Land Use and Topography Influence in a Complex Terrain Area: A High Resolution Mesoscale Modelling Study over the Eastern Pyrenees using the WRF Model

B. Jiménez-Esteve[a,b], M. Udina[a], M.R. Soler[a], N. Pepin[d], J.R. Miró[c]

[a] *Departament de Física Aplicada, Meteorologia. Universitat de Barcelona. C/ Martí i Franquès, 1. 08028 Barcelona.*

[b] *Institue for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland*

[c] *Applied Research and Modelling Department, Meteorological Service of Catalonia, Barcelona, Spain*

[d] *Department of Geography, University of Portsmouth, United Kingdom*

Abstract

Different types of land use (LU) have different physical properties which can change local energy balance and hence vertical fluxes of moisture, heat and momentum. This in turn leads to changes in near-surface temperature and moisture fields. Simulating atmospheric flow over complex terrain requires accurate local-scale energy balance and therefore model grid spacing must be sufficient to represent both topography and land-use. In this study we use both the Corine Land Cover (CLC) and United States Geological Survey (USGS) land use databases for use with the Weather Research and Forecasting (WRF) model and evaluate the importance of both land-use classification and horizontal resolution in contributing to successful modelling of surface temperatures and humidities observed from a network of 39 sensors over a 9 day period in summer 2013. We examine case studies of the effects of thermal inertia and soil moisture availability at individual locations. The scale at which the LU classification is observed influences the success of the model in reproducing observed patterns of temperature and moisture. Statistical validation of model output demonstrates model sensitivity to both the choice of LU database used and the horizontal resolution. In general, results show that on average, by a) using CLC instead of USGS and/or b) increasing horizontal resolution, model performance is improved. We also show that the sensitivity to these changes in the model performance shows a daily cycle.

Email addresses: bernat.jimenez@env.ethz.ch (B. Jiménez-Esteve), mudina@meteo.ub.edu (M. Udina), rosa@meteo.ub.edu (M.R. Soler), nicholas.pepin@port.ac.uk (N. Pepin), jrmiro@meteo.cat (J.R. Miró)

Keywords: Land use, topography, high resolution, mesoscale modelling, complex terrain.

1. Introduction

The land use (LU) is an important parameter in atmospheric models which describes the properties of the land including modifications due to human activities. It regulates the exchanges of heat, moisture and momentum between the soil and the air, which in numerical models determine the calculation of meteorological magnitudes (e.g. temperature, humidity) near the surface. Another important parameter when simulating atmospheric flows near the surface, especially in areas of complex terrain, is the topographic relief which is mainly determined by the model horizontal grid spacing chosen in the numerical model configuration.

The LU classification includes information about whether a region is covered by urban areas, forests, wetlands, croplands or water and also how people use land, e.g. extensively or intensively, for urban development or for conservation. In the Weather Research and Forecasting (WRF) atmospheric model (Skamarock et al., 2008) the default LU categories are determined using the USGS global land-use map (Anderson, 1976) with a spatial resolution 30" ($\sim$1 km in mid-latitudes). In Europe, the more recent Coordination of Information on the Environment (CORINE) land-use dataset (EEA, 2000) exists with 100-m spatial resolution, although it has to be adapted for use in the WRF model.

LU is classified into discrete categories, each characterized by six physical parameters: the roughness length z_0, thermal inertia λ_T, soil moisture availability M, albedo α, surface heat capacity C and surface emissivity ϵ. All of these parameters play a role in calculating surface heat and moisture fluxes in land-surface models (LSMs), gathering information from the surface layer scheme (the exchange coefficients), the radiation scheme (radiative forcing) and from the microphysics and convective schemes (precipitation forcing). Calculated moisture, heat and momentum fluxes at the lower level of the model are used as the lower boundary condition for the planetary boundary layer (PBL) scheme.

So far, it has been demonstrated that LU changes can have some significant effect on meteorological simulations (Pineda et al., 2004; Cheng & Byun, 2008). Recently, several authors investigated the effect of LU on wind speed and wind direction forecasts. In particular, De Meij & Vinuesa (2014) and Santos-Alamillos et al. (2015) showed that Corine Land Cover database in WRF simulations improved the wind speed forecast due to a better representation of the urban fraction. It has also been shown that different LSMs can lead to different results when resolving cold air pools or low level jets (Prabha et al., 2011) and, recently, Cuxart et al. (2016) has shown how surface heterogeneity influences the surface energy budget. Better representation and periodic updating of LU changes has been shown to improve the performance of meteorological and air quality

models (Cheng et al., 2013; Lam et al., 2006; Civerolo et al., 2007). In addition, precipitation and wind fields have also been shown to be dependent on LU data (Cheng et al., 2013).

In areas of complex terrain with steep slopes it is necessary to increase grid resolution in order to correctly resolve the topographic relief. In the first instance, this should help the model to better forecast magnitudes at the surface (such as temperature, humidity or wind). If resolution is increased too rapidly however, limitations in mesoscale atmospheric models can arise. For steep terrain, smoothing may be necessary because errors in calculation of the horizontal pressure gradient can occur when there are large differences in elevation between adjacent grid cells (Arnold et al., 2012). Several studies have explored the effect of changing grid resolution on simulating wind flows in complex terrain areas including the Italian Alps (Giovannini et al., 2014) and on the northern and southern side of the Pyrenees (Jiménez & Cuxart, 2014; Pagès et al., 2016; Udina et al., 2017). However, there are few studies which have applied the mesoscale approach at grid sizes below 1 km (Horvath et al., 2012; Seaman et al., 2012; Udina et al., 2017). This is because when mesoscale models are run over high resolution grids of hundreds of meters, one has to assume that the energy-containing turbulence is much smaller than the grid size for the mesoscale approach (Wyngaard, 2010), which depends on the time of the day and the size of turbulent eddies (Cuxart, 2015). Alternatively, if the size of the turbulent eddies is similar or larger than the grid size the large-eddy simulation approach may be required. As most widely available LU datasets do not have a resolution below 1 km, few mesoscale modeling studies have been able to evaluate the success in simulating surface magnitudes for grid spacings of hundreds of meters.

In this work, we first adapt a new LU dataset for use with the WRF model in order to explore the influence of LU and topography on prediction of surface magnitudes at resolutions finer than 1 km. Our objective is to contribute to a better understanding of model performance in areas of complex terrain. In particular, the study concentrates on the eastern Pyrenees, an area with complex microclimate where air temperature and wind flows are strongly constrained by topography. There is a strong diurnal variability in temperature (Pepin & Kidd, 2006) with frequent temperature inversion conditions at night which usually disperse during the day (Pagès & Miró, 2010).

The main objectives of the present work are:

1. to study how changing from one land use database to an updated one can improve forecast skill of surface temperature and humidity;
2. to quantify the effect of increasing horizontal resolution down to 500 m in an area of complex terrain; and
3. to try to distinguish between the effect of changing the resolution and of changing the LU database on model performance.

The present work is organized as follows: The discussion of general synoptic

conditions during the study period, model configuration and observational data are described in section 2. The simulation results including the sensitivity analysis of both land use and topography, along with a statistical evaluation are included in section 3. Finally section 4 provides a brief discussion and summary of the results.

2. Methodology

2.1. Reclassification of CLC to USGS land use categories

We adapted the Corine Land Cover (CLC) database in order to be used in the WRF meteorological simulations. This was done by reclassifying each CLC category to the most similar USGS category following the method of Pineda et al. (2004), they created a correspondence table between the two land use datasets. The process involves: (i) downloading raster data on land cover for the CLC2006 inventory (Büttner & Kosztra, 2007) (ii) changing the projection of the raster file to WGS84 used by WRF Preprocessing System (WPS) (iii) reclassifying the LU to USGS corresponding categories following Table 1 of Pineda et al. (2004), (iv) converting the raster image to an ASCII file and (v) rearranging the data and converting to WPS binary format. The end result is that the 44 CLC LU categories are reclassified into only 13 different LU categories corresponding to the USGS classification.

2.2. Model configuration

The model run simulated a 10 day summer period, although only 9 days were used for statistical study because the first 24 hours are used for spin-up. The study area consists of the valley of 'La Cerdanya' in the Eastern Pyrenees. It is a classic glacial rift with a characteristic U shape and is approximately orientated north-east to south-west. Its base is around 1000 m above sea level and it is surrounded by mountain ranges with peaks over 2500 m on the South and North sides. It is limited by the Segre gorge downstream to the west and the plain of Mont-Lluís (1500 m) to the east. These characteristics make the central part of the valley (upstream of the gorge) a one where strong thermal inversions can occur if the synoptic pattern is favorable (Pagès et al., 2016).

WRF version 3.6.1 was used for meteorological simulations. The WRF model (Skamarock et al., 2008) is a mesoscale numerical weather prediction system designed to serve both operational forecast and atmospheric research needs. The model configuration (Table 1) consists of four domains centered on the Cerdanya valley (Figure 1). Domains D1 (9 km), D2 (3 km) and D3 (1 km) are run from 0000 UTC 01 August to 0000 UTC 11 August 2010 with output files saved every half-hour. A finer resolution domain D4 (155 x 155, 0.5 km grid spacing) is nested from domain D3 in one-way nesting, starting on 0000 UTC 02 August and finishing at 0000 UTC 11 August, with input boundary conditions set every

Table 1: Model configuration options used for WRF simulations

	D1	D2	D3	D4
Horitzontal grid:	9 km	3 km	1 km	500 m
Dimensions (x, y, z):	(202, 202, 48)	(202, 202, 48)	(202, 202, 48)	(155, 155, 48)
Time step:	27 s	9 s	3 s	0.75 s
Initial and boundary conditions	ERA-Interim reanalysis from ECMWF (every 6h)			d3 external domain ndown (every 30min)
Simulated period	From 0000UTC 01 August 2013 to 0000UTC 11 August 2013			From 0000UTC 02 Aug 2013 to 0000UTC 11 August 2013
Radiation	Dudhia scheme for short-wave radiation. RRTM for long-wave radiation.			
Land surface	NOAH land-surface model (4 subsoil layers)			
Microphysics	New Thompson graupel scheme			
Cumulus	Kain-Fritsch (new Eta) (only for d1)			
PBL	Yonsei University scheme (YSU)			
Surface Layer	MM5 similarity surface layer (revised Monin-Obuckov)			
LU data sim 1.	USGS	USGS	USGS	USGS
sim 2.	USGS	CLC	CLC	CLC

30 minutes from the domain 3 output files. All 4 domains are one-way nested, which means that outer domains provide lateral conditions to inner domains but not vice-versa. In the vertical, 48 sigma levels are used from the ground up to 100 hPa for all domains with the first level 1.9 m above the surface, and the first 20 levels all within the first 250 m. Time steps for the simulations are 27 s (9-km), 9 s (3-km), 3 s (1-km) and 0.75 s for the highest resolution domain (0.5-km).

Initial and boundary conditions for domain D1 are taken from ERA-Interim reanalysis provided by ECMWF (European Centre for Medium-Range Weather Forecasts), which is presented as a gridded data set at approximately 0.125° spatial resolution, interpolated from original 0.75° data, and with 38 atmospheric levels. Boundary conditions are forced every 6 h.

The NOAH land surface scheme (Chen & Dudhia, 2001) is the land surface model used in this simulation along with revised MM5 similarity for the surface layer scheme (Jiménez et al., 2012) (new in WRF 3.6.1) based on the Monin-Obukhov similarity theory (MOST) (Obukhov, 1946; Monin & Obukhov, 1954). The first vertical model level is considered to be within the surface layer so the surface layer scheme computes the stability dependent coefficients which, together with the land surface model, permit a calculation of surface turbulent fluxes for the PBL scheme (the Yonsei University PBL - (Hong et al., 2006)). This has first order closure where turbulent fluxes and variances are determined using vertical gradients in the absence of the Turbulent Kinetic Energy (TKE) prediction equation. The entrainment is made proportional to the surface buoyancy flux in line with results from studies with large-eddy models (Noh et al., 2003). The physics package also includes the rapid radiative transfer model (RRTM) scheme for long-wave radiation (Mlawer et al., 1997); the Dudhia scheme for short-wave radiation (Dudhia, 1989); the new Thompson microphysics scheme (Thompson et al., 2004); and the Kain-Fritsch cumulus scheme (Kain, 2004).

The Shuttle Radar Topography Mission (SRTM) 90 m topography data (Farr et al., 2007) had previously been adapted to the WRF model topography in order to be able to perform high-resolution simulations with grid sizes smaller than 1 km. The finer resolution domains of 1-km (D3) and 0.5-km (D4) (Figure 1b) were run using the two LU datasets, CLC and USGS, to study the effect of both changing horizontal resolution and the choice of LU dataset in this complex terrain area. Domain D3 (1-km grid resolution) is used to compare the difference between two LU datasets since its relatively high resolution permits us to compare with available observations (see section 2.3). Domain D4 (0.5 km) is used to compare with the 1-km simulation in order to study the topographic effect.

2.3. Observational data: 'La Cerdanya' transects

There are 50 sensors measuring temperature ($^{\circ}C$) and relative humidity (%) located at different altitudes across La Cerdanya. The Hobo U23-001 sensors

have been in operation since July 2012. A similar type of sensor was evaluated by Whiteman et al. (2000) and shown to be adequate for field temperature monitoring. Data has been recorded every 30 minutes since installation. The observational data was collected by the University of Portsmouth (UK) who, together with the Meteorological Service of Catalonia, are carrying out a campaign to quantify the temperature and humidity patterns across Cerdanya (Pagès et al., 2016). The main objective is to understand the mechanisms and extent of thermal inversions and to what extent climate change may affect this process.

To protect the sensors from atmospheric conditions (particularly direct solar irradiance) they are attached inside an open white cylindrical PVC pipe about 30 cm in length, installed at a standard height 1.5 meters above ground level in evergreen trees, the top end of the tube facing to the north with an inclination of around 45 degrees. Data from 39 sensors are used in this study. These are grouped in 8 transects, 7 of them cover an elevational range from the valley bottom ($\sim$ 1000 m) to above treeline ($\sim$ 2400 m) (Cad, Eyne, Font, Lles, Mal, Mas, Rom) while the final transect (Val) goes along the valley axis (Table 2). Numbers, from 1 to 5, indicate the position in each transect from bottom (5) to top (1). Figure 2 shows the location of each sensor and corresponding transect code.

2.4. Episode synoptic characterization and and analytical approach

This study focuses on a 9 day period (2-10 August 2013). The synoptic situation over the Pyrenees can be classified into 3 sub-periods. The first from 2nd to 6th August was dominated by high pressure over central Europe and a low located in the north Atlantic. This helped to develop a heat wave over the Iberian Peninsula with temperatures well above $20°C$ at 850 hPa. The second period from 7th to 8th August was dominated by the passage of a mid and upper-level trough. During this period the temperatures decreased but very little precipitation fell in the zone, although clouds increased. After the passage of this trough, fair weather returned and temperatures recovered during the 9th and 10th August. Overall the majority of days were dominated by clear skies.

First, we perform a physical interpretation based on single transect points as examples across 'La Cerdanya' focusing on two main influences that alter our results: land use and topographic resolution. We analyse one case study in which the changes in temperature and humidity are due to change in land use type and two more case studies that illustrate a better performance of the model due to a better resolution of the relief. Secondly we perform a more extensive statistical analysis to evaluate general model performance and third we examine the geography of model performance across sites. Finally, we develop a sensitivity method which distinguishes the relative importance of the effects of land use versus horizontal resolution and we analyse when these contributions are more or less important.

Table 2: Height and Land use index for each transect point and for the different simulations.

	Height (m)			LU INDEX			
sensor code	real	1x1 km	500x500 m smoothed	CLC 1x1 km	CLC 500x500 m	USGS 1x1 km	USGS 500x500 m
Cad1	2228.9	1771.0	1970.6	14	7	5	2
Cad2	1953.8	1436.1	1653.7	14	14	10	5
Cad3	1703.8	1308.2	1510.6	9	14	10	5
Cad4	1451.5	1170.2	1287.7	9	9	2	2
Cad5	1221.5	1062.2	1143.0	6	9	11	2
Eyne1	2282.5	2066.5	2093.9	14	14	2	2
Eyne2	2050.8	1810.9	1964.1	14	14	2	2
Eyne3	1846.3	1810.9	1849.2	14	14	2	2
Eyne4	1668.0	1754.8	1718.3	14	2	2	15
Font0	2381.1	2230.0	2289.6	7	7	5	5
Font1	2095.1	1998.3	2120.9	14	14	2	2
Font2	1836.2	1582.1	1658.6	14	14	2	2
Font3	1610.3	1816.2	1784.3	14	7	2	2
Font4	1318.8	1202.6	1244.0	9	7	2	2
Font5	1033.8	1136.1	1203.1	7	9	2	2
Lles1	2334.4	2568.1	2470.4	7	7	7	7
Lles2	2078.2	2143.4	2066.9	14	14	7	7
Lles3	1830.9	1977.8	1826.3	14	9	15	15
Lles4	1588.5	1811.0	1666.1	14	14	15	15
Lles5	1306.5	1281.8	1301.8	9	9	15	15
Mal1	2361.1	2380.0	2380.0	9	9	7	7
Mal2	2099.6	2160.3	2107.8	7	7	2	10
Mal3	1847.0	1941.7	1953.1	7	9	2	2
Mal4	1585.1	1698.7	1652.0	9	9	2	2
Mal5	1320.0	1309.0	1286.0	9	9	2	2
Mas1	2484.4	2187.3	2238.8	7	7	3	3
Mas2	2123.9	1921.3	1974.3	14	14	2	2
Mas3	1753.1	1579.4	1710.6	14	14	11	2
Mas4	1410.1	1305.9	1389.8	9	14	2	11
Mas5	1212.1	1266.5	1203.8	9	9	2	2
Rom1	2104.7	2020.4	2079.9	14	9	15	2
Rom2	1789.6	1934.8	1845.1	7	7	15	15
Rom4	1540.1	1623.7	1584.4	1	1	2	2
Rom5	1365.7	1494.2	1471.1	7	2	2	2
Val1	738.9	966.8	834.9	2	11	15	15
Val2	912.9	1188.3	1064.4	2	6	2	2
Val3	1015.3	1123.1	1056.5	3	3	15	15
Val4	1109.0	1081.3	1089.9	6	6	2	2
Val5	1265.5	1390.8	1301.7	7	6	2	2

3. Results

3.1. A case of LU sensitivity

Figure 3 shows the LU index from the WRF output files for the two experiments: USGS (Fig. 3a, c) and CLC (Fig. 3b, d) over the two domains using 1-km (part of D3) and 0.5-km (D4) horizontal grid spacing (note the same area is shown in each case) The index corresponds to the dominant LU category at each grid point (see the key). The CLC LU dataset presents a more realistic distribution, clearly distinguishing the valley area, dominated by Dryland and Irrigated Cropland and Pastures, and some Urban and Built-up Land. In contrast, the USGS classification in the valley is less diverse and does not have any urban grid points. When we increase the USGS horizontal resolution (Fig. 3c) we do not see any improvement in LU detail because the original database resolution is only 1 km. On the other hand, we do obtain a better representation for CLC-500 m (Fig. 3d). The urban area (LU category 1) in domain D4 is 0% for USGS but 0.9% for CLC. The 'Dryland Cropland and Pasture' (LU category 2) changes from 24% in USGS to only 5% in CLC. Forest area expands significantly in CLC. For example, 'Evergreen Needleleaf Forest' (category 14) represents 6% of pixels in USGS but 25% in CLC, closer to reality. On the other hand mixed forest (category 5) reduces from 14% to 7% and 'Wooded tundra' (category 21), present in USGS, disappears completely in CLC. 'Barren or Sparsely Vegetated' (category 19) terrain goes from 0% to more than 11%. This change occurs above treeline in the alpine zone which dominates high elevations (>2000 m). In the USGS, wooded tundra (category 21) is assigned to the highest elevation regions, which has too high a thermal inertia and is not realistic.

Figure 4 shows maps of physical parameters derived from the land use classifications at 1 km resolution. Lower thermal inertia (Fig. 4e) of category 19 'Barren or Sparsely Vegetated' compared to category 21 'Wooded Tundra' suggests that higher temperatures will be reached in the CLC simulations compared with USGS. Also increased urbanization in CLC will also modify the physical proprieties (De Meij & Vinuesa, 2014; Santos-Alamillos et al., 2015). Soil moisture availability also differs among the two simulation experiments (Fig. 4c). In general, but especially in the elevated terrain areas, M is larger in USGS which will lead to increased evaporation. However, in 'La Cerdanya' valley the opposite is the case, and higher soil moisture availability and lower thermal inertia occurs in the CLC database(Fig. 4c,e).

Due to advective processes, LU changes can potentially influence the entire domain. However looking closer at direct impacts helps us to understand the main mechanisms that lead to changes in surface variables. We choose the sensor at Mal5 (marked in bold in Table 2) to illustrate some of these effects.

Mal5 is located at 1320 m a.s.l. In domain D3 the nearest pixel has an elevation of 1309 m, so there is minimal altitude error. LU changes from category 2 (Dryland Cropland and Pasture) in USGS to category 9 (Mixed Shrubland/Grassland)

in CLC which has drier characteristics. Fig. 5 shows a) temperature, and b) specific humidity at 2 meters above the surface for D3. Temperatures are higher in the CLC simulation than in USGS, both during day and night. However, the two simulations underestimate considerably the observed daily maximum temperature by about 4 degrees Celsius.

The CLC experiment has lower values of specific humidity (Fig. 5b), because LU category 9 has lower M, which means less evaporation takes place at the surface. In Fig. 6 latent heat flux (LH) and sensible heat flux (SH) are represented for this point for both CLC and USGS simulations. Clearly the CLC simulation has much lower LH but higher SH, indicative of a drier environment.

Skin temperature (Ts) also shows pronounced differences, being closely related to thermal inertia and also albedo and emissivity of the surface (Fig. 7). Ts is higher during the day in the CLC simulation because LU category 9 has a lower thermal inertia than category 2. The difference in daytime maximum is often more than $5°C$. This explains why 2 meter air temperature is also higher in CLC and why CLC has greater SH than USGS, since SH is proportional to the vertical temperature gradient.

3.2. Topography and horizontal resolution sensibility

In complex terrain resolving the relief accurately is very important. We expect increased spatial resolution to improve the simulations of magnitudes near the surface (such as temperature and humidity). However, a fine resolution can lead to steeper slopes that become a limitation for numerically resolving equations with terrain-following sigma coordinates. The topographic bias for the 1 km and 500 m grids are shown in Fig. 8, for the same sub-area as in Figure 3. The 500 m grid has been smoothed to avoid possible instability divergence errors that normally arise if there are slopes of more than 45° (Arnold et al., 2012). The errors in D4 are reduced significantly despite this smoothing.

To investigate such influences two data points are chosen: one near the mountain peaks which presents a great improvement (reduction) in height error at 500 m compared with 1 km (Cad2) and another at the bottom of the valley (Val4). Cad2, is located at 1954 m (see Fig. 2). In domain D3 (1-km) the model elevation is 1436 m (absolute error of more than 500 m). In domain D4 (0.5-km), model elevation is 1654 m, reducing the error by 300 m. As expected, time evolution of temperature is much better represented in D4 than in D3 (Fig. 9). The LU category does not change in the two simulations and is consistent at category 14 'Evergreen Needleleaf Forest'. There is a clear improvement in temperature evolution for the CLC-500m simulation (yellow line in Fig. 9a), which may be attributed to the smaller elevation error. For both the USGS-1km and CLC-1km simulations (blue and red lines in Fig. 9a) the temperature is overestimated, a consequence of the lower altitude in the 1-km resolution. Thus, finer resolution improves the temperature forecast. On the other hand, we can also see remarkable differences between the USGS-500m and the CLC-500m

(green and yellow lines in Fig. 9a). CLC-500m is much closer to observations, which indicates also an improvement of the temperature forecast due to a better representation of the LU in the surrounding area. Humidity differences between simulations are inconsistent and no noticeable improvement can be seen from higher resolution in Fig. 9. This is likely because changes in physical parameters are small and do not change surface evaporation much.

Another informative point to analyze is Val4, located in the center of the valley at 1109 m, corresponding to the automatic weather station of 'Das'. In D3 the pixel elevation is 1081 m and in D4 is 1090 m. Therefore there is minimal error in either domain. In addition, the LU category is the same for each domain; 'Cropland/Woodland Mosaic' (category 6) for CLC and 'Dryland Cropland and Pasture' (category 2) for USGS. Physical parameters are very similar between these 2 categories. Looking at the observed temperature curve (Fig. 10) strong night-time inversions are recorded. Neither of the two experiments is able to capture this strong inversion during any night and all models overestimate the minimum temperatures. This is a known problem for surface layer schemes using MOST, which maintain a minimum value for the friction velocity in order to prevent the heat flux from being zero under very stable conditions (runaway cooling) and therefore they are not able to reproduce the very stable conditions at the surface (Bravo et al., 2008; Jiménez et al., 2012). None of the simulations reaches the low observed minimum temperatures but the higher resolution simulation (green and yellow lines in Fig. 10a) is able to get somewhat closer. Inversions are strongly controlled by topography, especially in complex terrain, so this is not surprising. Maximum temperature values are not improved in the finer simulations because they are related to other factors such as land use. In addition, the minimum temperature in the USGS-500m simulation (green line in Fig. 10) is closer to the observations than the CLC-500m, which may be a consequence of the drier representation of the valley (not at the exact point of the station where they are very similar) in the USGS database (see Sect. 3.1), with higher soil moisture potential availability and lower thermal inertia (Fig. 4c,e), which means lower evaporation and, thus, stronger air cooling.

Relative humidity (RH) is directly related to air temperature because cold air can hold less vapour than warm air. A better performance of the model in simulating minimum temperatures often occurs when RH is higher at night, which is the case for the USGS-500m simulation. On many nights the air chills until saturation of water vapour is obtained and fog forms, then slowing down the fall in air temperature. Thus the minimum temperature observed is controlled by when 100% RH is obtained in such cases. For specific humidity (q), the results are less clear, there are some periods like the nights of 2nd/3rd and 9th/10th August where the higher resolution gives better results, but this is not universal and further analysis must be done.

In summary, the Val4 example shows that the improvement of the model through increasing horizontal resolution is often not directly related to reducing the local elevation error, but moreover a consequence of a better representation of the

surrounding landscape and the whole domain. Nocturnal temperature inversions depends on terrain morphology, which determines where cool pools are formed (Miró et al., 2010; Miró et al., 2017).

3.3. Statistical parameters evaluation of the model.

Beyond the above case studies, a more extensive analysis involves calculation of statistical parameters to describe model performance in estimating observed temperature (T), relative humidity (RH) and specific humidity (q) at 2 m. We calculate mean bias (MB) defined as $MB = \frac{1}{N}\sum_i (P_i - O_i)$, the root mean squared error (RMSE), $RMSE = \sqrt{\frac{\sum_i (P_i - O_i)^2}{N}}$ and the correlation coefficient (r), $r = \frac{\sum_i (P_i - \overline{P})(O_i - \overline{O})}{\sqrt{\sum_i (P_i - \overline{P})^2 \sum_i (O_i - \overline{O})^2}}$, where P_i is the model forecast, O_i the observed value and N is the total number of values. Mean bias gives an idea of the mean deviation of the model. If MB > 0 the model tends to overestimate and if MB < 0 it underestimates values observed; The RMSE is a measure of the model absolute deviation and r measures the correlation between observations and predictions. The model correlates with observations if the value is near 1 ($r \sim 1$).

Table3 shows the statistics averaged for 9 simulated days and over all stations (39 data sensors). Because specific humidity is not directly recorded at the sensors, it was calculated using model pressure at each point, which can introduce small errors.

USGS experiments tend to underestimate temperature more than CLC. Although biases are small when averaging over all locations, there is a noticeable improvement in the CLC experiments, with the MB closer to zero. CLC-500m is the simulation that obtains the best RMSE value (bold highlighted in Table 3) and even CLC-1km has a lower RMSE than USGS-500m, meaning that LU has a more powerful effect on 2 m air temperature compared to improving horizontal resolution in this case. Based on correlation, CLC-1km is surprisingly the best experiment, maybe because the higher resolution CLC-500m generates more short time scale variation which does not always follow the observations.

Regarding relative humidity all experiments show us a clear underestimation of this surface value. This might not all be due to model inadequacies because sensors are installed in trees which would suffer increased RH in comparison with mean conditions over the whole grid cell, due to local transpiration. Despite this, the MB of the two USGS experiments, averaged over all the data transect points at 'La Cerdanya' valley, presents better values than the two CLC experiments. In contrast, the best RMSE value is obtained for the CLC-500m simulation, followed by USGS-500m, meaning that overall RH is more sensitive to horizontal resolution than to LU. CLC-500m also presents the highest correlation, with a clear improvement over USGS-1km.

Because RH is highly dependent on temperature, it is convenient to use q, which does not depend directly on the temperature, in order to evaluate more

Table 3: Statistic parameters comparison of temperature (T), relative humidity (RH) and specific humidity (q) for the average values over all the observation sites and for the 4 different experiments. Values highlighted in bold correspond to the best statistic for each experiment.

		USGS		CLC	
Mag.	Param.	1 km	500m	1km	500m
T ($°C$)	MB	-0.375	-0.249	-0.161	**-0.067**
	RMSE	2.722	2.716	2.655	**2.599**
	corr.	0.857	0.887	**0.920**	0.887
RH (%)	MB	-3.678	**-3.352**	-5.201	-5.266
	RMSE	17.489	17.185	17.323	**17.034**
	corr.	0.489	0.558	0.584	**0.608**
q ($g\,kg^{-1}$)	MB	-1.059	**-0.976**	-1.201	-1.171
	RMSE	2.220	**2.104**	2.246	2.162
	corr.	0.472	0.574	**0.599**	0.588

precisely water vapour processes in the model. On average, q is better resolved in USGS-500m experiment, except for the correlation, for which CLC-1km is the best simulation (Table 3). This confirms that CLC experiments tend to underestimate humidity more than USGS, which in general can be attributed to lower moisture availability in the CLC LU database (see Fig. 4) over the whole domain. This in turn leads to lower evaporation, although in the valley bottom it is the contrary. Higher correlations for the CLC experiments suggest that the geographical distribution is more accurate in the adapted LU database.

We also extend our analysis to include maximum and minimum temperatures (Table 4). In general maximum temperatures are underestimated in all model simulations, by an order of 2 or 3 degrees. CLC-500m is the one that performs best, with a reduction of more than $0.3°C$ in mean bias in comparison with USGS-1km. This improvement is broadly similar to that observed in model bias when temperature is evaluated at all times. However the improvement in RMSE for maximum temperature, which reduces by $0.3°C$, is much more than for temperature in general ($0.1°C$). We can also conclude that LU has a stronger impact compared with grid size which is less important. In contrast minimum temperatures get worse when the LU database is changed or the horizontal grid size decreased. Local improvement of minimum temperatures along the valley axis (section 3.2) is not achieved across the whole domain and appears to be compensated for by worse prediction skill at higher elevations. A more detailed

discussion can be found in section 3.5.

A general conclusion from the statistical evaluation is that CLC LU distribution obtains better results of near surface air temperature and that increasing horizontal resolution of the model usually improves the forecast at least at those points where we have observations. Regarding RH, the results are less clear, as we have already seen in particular case studies (Sect. 3.1). Correlations do appear to improve for both air temperature and humidity in the CLC simulations.

Table 4: Statistic parameters comparison of daily maximum temperature (T_{max}) and minimum temperature (T_{min}) for the averaged values over all the observation sites and for the 4 different experiments. Values highlighted in bold correspond to the best statistic for each experiment.

		USGS		CLC	
Mag.	Param.	1 km	500m	1km	500m
T max ($°C$)	MB	-2.746	-2.648	-2.476	**-2.439**
	RMSE	3.853	3.800	3.660	**3.561**
T min ($°C$)	MB	**1.325**	1.461	1.488	1.652
	RMSE	**2.068**	2.120	2.156	2.147

3.4. Spatial patterns of the model bias and RMSE.

We also examined the spatial patterns of bias and RMSE in the four simulations in more detail. Table 5 lists the explanatory factors which were significant (at p=0.2) in a regression model predicting mean bias or RMSE from topographic variables. The derivations of the topographic variables of aspect and exposure (relative elevation) are explained in Appendix A. Northerly aspect is a proxy for radiation input with high/low numbers representing shady/sunlit slopes. Westerly aspect is a proxy for exposure to the prevailing winds with high/low numbers representing exposed/sheltered slopes.

There is very little difference between the four simulations (differing resolutions and USGS vs CLC). Thus we do not discuss these model differences in detail. In most cases the same variables appear in most regression models for each predictand. On average about two-thirds of the temperature bias can be explained by a combination elevation, model elevation error, and north/south aspect. Positive bias (simulation too warm) is common at low elevations and also when model elevation is lower than actual elevation (as expected) and on northerly aspects with limited radiation input. This suggests that the model is not good at including cold air drainage effects in valley bottoms, and that shading effects are slightly under-simulated. Temperature RMSE shows a more complex pattern

Table 5: Results of stepwise regression to predict the MB and RMSE for temperature (T), specific humidity (q) and relative humidity (RH). p-values<0.2 are shown for each predictor. Colors indicate the sign of the regression coefficient for each parameter (blue/red: negative/positive). The last column shows the R^2 value for the multiple regression, only considering the predictors with $p > 0.2$.

Predictand	Station Elevation (m)	Model Elevation Error (m)	N-Aspect	W-Aspect	Exposure (360 m radius)	Exposure (180 m radius)	**Total R^2**
Temperature MB							
USGS 1 km	0.005	0.000	0.073				0.659
USGS 500 m	0.000	0.000					0.549
CLC 1 km	0.003	0.000	0.102				0.662
CLC 500 m	0.000	0.001	0.071				0.605
Temperature RMSE							
USGS 1 km	0.150	0.037					0.130
USGS 500 m	0.010	0.001		0.072	0.174	0.187	0.395
CLC 1 km							
CLC 500 m	0.184	0.033	0.036	0.070	0.072	0.067	0.469
Q Bias							
USGS 1 km	0.000	0.079					0.820
USGS 500 m	0.000	0.148					0.738
CLC 1 km	0.000	0.063				0.198	0.846
CLC 500 m	0.000	0.156				0.113	0.777
Q RMSE							
USGS 1 km	0.000				0.131	0.163	0.676
USGS 500 m	0.000						0.577
CLC 1 km	0.000				0.142	0.158	0.718
CLC 500 m	0.000						0.635
RH Bias							
USGS 1 km	0.000		0.010				0.585
USGS 500 m	0.000		0.058				0.434
CLC 1 km	0.000		0.017				0.601
CLC 500 m	0.000		0.107	0.086			0.573
RH RMSE							
USGS 1 km	0.128	0.053	0.121	0.085			0.339
USGS 500 m		0.048		0.021			0.207
CLC 1 km	0.027	0.092	0.073	0.125			0.395
CLC 500 m	0.017			0.112			0.247

with exposure values included in the regression models at 500 m, but this is influenced by multi-collinearity effects and the r2 is low.

Models for the specific humidity patterns are clearer, with 70- 90% of the bias explained by station elevation and model elevation error alone. All four model simulations overestimate humidity at high elevations, and underestimate it in valley bottoms, possibly again a result of the lack of trapping of moisture below inversions in simulations compared with reality. Most sensors are installed in vegetation so there may also be an additional observational effect here which is compounded in sheltered valley bottoms with reduced ventilation. q RMSE is also well explained, mostly by elevation, but also partly by exposure. RMSE is larger at low elevations because the stagnant air at lower elevation sites will show rapid changes in humidity as a result of variable evapotranspiration, which is not simulated well by the model.

Finally the results for relative humidity are slightly more complex, because they depend both on specific humidity and air temperature. Nevertheless around 40-60% of the bias can still be explained by topographic variables, in this case elevation and northerly aspect. High elevations and southerly aspects are too humid in the model, while valley bottoms and northerly aspects are too dry. Vegetation contrasts between more xeric south facing environments, and wooded north-facing slopes which are not simulated by the model probably account for much of the bias. RMSE is less clearly explained with lower r^2.

Overall the signs of the variables are coherent, do not change between model simulations, and what would be expected due to additional microclimate effects. We also re-ran the regression models with north-westerly (NW) and north-easterly (NE) aspects (parallel and perpendicular to the valley orientation) and the differences in results were negligible (not shown). The NE/NW aspects behaved like N/W aspects respectively.

3.5. Statistical analysis of horizontal resolution and LU database contributions to model performance.

In order to differentiate when improvement in the LU database is responsible for better model performance from when it is due to topographic resolution, we average results from domains with different resolutions (1 km/ 500 m) and simulations with different LU databases (CLC/USGS). Taking into account that RMSE is a measure of model dispersion, we can define the average RMSE of a given variable (temperature or humidity) of the domains with similar LU as:

$$RMSE_{CLC} = \frac{RMSE(CLC, 500m) + RMSE(CLC, 1km)}{2} \quad (1)$$

$$RMSE_{USGS} = \frac{RMSE(USGS, 500m) + RMSE(USGS, 1km)}{2} \quad (2)$$

Similarly, we averaged the RMSE data of equal horizontal resolution, but different LU:

$$RMSE_{1km} = \frac{RMSE(CLC, 1km) + RMSE(USGS, 1km)}{2} \quad (3)$$

$$RMSE_{500m} = \frac{RSME(CLC, 500m) + RMSE(USGS, 500m)}{2} \quad (4)$$

Then we compute the improvements solely due to changes in LU and those due to changes in horizontal resolution (HR) as:

$$D_{LU} = RMSE_{USGS} - RMSE_{CLC} \quad (5)$$

$$D_{HR} = RMSE_{1km} - RMSE_{500m} \quad (6)$$

where D_{LU} is the reduction in RMSE due to LU change and D_{HR} is the reduction due to horizontal resolution. The sign of D_{LU} tells us when the CLC database improves the model ($D_{LU} > 0$), or when it performs worse ($D_{LU} < 0$). DLU is independent of resolution because we have averaged these effects. In the same way, D_{HR} shows us when horizontal resolution is improving our results ($D_{HR} > 0$) or when it is worsening them ($D_{HR} < 0$). Fig. 11 shows the temporal evolution of D_{LU} and D_{HR} for a) temperature and b) specific humidity over the 9 day period, averaged from all transect sensors every 30 minutes.

For temperature there is a clear diurnal cycle in D_{LU} (blue line in Fig. 11a), which has a maximum value at midday and a minimum during nighttime hours. Thus, changing from USGS LU database to CLC database has the most success during the day time and is less influential at night. On the other hand, DHR fluctuates more rapidly and has more abrupt changes. Changing horizontal resolution from 1-km to 500-m usually has positive effects in the daytime but some negative impacts at night, most frequently between 22UTC and 5UTC. On average a finer resolution does not improve results during the night, although at specific local points, for instance within the valley, it helps to resolve better the cooling processes and local circulations (as seen in Sect. 3.2).

For specific humidity contrasting patterns are shown. On one hand, D_{LU} remains near zero and is often slightly negative so, on average, worse results are given for specific humidity through changing the LU database from USGS to CLC (Fig. 11b). The spatial heterogeneity of LU categories in both classifications, and the lack of a consistent change between the two makes it difficult to reach a clear conclusion. For instance, as mentioned above, CLC tends towards more moist physical parameters (larger soil moisture and higher thermal inertia) within the valley, but is drier in the rest of the domain, specially at high elevations. This leads to opposing influences in terms of the variation of specific humidity. On the other hand, the D_{HR} value is often substantially positive, meaning that the 500-m horizontal resolution simulation performs consistently better than the 1-km one in terms of humidity.

We also compute the mean daily cycle of the RMSE differences over the 9 simulated days (Fig. 12), from 00UTC to 2330UTC. Because in general terms

the 9 days are very similar, we expect to obtain a daily pattern. In this case statistical significance is assessed using the t-test. P-values are calculated for each 30 minute window of the daily cycle (n=48) and values less than 0.1 were deemed significant.

For temperature the D_{LU} has a clear daily cycle. In general, CLC performs better than USGS ($D_{LU} > 0$) during the day, whereas in general effects are statistically insignificant ($D_{LU} = 0$) during the night (Fig. 12a). During the day, when incoming radiation is strong, the radiative terms in the energy budget equation play an important role controlling the surface fluxes of latent and sensible heat, which strongly control the simulated temperature. A better representation of physical parameters becomes more important. The D_{HR} curve shows a significant peak between 0600 and 0900 UTC, which may be because the 500-m simulations capture better the mesoscale/microscale processes that lead to the formation of temperature inversions which are at their strongest in the hours around sunrise. A second statistically significant broader peak is observed at sunset. This means that improved representation of topography in the 500-m simulations leads to a better performance during morning and evening hours. In addition, the two peaks in improvement occur when the sun is at a low elevation and therefore aspect effects are at a maximum, along with the influence of topographic shading. Increased horizontal resolution improves the representation of both of these effects, whereas in the middle of the day they are less critical. Unexpectedly, during most of the night the 500-m simulations do not improve upon the 1-km simulations. We speculate that the model requires a certain time to fully develop the stable layer during nights.

For specific humidity the results are less clear. D_{LU} is near zero, and can be slightly negative during morning hours, which is due to the lack of soil moisture availability in the CLC dataset averaged over the whole domain. This causes less evaporation of the liquid water in the soil layers of the model (Fig. 12b). In contrast, the 500-m simulations (D_{HR}) obtain better results for most of the 24 hr period, but especially during the night. This was especially true on stable nights, 2-6 Aug and 9-10 Aug (see Fig. 11). Although these results are not usually significant at the 90% confidence level, they reach occasionally 75% (p=0.25), which suggests that after two or three more days of clear skies and stable conditions results could easily become significant. This improvement is a consequence of the better representation of relief, which together with the smaller grid size contributes to a more precise performance of the model during night, when the predominant winds are weak, horizontal and caused by local complex terrain processes. Also night-time q is dependent on cold air ponding and trapping of air in topographically confined inversions which can be simulated more effectively using a higher grid resolution.

4. Summary and Conclusions

In this study, we have adapted the CLC (Corine Land Cover) 2006 database to be used in WRF, a three-dimensional mesoscale atmospheric model, using a reclassifying procedure based on previous research. This permitted an increase in the horizontal resolution of the LU database and an update to better represent land surface characteristics that may have changed in recent years and that USGS (U.S. Geological Survey) may not include.

We have evaluated the effects of land use database and horizontal resolution on the modelling of near surface atmospheric variables, including 2 m air temperature and humidity. We use two simulations with different land use datasets (USGS versus CLC) and compare model magnitudes with observed data across 'La Cerdanya'. We also study the effect of using different horizontal grid resolutions, comparing results from two different nested domains, D3 with 1km horizontal resolution and D4 with 500m.

At each grid point LU category changes modify the local radiative and energy balance, thus impacting surface fluxes of latent and sensible heat. Furthermore, the skin temperature (do not confuse with the 2 m temperature) also experienced significant change at those points where the LU category changed significantly. The LU impact on the surface sensible and latent heat has been demonstrated through particular case studies (section 3.1). Local air temperature and specific humidity were shown to be sensitive to the type of land use database at these particular sites.

Horizontal resolution becomes important when simulating atmospheric behavior in complex terrain. A clear improvement of the surface temperature forecast at valley bottom and hill slope points was shown when using a finer grid spacing of 500 m. In the first case (Cad2), a reduced model elevation error at the point reduced the temperature bias. In the second case, there was an improvement in simulating minimum temperatures at the bottom of the valley during stable night conditions, meaning that temperature inversions were captured more realistically. This also improved the prediction of relative humidity.

Statistical validation across all 39 observation points showed that temperature predictions were more accurate using CLC compared with USGS. However, this was not the case for specific humidity, where USGS performs slightly better on average. Regarding the effect of different horizontal resolutions, specific humidity RMSE and bias are reduced when using finer resolution, which is possibly due to better modeled surface flow. The 500 m resolution also shows an improvement of temperature forecasts due to reduced elevation errors and better simulation of cool pool formation). The forecast of maximum temperatures is primarily improved through change in LU using CLC. In contrast, minimum temperatures show no improvement or a worsening in general, which contrast with the local effect at the bottom of the valley seen in the Val4 transect point.

We also examined the spatial patterns of MB and RMSE in the four simulations

using a stepwise multi-regression model. On average about two-thirds of the temperature bias can be explained by a combination of elevation, model elevation error, and aspect. In the case of specific humidity 70-90% of the bias is explained by station elevation and model elevation error alone. The model simulations overestimate humidity at high elevations, and underestimate it in valley bottoms. The q RMSE is also well explained, mostly by elevation, but also partly by exposure.

Finally we averaged the daily cycle of the RMSE values for different simulations, enabling us to compare benefits of modernising LU versus increasing horizontal resolution. In general, LU change reduces temperature errors during daylight hours and has a small negative effect during night. A finer grid size of 500 m causes a noticeable improvement of temperature at sunrise, when temperature inversions form in the valley, and before sunset, when local microclimate effects are enhanced. During the night there is no or minimal improvement. In contrast, specific humidity shows contrasting results in that CLC increases model deviation during morning but does not have general effect during the rest of the day, due to the lack of moisture availability. In addition, a finer grid size caused a noticeable reduction of the averaged model absolute error, which we speculate could be due to better resolution of mesoscale valley flows.

In summary using both an appropriate LU dataset and an appropriate horizontal grid size is important to improve numerical weather simulations and enhance weather forecasting capability in areas of complex terrain like the Pyrenees.

Appendix A. Topographic variables

The topographic variables used in section 3.4 and their calculation are described here for convenience:

First we define the slope of the terrain at a 2D grid point (i, j) (where i and j increase to north and to east respectively) as:

$$N_{slope}(i,j) = Z(i,j-1) - Z(i,j+1) \tag{A.1}$$
$$W_{slope}(i,j) = Z(i+1,j) - Z(i-1,j) \tag{A.2}$$

Note a high positive number is (N/W) facing and high negative number is (S/E) facing respectively.

Then we define the aspect at this point as:

$$N_{asp} = atan(Nslope/d) \tag{A.3}$$
$$W_{asp} = atan(Nslope/d) \tag{A.4}$$

where d is the horizontal distance between the two cells either side of the nearest sensor point. In our case, we use a DEM (Digital Elevation Model) with a

resolution of 90 m, then $d = 180m$. Note that the aspect in (N/W) direction is an angle (positive is north/west and negative is south/east facing respectively).

Exposure variables represent whether the sensor grid cell is higher or lower than its surroundings. To calculate it, we used the following equation:

$$Z_{ring1}(i,j) = \frac{\begin{array}{c} Z(i+1,j-1) + Z(i+1,j) + Z(i+1,j+1) + Z(i,j-1) \\ + Z(i,j+1) + Z(i-1,j-1) + Z(i-1,j) + Z(i-1,j+1) \end{array}}{8} \tag{A.5}$$

This is the mean elevation of the surrounding 8 grid cells, in the DEM 90 m away from the station point. Exposure is calculated by:

$$exp_l(i,j) = Z(i,j) - Z_{ring1} \tag{A.6}$$

A positive value is higher than its surroundings (hence the name exposure index), this would mean a ridge top or exposed location to the free atmosphere. A negative number would mean a topographic bowl or concavity. Note that exp_l is on small scale (radius of 90 m) hence l for local.

We can also do the same thing for the second ring:4

$$Z_{ring2}(i,j) = \frac{\begin{array}{c} Z(i+2,j-2) + Z(i+2,j) + Z(i+2,j+2) + Z(i,j-2) \\ + Z(i,j+2) + Z(i-2,j-2) + Z(i-2,j) + Z(i-2,j+2) \end{array}}{8} \tag{A.7}$$

and

$$exp_m(i,j) = Z(i,j) - Z_{ring2} \tag{A.8}$$

Of course this is at a moderate scale of 180 m radius (hence m for medium).

Acknowledgements

This research was supported by the Spanish Government through >> projects CGL2012-37416-C04-04 and CGL2015-65627-C3-2-R

Sensors for field observations of temperature and relative humidity were provided by the University of Portsmouth Research Development.

References

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data* volume 964. US Government Printing Office.

Arnold, D., Morton, D., Schicker, I., Seibert, P., Rotach, M., Horvath, K., Dudhia, J., Satomura, T., Müller, M., Zängl, G. et al. (2012). High resolution modelling in complex terrain. In *Report on the HiRCoT 2012 Workshop, Vienna* (pp. 21–23).

Bravo, M., Mira, T., Soler, M., & Cuxart, J. (2008). Intercomparison and evaluation of MM5 and meso-nh mesoscale models in the stable boundary layer. *Boundary-Layer Meteorol*, *128*, 77–101.

Büttner, G., & Kosztra, B. (2007). Clc2006 technical guidelines. *European Environment Agency, Technical Report*, .

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface-hydrology model with the penn state-ncar mm5 modeling system. part i: Model implementation and sensitivity. *Monthly Weather Review*, *129*, 569–585.

Cheng, F.-Y., & Byun, D. W. (2008). Application of high resolution land use and land cover data for atmospheric modeling in the houston–galveston metropolitan area, part i: Meteorological simulation results. *Atmospheric Environment*, *42*, 7795–7811.

Cheng, F.-Y., Hsu, Y.-C., Lin, P.-L., & Lin, T.-H. (2013). Investigation of the effects of different land use and land cover patterns on mesoscale meteorological simulations in the taiwan area. *Journal of Applied Meteorology and Climatology*, *52*, 570–587.

Civerolo, K., Hogrefe, C., Lynn, B., Rosenthal, J., Ku, J.-Y., Solecki, W., Cox, J., Small, C., Rosenzweig, C., Goldberg, R. et al. (2007). Estimating the effects of increased urbanization on surface meteorology and ozone concentrations in the new york city metropolitan region. *Atmospheric environment*, *41*, 1803–1818.

Cuxart, J. (2015). When can a high-resolution simulation over complex terrain be called les? *Frontiers in Earth Science*, *3*, 87.

Cuxart, J., Wrenger, B., Martínez-Villagrasa, D., Reuder, J., Jonassen, M. O., Jiménez, M. A., Lothon, M., Lohou, F., Hartogensis, O., Dünnermann, J., Conangla, L., & Garai, A. (2016). Estimation of the advection effects induced by surface heterogeneities in the surface energy budget. *Atmospheric Chemistry and Physics Discussions*, *2016*, 1–16. URL: `http://www.atmos-chem-phys-discuss.net/acp-2015-1051/`. doi:10.5194/acp-2015-1051.

De Meij, A., & Vinuesa, J. (2014). Impact of srtm and corine land cover data on meteorological parameters using wrf. *Atmospheric Research*, *143*, 351–370.

Dudhia, J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, *46*, 3077–3107.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L. et al. (2007). The shuttle radar topography mission. *Reviews of geophysics*, *45*.

Giovannini, L., Antonacci, G., Zardi, D., Laiti, L., & Panziera, L. (2014). Sensitivity of simulated wind speed to spatial resolution over complex terrain. *Energy Procedia*, *59*, 323–329.

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

Horvath, K., Koracin, D., Vellore, R., Jiang, J., & Belu, R. (2012). Sub-kilometer dynamical downscaling of near-surface winds in complex terrain using wrf and mm5 mesoscale models. *Journal of Geophysical Research: Atmospheres (1984–2012)*, *117*.

Jiménez, M. A., & Cuxart, J. (2014). A study of the nocturnal flows generated in the north side of the pyrénées. *Atmospheric Research*, *145*, 244–254.

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., & García-Bustamante, E. (2012). A revised scheme for the wrf surface layer formulation. *Monthly Weather Review*, *140*, 898–918.

Kain, J. S. (2004). The kain-fritsch convective parameterization: an update. *Journal of Applied Meteorology*, *43*, 170–181.

Lam, J., Lau, A., & Fung, J. (2006). Application of refined land-use categories for high resolution mesoscale atmospheric modelling. *Boundary-layer meteorology*, *119*, 263–288.

Miró, J. R., Pagès, M., & Kossman, M. (2010). Cold-air pool detection tools in the pyrenees valleys. In *14th Conference on Mountain Meteorology*.

Miró, J. R., Peña, J. C., Pepin, N., Sairouni, A., & Aran, M. (2017). Key features of cold-air pool episodes in the northeast of the iberian peninsula (cerdanya, eastern pyrenees). *International Journal of Climatology*, (pp. n/a–n/a). URL: `http://dx.doi.org/10.1002/joc.5236`. doi:10.1002/joc.5236.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres*, *102*, 16663–16682.

Monin, A., & Obukhov, A. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. *Contrib. Geophys. Inst. Acad. Sci. USSR*, *151*, 163–187.

Noh, Y., Cheon, W., Hong, S., & Raasch, S. (2003). Improvement of the k-profile model for the planetary boundary layer based on large eddy simulation data. *Boundary-layer meteorology*, *107*, 401–427.

Obukhov, A. (1946). Turbulence in the atmosphere with inhomogeneous temperature. *Trudy geofiz. inst. AN SSSR*, *1*.

Pagès, M., Pepin, N., & Miró, J. R. (2016). Measurement and modelling of temperature cold pools in cerdanya. *Meteorological Applications*, . URL: `http://eprints.port.ac.uk/22646/`.

Pagès, M., & Miró, J. R. (2010). Determining temperature lapse rates over mountain slopes using vertically weighted regression: a case study from the pyrenees. *Meteorological Applications*, *17*, 53–63. URL: `http://dx.doi.org/10.1002/met.160`. doi:10.1002/met.160.

Pepin, N., & Kidd, D. (2006). Spatial temperature variation in the eastern pyrenees. *Weather*, *61*, 300–310.

Pineda, N., Jorba, O., Jorge, J., & Baldasano, J. (2004). Using noaa avhrr and spot vgt data to estimate surface parameters: application to a mesoscale meteorological model. *International Journal of Remote Sensing*, *25*, 129–143.

Prabha, T. V., Hoogenboom, G., & Smirnova, T. G. (2011). Role of land surface parameterizations on modeling cold-pooling events and low-level jets. *Atmospheric Research*, *99*, 147–161.

Santos-Alamillos, F., Pozo-Vázquez, D., Ruiz-Arias, J., & 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.

Seaman, N. L., Gaudet, B. J., Stauffer, D. R., Mahrt, L., Richardson, S. J., Zielonka, J. R., & Wyngaard, J. C. (2012). Numerical prediction of submesoscale flow in the nocturnal stable boundary layer over complex terrain. *Mon Weather Rev*, *140*.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., & Powers, J. G. (2008). *A description of the advanced research WRF version 3*. Technical Report DTIC Document.

Thompson, G., Rasmussen, R. M., & Manning, K. (2004). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. part i: Description and sensitivity analysis. *Monthly Weather Review*, *132*, 519–542.

Udina, M., Soler, M. R., & Sol, O. (2017). A modeling study of a trapped lee-wave event over the pyrénées. *Monthly Weather Review*, *145*, 75–96.

Whiteman, C. D., Hubbe, J. M., & Shaw, W. J. (2000). Evaluation of an inexpensive temperature datalogger for meteorological applications. *Journal of Atmospheric and Oceanic Technology*, *17*, 77–81.

Wyngaard, J. C. (2010). *Turbulence in the Atmosphere*. Cambridge University Press.

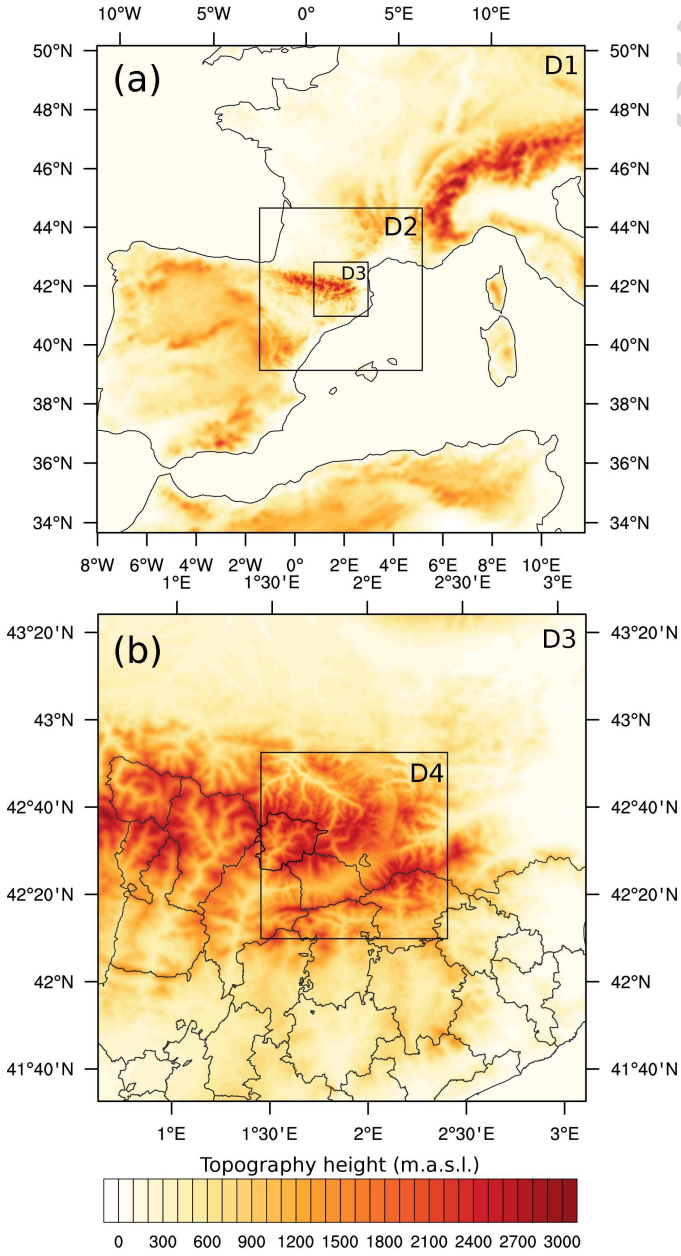

Figure 1: Topography of the studied area. (a) Nested model domains D1, D2 and D3: (b) Zoom at D3 and D4. Horizontal resolutions are 9, 3, 1 and 0.5 km respectively.

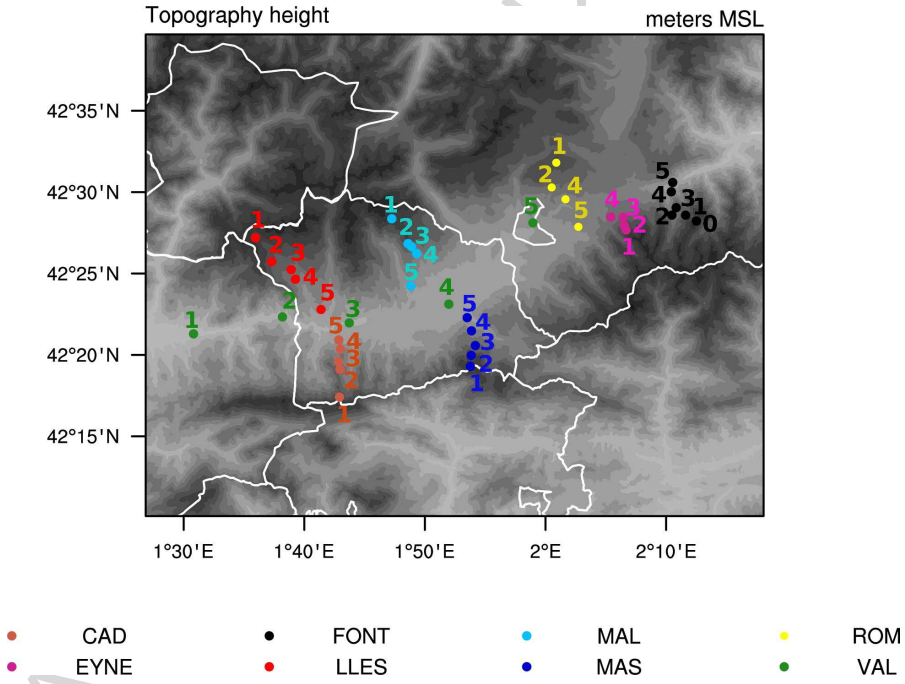

Figure 2: Localization of the 7 sensor transects at 'La Cerdanya' valley. The transects extend from above treeline (∼ 2400 m) down to the valley bottom (∼ 1000 m), except for VAL transect which goes along the valley axis, from SW to NE.

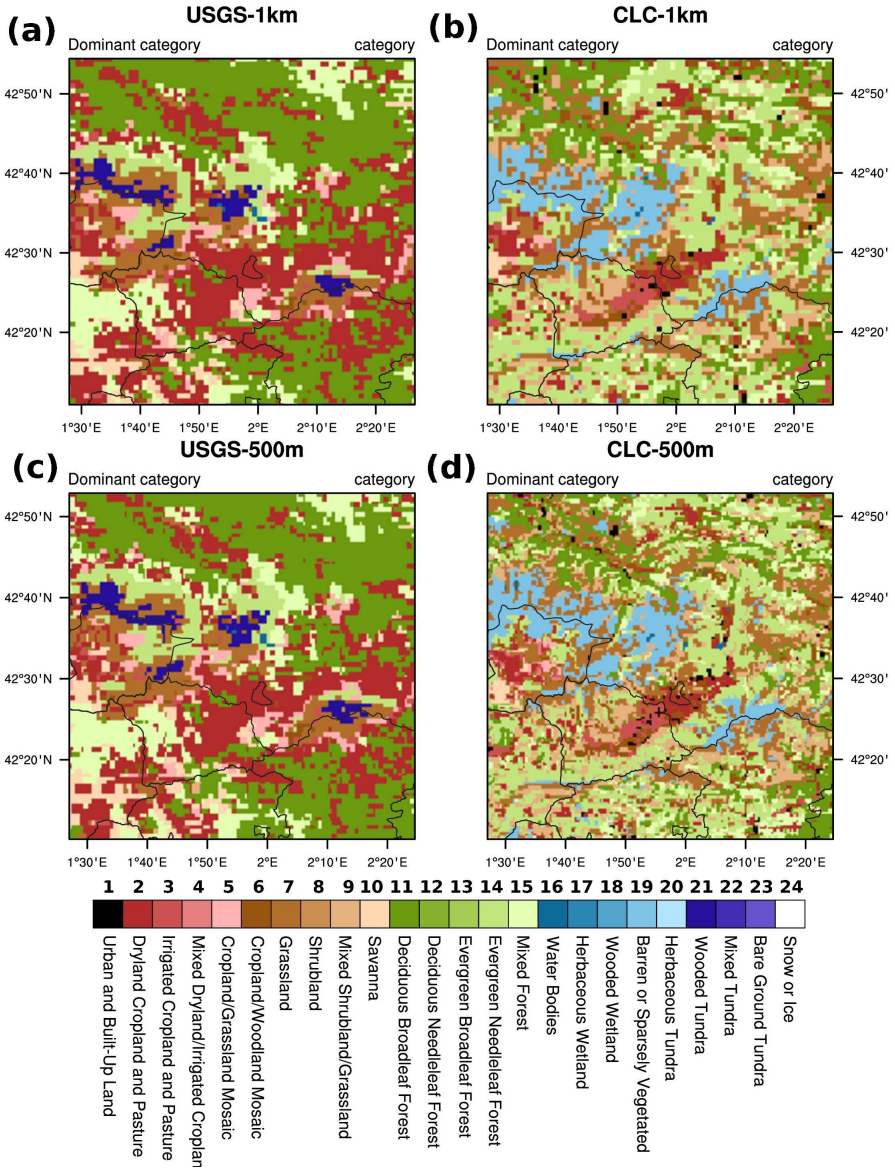

Figure 3: LU index at each grid point at 'La Cerdanya' valley (domain D4) in the two experiments: (a,c) USGS and (b,d) CLC. (a,b) 1x1 km and (c,d) 500x500 m horizontal resolution domains.

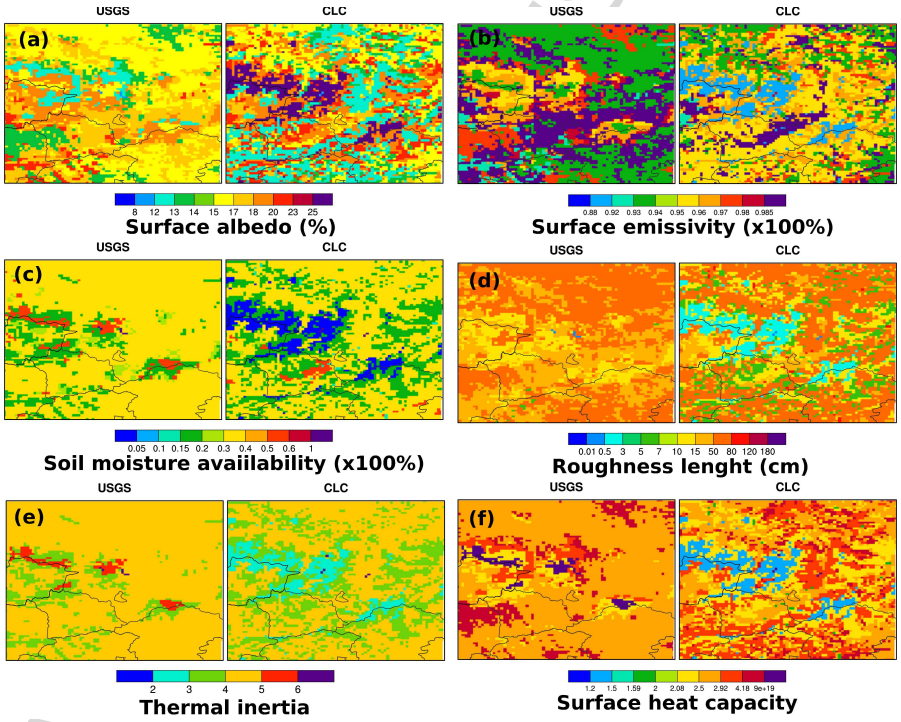

Figure 4: Spatial distribution over 'La Cerdanya' valley area of the 6 main physical parameters for the USGS and CLC LU datasets. (a) Surface albedo (α) in %, (b) surface emissivity (ϵ) in $\times$100%, (c) soil moisture availability (M) in $\times$100%, (d) roughness length (z_0) in cm, (e) thermal inertia (λ_T) in $4.184 \times 10^2 Jm^{-2}K^{-1}s^{-1/2}$ and (f) surface heat capacity (C) in $\times 10^5 Jm^{-3}K^{-1}$. Horizontal grid resolution is 1x1 km.

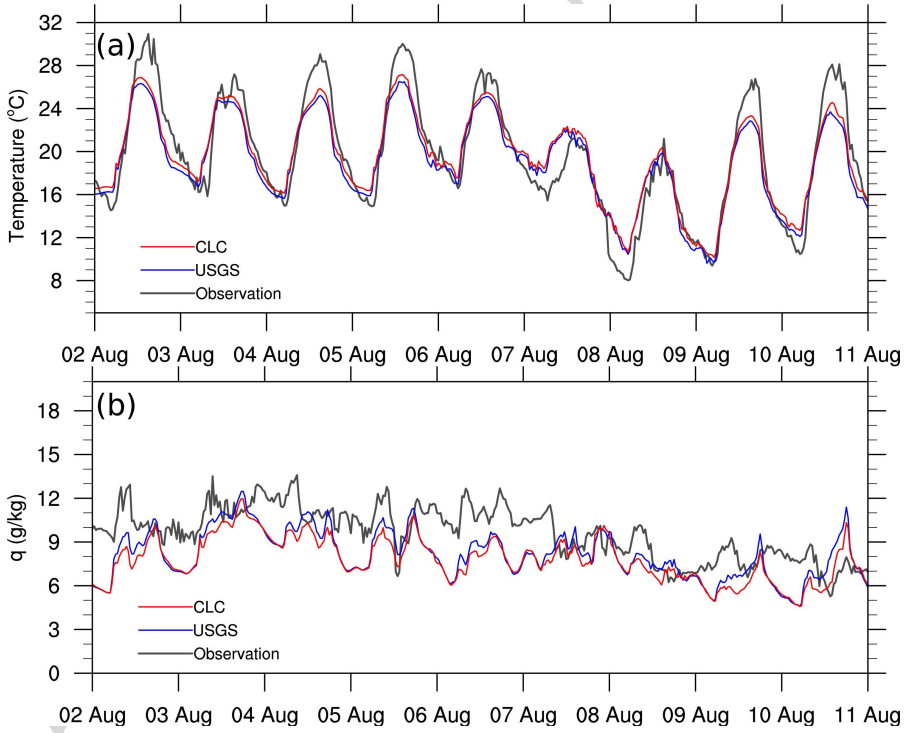

Figure 5: (a) Temperature and (b) specific humidity at 2 m at Mal5 transect point. Observation (grey), CLC (red) and USGS (blue) at the domain D3.

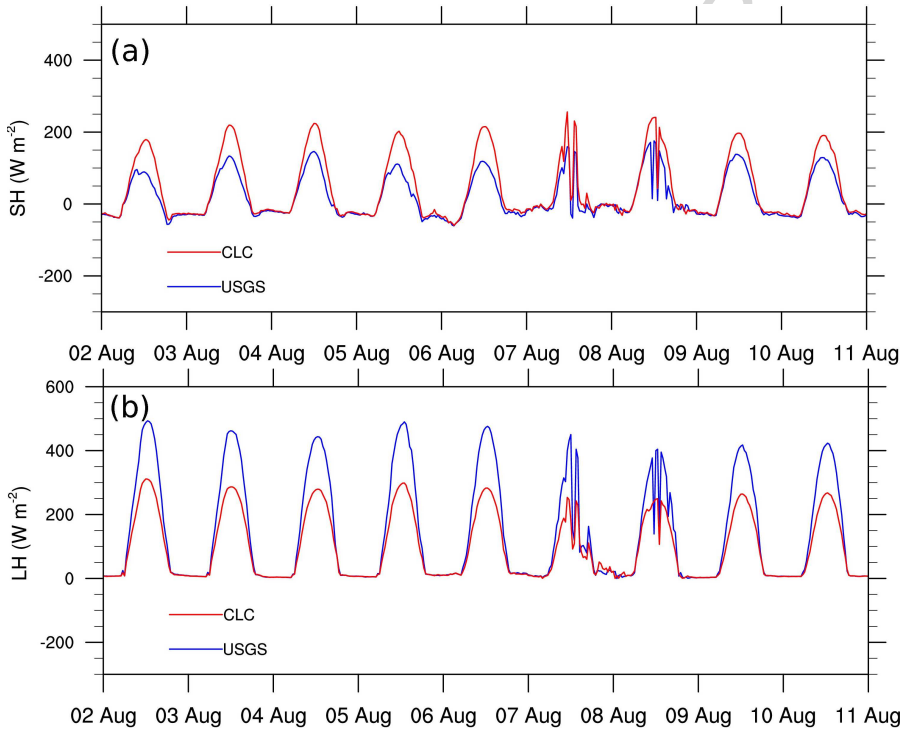

Figure 6: (a) Sensible heat flux (SH) and (b) latent heat flux (LH) at the surface at Mal5 transect point. CLC (red) and USGS (blue) at the domain D3.

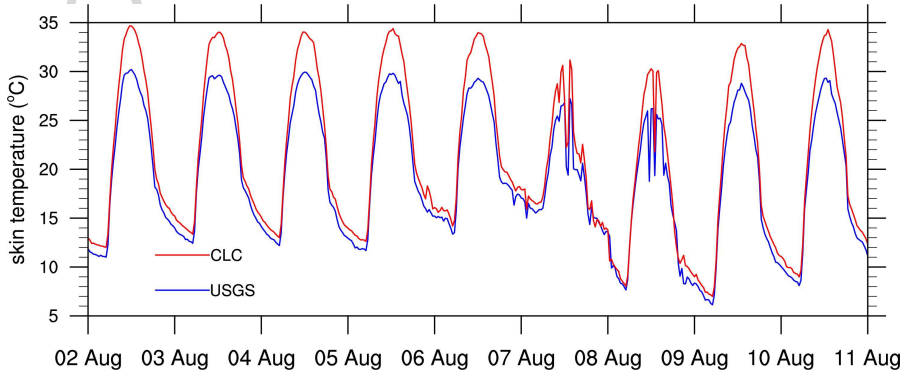

Figure 7: Skin temperature at Mal5 transect point. CLC (red) and USGS (blue) at the domain D3.

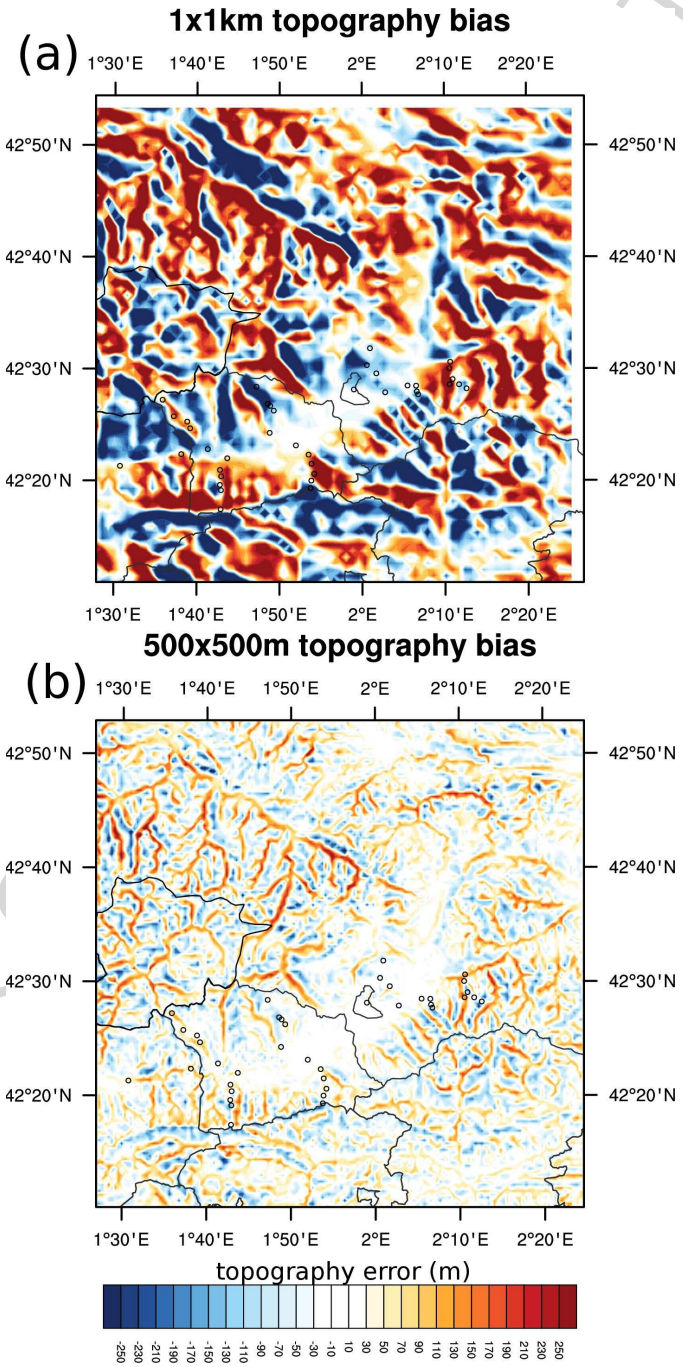

Figure 8: Elevation error (bias) in meters in 'La Cerdanya' valley for: (a) domain D3 (1km) and (b) domain D4 (500m smoothed).

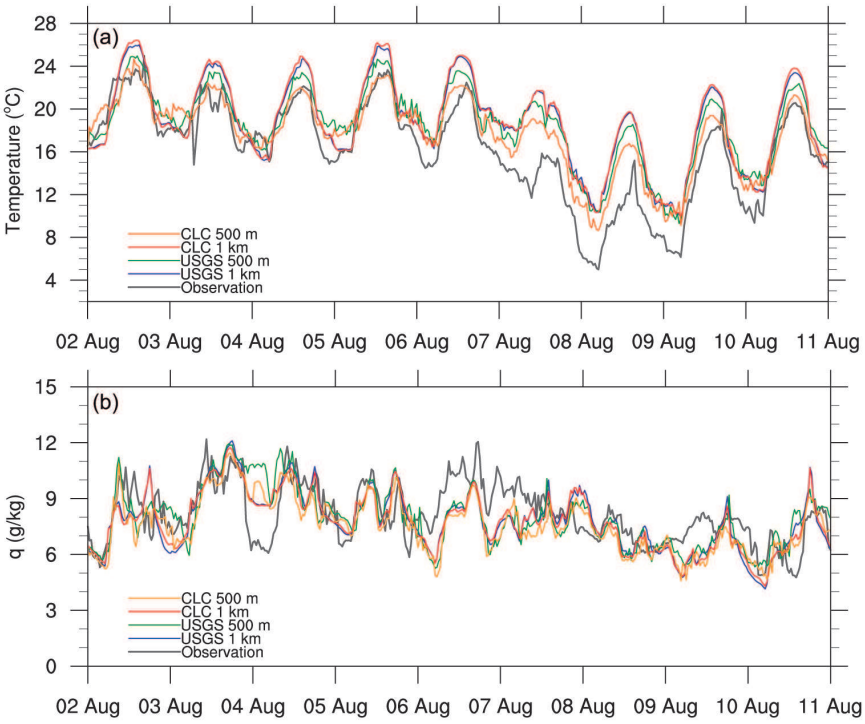

Figure 9: (a) Temperature and (b) specific humidity (q) at 2 m at Cad2 transect point. Observation (grey), CLC-1km (red), CLC-500m (yellow), USGS-1km (blue) and USGS-500m (green) .

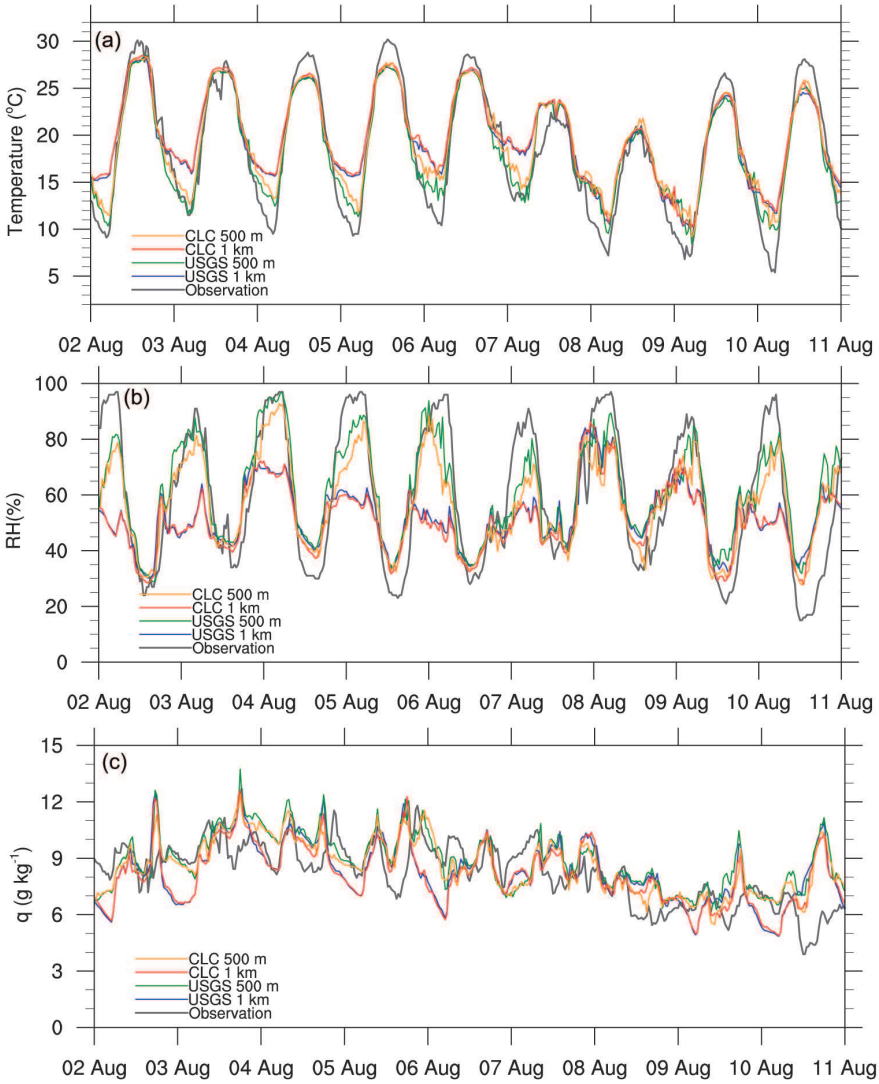

Figure 10: (a) Temperature, (b) relative humidity and (c) specific humidity at 2 m at Val4 transect point. Observation (grey), CLC-1km (red), CLC-500m (yellow), USGS-1km (blue) and USGS-500m (green)

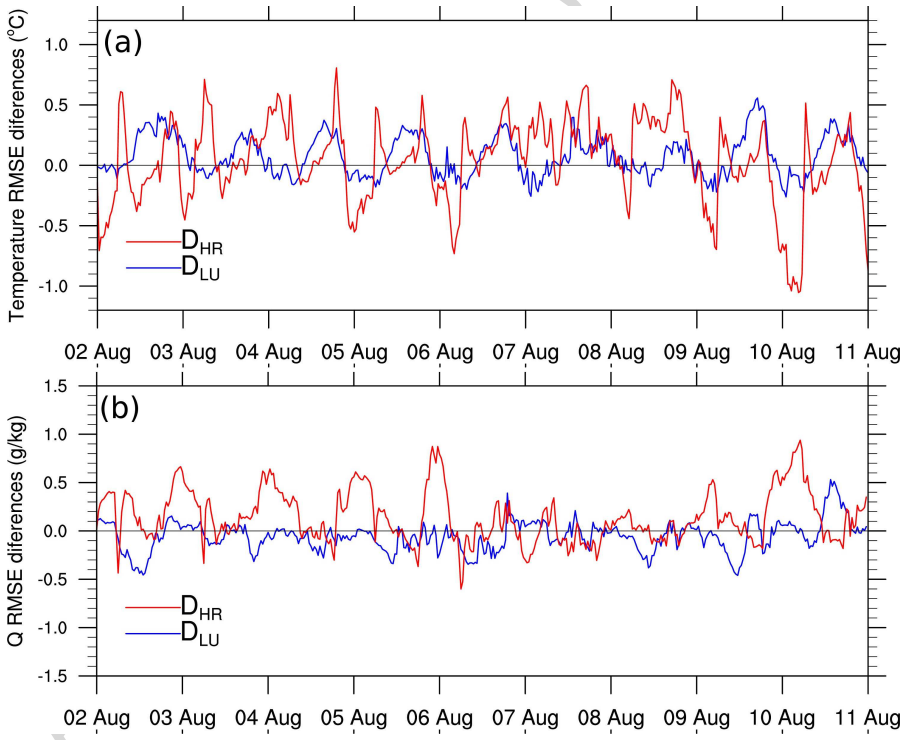

Figure 11: (a) RMSE temperature differences and (b) specific humidity differences: D_{HR} (red) and D_{LU} (blue). RMSE have been averaged every 30 min for all transect sensor points.

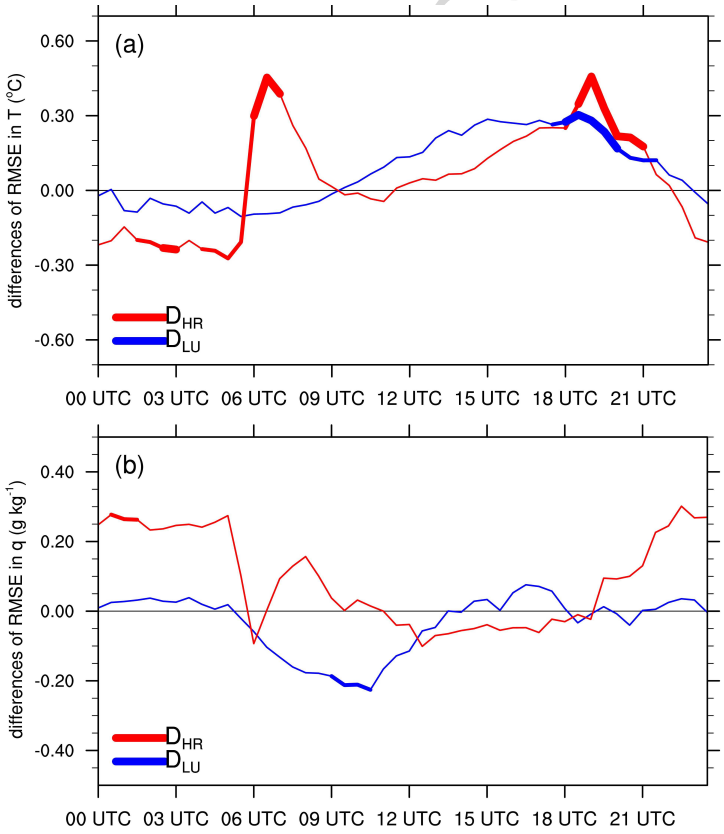

Figure 12: Daily averaged cycle of (a) RMSE temperature differences and (b) specific humidity differences: D_{HR} (red) and D_{LU} (blue). RMSE have been averaged every 30 min for all transect sensor points. Wide lines and medium wide lines indicate that the differences are 90% and 75% significant according to t-test respectively.

Hi [illegible]

- [illegible]
- [illegible]
- [illegible]
- T[illegible]
- T[illegible]