overestimated the influence of other surface parameters could be minimized and the spatial variability in atmospheric parameters could smoothen. Additional fine-resolution maps of several soil variables would therefore help to improve the simulations.

## 6. Conclusion

This study introduced a new method for spatial aggregation on land cover data sets. This method, called MLM aggregation, was developed to overcome difficulties with conventional aggregation techniques. It focused on discrepancies between maps with many classes and maps with a limited amount of classes, with possible artifacts caused by a too high level of detail in the initial data set. The MLM avoids such drawbacks while maintaining the richness of the initial data set. We applied this technique on the European Corine Land Cover database, which is well suited as it combines a high-resolution data set and a high number of classes with a hierarchical classification of the land cover categories in three distinct levels. We took advantage of this hierarchy, allowing us to choose a category always in agreement with its upper hierarchical classes in an aggregated cell. There has not yet been a spatial aggregation method that considers such a categorical interdependency to the authors' knowledge. Furthermore, to enhance the benefit from the high number of classes in the CLC data set, we made a freely available physical parameter table (LUT) with 44 classes, which can be used in WRF simulations. For the time being, our method was applied on the CLC data set, but it could be extended to other data sets, with the only condition that one or two parent levels of grouped categories were defined beforehand, not limited to land cover data. For example, it could be applied to sorting in a wide variety of sectors.

The impact of the aggregation method was evaluated through counting the number of cells of a given domain for which the retained category differs between the MLM and the conventional aggregation, both starting with the same original data set (CLC). These metrics were computed for different resolutions, over a $6° \times 4°$ area in southeastern France. For the smallest resolution (111 m), we observed a difference of only 0.03%, while the mismatch became 0.8% at 333-m resolution and reached 16.6% at 9-km resolution. These values are averages over the whole domain and can vary locally when a smaller area is considered, depending on the local LC variability.

A comparative study between various configurations combining different original land cover data sets (CLC, MODIS, and USGS), different classification and different aggregation methods was carried out through a series of WRF simulations. The WRF model was run with nested domain resolutions ranging from 27 to 1 km. The latter for the finest domain centered on an area documented through a field campaign conducted in the 2013 winter. Over the finest domain, we noted considerable differences between some configurations regarding average roughness length and albedo, which directly impact momentum flux (and therefore wind in the atmospheric boundary layer) and surface energy budget (hence the development of the atmospheric boundary layer).

To test the impact on the atmospheric flow of the various land cover configurations, we analyzed a 3.5-day period with clear sky, calm wind, diurnal convective boundary layer, and nocturnal valley wind development along the French DV. In the finest domain, the simulations were evaluated against observations collected during the 2013 field campaign. Regarding the simulated net radiation, CLC and MODIS outperform USGS. Though for humidity CLC does not show better results than MODIS and USGS. For stratification, USGS runs show a lower bias and RMSE. Concerning wind speed and direction, CLC with 24-class LUT and MODIS outperform the USGS simulation. However, CLC with 24-class LUT showed results worse than $USGS_{24}$ with the current model version.

When we turned to comparisons with observations, the different CLC aggregation methods showed limited differences in bias and RMSE of the predicted variables. However, the MLM method generally showed better results. The WRF LUT with 44 classes did improve the weighted average, even though this LUT contains several not yet well differentiated parameters. We must keep in mind that comparisons between observations and simulation were restricted to a single location with the same land cover type regardless of the aggregation method. Therefore, the differences among the various aggregation method simulations regarding atmospheric parameters do not have a local origin but are the result of the horizontal advection from upwind areas where surface conditions differ.

A second experiment on WRF input resolution of land cover maps showed that the best results were obtained with a single map having the same resolution as the finest WRF domain (here 111 m). Surprisingly, the promising Mosaic option was outperformed.

These simulations evaluations only span a relatively short period that cannot represent many different meteorological conditions. An extended time series would cope with this restriction and would enable us to compute confidence intervals on the metrics used to compare with observations. Additionally, multiple locations need to be considered, to verify whether our conclusion favoring the MLM still holds in different locations. Coarse simulations (9-km horizontal resolution) over a greater area such as the whole of France should be undertaken to better quantify the effect of aggregation differences. Such simulations are proper tools to further the knowledge of the impacts of land cover data sets.

## Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

## Data Availability Statement

The observational data of KASCADE-2013 measurements were used for verification and are available at: https://kascade.sedoo.fr/database/. The 44-class LUT used for simulations in WRF is available (de Bode, 2022a; https://doi.org/10.5281/zenodo.6027582). This link includes the creation and conversion applied. An aggregation script in python for the aggregation of SLM and MLM is available (de Bode, 2022b; https://doi.org/10.5281/ZENODO.6027704). Figures are made with anaconda python version 4.1.5.

**Acknowledgments**
We thank the reviewers for their contribution to our work, improving its quality. We acknowledge the CEA MRISQ project for the funding of the research.

## References

Aldwaik, S. Z., Onsted, J. A., & Pontius, R. G., Jr. (2015). Behavior-based aggregation of land categories for temporal change analysis. *International Journal of Applied Earth Observation and Geoinformation*, *35*, 229–238. https://doi.org/10.1016/j.jag.2014.09.007

Anderson, J. R., Hardy, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensor data*. U.S. Government Printing Office.

Bock, R. D., & Haggard, E. A. (1968). The use of multivariate analysis of variance in behavioral research. In *Handbook of measurement and assessment in behavioral sciences* (Vol. 1 (8)). Reading, MA: Addison-Wesley.

Broxton, P. D., Zeng, X., Sulla-Menashe, D., & Troch, P. A. (2014). A global land cover climatology using MODIS data. *Journal of Applied Meteorology and Climatology*, *53*(6), 1593–1605. https://doi.org/10.1175/JAMC-D-13-0270.1

Büttner, G. (2014). CORINE Land Cover and land cover change products. In I. Manakos, & M. Braun (Eds.), *Land use and land cover mapping in Europe: Practices & trends, remote sensing and digital image processing* (pp. 55–74). Dordrecht, Netherlands: Springer.
Cao, Y., & Fovell, R. G. (2016). Downslope windstorms of San Diego County. Part I: A case study. *Monthly Weather Review*, *144*(2), 529–552. https://doi.org/10.1175/MWR-D-15-0147
Cerenzia, I. (2017). Challenges and critical aspects in stable boundary layer representation in numerical weather prediction modeling: Diagnostic analyses and proposals for improvement. Tesi di dottorato, alma
Copernicus. (2020). CORINE Land Cover—Copernicus Land Monitoring Service [Dataset]. Retrieved from https://land.copernicus.eu/pan-european/corine-land-cover
Daniels, M. H., Lundquist, K. A., Mirocha, J. D., Wiersema, D. J., Chow, F. K., & Chow, F. K. (2016). A new vertical grid nesting capability in the weather research and forecasting (WRF) model. *Monthly Weather Review*, *144*(10), 3725–3747. https://doi.org/10.1175/MWR-D-16-0049.1
de Bode, M. (2022a). MichieldeB/Python-scripts: Transfer to zenodo (v1.0) [Software]. Zenodo. https://doi.org/10.5281/ZENODO.6027704
de Bode, M. (2022b). MichieldeB/WRF_tables: Github_WRF_tables. (v1.0) [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.6027582
de Bode, M., Hedde, T., Roubin, P., & Durand, P. (2021). Fine-resolution WRF simulation of stably stratified flows in shallow pre-alpine valleys: A case study of the KASCADE-2017 campaign. *Atmosphere*, *12*(8), 1063. https://doi.org/10.3390/atmos12081063
De Meij, A., & Vinuesa, J.-F. (2014). Impact of SRTM and Corine Land Cover data on meteorological parameters using WRF. *Atmospheric Research*, *143*, 351–370. https://doi.org/10.1016/j.atmosres.2014.03.004
Duan, H., Li, Y., Zhang, T., Pu, Z., Zhao, C., & Liu, Y. (2018). Evaluation of the forecast accuracy of near-surface temperature and wind in Northwest China based on the WRF model. *Journal of Meteorological Research*, *32*(3), 469–490. https://doi.org/10.1007/s13351-018-7115-9
Duine, G.-J. (2015). *Characterization of down-valley winds in stable stratification from the KASCADE field campaign and WRF mesoscale simulations (PhD)*. Université de Toulouse. Université Toulouse III-Paul Sabatier.
Duine, G.-J., Dupuy, F., Hedde, T., Roubin, P., Pardyjak, E., & Durand, P. (2017). KAtabatic winds and stability over CADarache for dispersion effluents. Retrieved from https://kascade.sedoo.fr/
Duine, G.-J., Hedde, T., Roubin, P., Durand, P., Lothon, M., Lohou, F., et al. (2017). Characterization of valley flows within two confluent valleys under stable conditions: Observations from the KASCADE field experiment. *Quarterly Journal of the Royal Meteorological Society*, *143*(705), 1886–1902. https://doi.org/10.1002/qj.3049
EEA. (2020). CLC 2012—Copernicus Land Monitoring Service [Dataset]. Retrieved from https://land.copernicus.eu/pan-european/corine-land-cover/clc-2012
Faroux, S., Tchuenté, A. T. K., Roujean, J.-L., Masson, V., Martin, E., & Le Moigne, P. (2013). ECOCLIMAP-II/Europe: A twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models. *Geoscientific Model Development*, *6*(2), 563–582. https://doi.org/10.5194/gmd-6-563-2013
Golzio, A., Ferrarese, S., Cassardo, C., Diolaiuti, G. A., & Pelfini, M. (2021). Land-use improvements in the weather research and forecasting model over complex mountainous terrain and comparison of different grid sizes. *Boundary-Layer Meteorology*, *180*(2), 319–351. https://doi.org/10.1007/s10546-021-00617-1
Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Lu, L., et al. (2013). Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *International Journal of Remote Sensing*, *34*(7), 2607–2654. https://doi.org/10.1080/01431161.2012.748992
Gudowicz, J., & Zwolinski, Z. (2016). The fourth level of the LULC classification in the integrated environmental monitoring in Poland. *Presented at the 2nd EARSeL SIG LU/LC and NASA LCLUC Joint Workshop, May 6, Prague, Czech Republic*.
Hartmann, D. L. (2015). *Global physical climatology* (Vol. 103). Newnes.
Hersbach, H., Bell, B., Paul, B., Biavati, G., Horányi, A., Sabater, J. M., et al. (2018a). ERA5 hourly data on model level from 1978 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.adbb2d47
Hersbach, H., Bell, B., Paul, B., Biavati, G., Horányi, A., Sabater, J. M., et al. (2018b). ERA5 hourly data on pressure levels from 1978 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*. https://doi.org/10.24381/cds.bd0915c6
Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., et al. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 National Land Cover Database. *ISPRS Journal of Photogrammetry and Remote Sensing*, *162*, 184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.019
Hong, S.-Y., Kim, J.-H., Lim, J.-o., & Dudhia, J. (2006). The WRF single moment microphysics scheme (WSM). *Journal of the Korean Meteorological Society*, *42*, 129–151.
Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*, D13103. https://doi.org/10.1029/2008JD009944
Ikİel, C., Dutucu, A. A., Ustaoglu, B., & Derya KiİİC, E. (2012). Land Use and Land Cover (LULC) classification using SPOT-5 image in the Adapazari plain and its surroundings, Turkey. *The Online Journal of Science and Technology (TOJSAT)*, *2*(2), 37–42.
Inglada, J., Vincent, A., & Vincent, T. (2017). Theia OSO Land Cover Map 2106 [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.1048161
Inglada, J., Vincent, A., & Thierion, V. (2018). 2017 Metropolitan France Land Cover Map—CESBIO—OSO [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.1993596
Jach, L., Warrach-Sagi, K., Ingwersen, J., Kaas, E., & Wulfmeyer, V. (2020). Land cover impacts on land-atmosphere coupling strength in climate simulations with WRF over Europe. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD031989. https://doi.org/10.1029/2019JD031989
Kain, J. S. (2004). The Kain-Fritsch convective parameterization: An update. *Journal of Applied Meteorology and Climatology*, *43*(1), 170–181. https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2
Kalverla, P. C., Duine, G. J., Steeneveld, G. J., & Hedde, T. (2016). Evaluation of the weather research and forecasting model in the Durance valley complex terrain during the KASCADE field campaign. *Journal of Applied Meteorology and Climatology*, *55*(4), 861–882. https://doi.org/10.1175/JAMC-D-15-0258.1
Kleczek, M. A., Steeneveld, G. J., & Holtslag, A. A. M. (2014). Evaluation of the weather research and forecasting mesoscale model for GABLS3: Impact of boundary-layer schemes, boundary conditions and spin-up. *Boundary-Layer Meteorology*, *152*(2), 213–243. https://doi.org/10.1007/s10546-014-9925-3
Kosztra, B., Büttner, G., Hazeu, G., & Arnold, S. (2019). *Updated CLC illustrated nomenclature guidelines* (pp. 1–124). Wien, Austria: European Environment Agency.
Li, D., Bou-Zeid, E., Barlage, M., Chen, F., & JamesSmith, A. (2013). Development and evaluation of a Mosaic approach in the WRF-Noah framework. *Journal of Geophysical Research: Atmospheres*, *118*, 11918–11935. https://doi.org/10.1002/2013JD020657

Li, H., Wolter, M., Wang, X., & Sodoudi, S. (2018). Impact of land cover data on the simulation of urban heat island for Berlin using WRF coupled with bulk approach of Noah-LSM. *Theoretical and Applied Climatology*, *134*(1), 67–81. https://doi.org/10.1007/s00704-017-2253-z

Lindvall, J., & Svensson, G. (2015). The diurnal temperature range in the CMIP5 models. *Climate Dynamics*, *44*(1/2), 405–421. https://doi.org/10.1007/s00382-014-2144-2

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., & Merchant, J. W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing*, *21*(6–7), 1303–1330. https://doi.org/10.1080/014311600210191

Miguez-Macho, G., Stenchikov, G. L., & Robock, A. (2004). Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. *Journal of Geophysical Research*, *109*, D13104. https://doi.org/10.1029/2003JD004495

Neumann, K., Herold, M., Hartley, A., & Schmullius, C. (2007). Comparative assessment of CORINE2000 and GLC2000: Spatial analysis of land cover data for Europe. *International Journal of Applied Earth Observation and Geoinformation*, *9*(4), 425–437. https://doi.org/10.1016/j.jag.2007.02.004

Oke, T. R. (2002). *Boundary layer climates*. Routledge.

Pineda, N., Jorba, O., Jorge, J., & Baldasano, J. M. (2004). Using NOAA AVHRR and SPOT VGT data to estimate surface parameters: Application to a mesoscale meteorological model. *International Journal of Remote Sensing*, *25*(1), 129–143. https://doi.org/10.1080/0143116031000115201

Prósper, M. A., Otero-Casal, C., Fernández, F. C., & Miguez-Macho, G. (2019). Wind power forecasting for a real onshore wind farm on complex terrain using WRF high resolution simulations. *Renewable Energy*, *135*, 674–686. https://doi.org/10.1016/j.renene.2018.12.047

Román-Cascón, C., Lothon, M., Lohou, F., Hartogensis, O., Vila-Guerau de Arellano, J., Pino, D., et al. (2021). Surface representation impacts on turbulent heat fluxes in the weather research and forecasting (WRF) model (v.4.1.3). *Geoscientific Model Development*, *14*(6), 3939–3967. https://doi.org/10.5194/gmd-14-3939-2021

Santos-Alamillos, F., Fransisco, J., Pozo-Vázquez, D., Ruiz-Arias, J. A., & Tovar-Pescador, J. (2015). Influence of land-use misrepresentation on the accuracy of WRF wind estimates: Evaluation of GLCC and CORINE land-use maps in southern Spain. *Atmospheric Research*, *157*, 17–28. https://doi.org/10.1016/j.atmosres.2015.01.006

Santos-Alamillos, F. J., Fransisco, J., Pozo-Vázquez, D., Ruiz-Arias, J. A., Lara-Fanego, V., & Tovar-Pescador, J. (2013). Analysis of WRF model wind estimate sensitivity to physics parameterization choice and terrain representation in Andalusia (southern Spain). *Journal of Applied Meteorology and Climatology*, *52*(7), 1592–1609. https://doi.org/10.1175/JAMC-D-12-0204.1

Schicker, I., Arnold Arias, D., & Seibert, P. (2016). Influences of updated land-use datasets on WRF simulations for two Austrian regions. *Meteorology and Atmospheric Physics*, *128*(3), 279–301. https://doi.org/10.1007/s00703-015-0416-y

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. (2019). A description of the advanced research WRF model version 4. https://doi.org/10.5065/1dfh-6p97

Steeneveld, G.-J., & de Bode, M. (2018). Unravelling the relative roles of physical processes in modelling the life cycle of a warm radiation fog. *Quarterly Journal of the Royal Meteorological Society*, *144*(714), 1539–1554. https://doi.org/10.1002/qj.3300

Sterk, H. A., Marina, G.-J. S., & Holtslag, A. M. (2013). The role of snow-surface coupling, radiation, and turbulent mixing in modeling a stable boundary layer over Arctic Sea ice. *Journal of Geophysical Research: Atmospheres*, *118*, 1199–1217. https://doi.org/10.1002/jgrd.50158

Sukoriansky, S. (2008). *Implementation of the quasi-normal scale elimination (QNSE) model of stably stratified turbulence in WRF*. Beer-Sheva.

Sukoriansky, S., Galperin, B., & Perov, V. (2005). Application of a new spectral theory of stably stratified turbulence to the atmospheric boundary layer over sea ice. *Boundary-Layer Meteorology*, *117*(2), 231–257. https://doi.org/10.1007/s10546-004-6848-4

Tabachnick, B. G., & Fidell, L. S. (2011). Multivariate Analysis of Variance (MANOVA). In M. Lovric (Ed.), (Red.), *International Encyclopedia of statistical science* (pp. 902–904). Springer. https://doi.org/10.1007/978-3-642-04898-2_394

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K. E., et al. (2004). Implementation and verification of the unified NOAH land surface model in the WRF model (Formerly Paper Number 17.5). In *20th Conference on weather Analysis and forecasting/16th Conference on numerical weather prediction* (pp. 11–15).

Varga, O. G., Pontius, R. G., Jr., Szabó, Z., & Szabó, S. (2020). Effects of category aggregation on land change simulation based on Corine Land Cover data. *Remote Sensing*, *12*(8), 1314. https://doi.org/10.3390/rs12081314

Velde, I. R. v. d., Steeneveld, G.-J., Schreur, B. G. J. W., & Holtslag, A. A. M. (2010). Modeling and forecasting the onset and duration of severe radiation fog under frost conditions. *Monthly Weather Review*, *138*(11), 4237–4253. https://doi.org/10.1175/2010MWR3427.1

Wang, W., & Gill, D. (2012). WRF nesting. Presented at the WRF Tutorial 2012 Brazil, October 18, Sao Paulo, Brazil.

Weaver, C. P., & Avissar, R. (2001). Atmospheric disturbances caused by human modification of the landscape. *Bulletin of the American Meteorological Society*, *82*(2), 269–282. https://doi.org/10.1175/1520-0477(2001)082<0269:ADCBHM>2.3.CO;2

Wyszogrodzki, A. A., Liu, Y., Jacobs, N., Childs, P., Zhang, Y., Roux, G., & Warner, T. T. (2013). Analysis of the surface temperature and wind forecast errors of the NCAR-AirDat operational CONUS 4-km WRF forecasting system. *Meteorology and Atmospheric Physics*, *122*(3), 125–143. https://doi.org/10.1007/s00703-013-0281-5

Yang, Z.-L. (2004). Modeling land surface processes in short-term weather and climate studies. In *Observation, theory and modeling of atmospheric variability* (Vol. 3, pp. 288–313). World Scientific Series on Asia-Pacific Weather and Climate. World Scientific.