Persistent Wintertime Thermal Inversion in the City. The case of Madrid. Part I. Meteorology

Alberto Martilli^{1}, Beatriz Sanchez², Domingo Rasilla³, Gianluca Pappaccogli⁴, Fernando Allende⁵, Fernando Martin¹, Carlos Román-Cascón⁶, Carlos Yagüe⁶, Felipe Fernandez⁵.*

¹ CIEMAT, Department of Environment, Spain

² Department of Geography, National University of Singapore, Singapore

³ Departamento de Geografía, Urbanismo y OT, Universidad de Cantabria, Spain

⁴ National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Lecce, Italy

⁵ Departamento de Geografía, Universidad Autónoma de Madrid, Spain

⁶ Departamento de Física de la Tierra y Astrofísica. Universidad Complutense de Madrid, Spain

**Corresponding Author: Alberto Martilli, Despacho 1.11, Edificio 70, CIEMAT, Avenida Complutense 40, 28040 Madrid, Spain.
E-mail: alberto.martilli@ciemat.es*

Abstract

Persistent wintertime inversions over cities are the most critical conditions for air quality, and also one of the most challenging situations to simulate with a meteorological model. In the first part of this study, the ability of the meteorological model WRF, coupled with the multilayer urban canopy parameterization BEP-BEM, to simulate the evolution of the Planetary Boundary Layer Structure over the city of Madrid, Spain, during a multiday inversion episode, is assessed. The model results are evaluated against airport soundings, fourteen meteorological stations within and around the urban area, and remotely sensed surface temperatures. The study indicates that PBL structure is determined by the interaction between the urban and rural heat and momentum fluxes, the topography, and the downward turbulent transport of heat. The best air temperature spatial distribution is obtained when the 6th order horizontal filter is applied only to wind and not to temperature, and when the soil moisture is reduced to 25% of the initial value provided by the global scale model. However, the comparison against satellite surface temperature data indicates that with such a dry soil the model underestimates the surface temperatures during the night and overestimates them during the day in rural areas. This error compensation points to a likely deficiency in the PBL and surface schemes during the persistent inversions. In urban areas, the simulations with the urban canopy parameterization tend to overestimate the nocturnal surface and air temperatures, while they reproduce wind speed correctly. Finally, a new methodology to assess the model's capability to reproduce the spatial variability of air temperature is introduced, and the results show that the use of the multilayer urban canopy scheme and the adjustment of the soil moisture are both needed to improve the reproduction of such spatial distribution.

Keywords:

Mesoscale modelling, thermal inversion, urban canopy paramterizations, air quality, complex terrain.

1. Introduction.

The highest concentrations of $\text{PM}_{2.5}$ and NO_2 in urban areas are usually observed during periods of several consecutive days characterized by low winds and strong thermal inversion that prevent the dispersion of pollutants. A layer with thermal inversion is a layer where temperature increases with height (which is contrary to the normal behavior in the atmosphere). In such layers, there is a very strong stability with little turbulent vertical exchange. These situations happen usually during winter and are characterized by the presence of a mass of warm air a few hundreds of meters above the surface. During the night, the radiative cooling of the ground surface induces cooling of the adjacent layers of the atmosphere. Moreover, for regions located in complex terrain (valley bottom, or close to a mountain ridge), the cooling is reinforced by the drainage of cold air induced by nocturnal katabatic or mountain flows. In these conditions, the inversion starts from the surface and is therefore called *surface-based inversion*. During the daytime, the solar radiation is absorbed by the surface and converted into heat to warm the layers of air directly at contact with the surface (Wei et al. 2013). Buoyancy-induced vertical motions are generated, but their vertical extension is limited by the presence of the mass of warm air above. The base of the inversion layer is then lifted few hundreds of meters above ground level. This is called a *capped inversion* layer because its base is elevated above the surface. In this context, the presence of the city, due to its roughness and large thermal capacity, and the anthropogenic heat emission is expected to behave differently than rural areas, potentially inducing large spatial variability (Fernando and Weil, 2010), and modifying the strength, thickness, and height of the inversion layers. It is the interplay between the inversion layer and the effect of the urban areas that determines the time evolution of the urban boundary layer structure, and eventually, the pollutant concentration levels in the city. In this context, modeling plays a relevant role both as a scientific tool to investigate the behavior of a complex non-linear system, as well as a practical tool to forecast future air quality levels and to assess the effectiveness of air pollution abatement strategies. However, the simulation of a strong inversion layer is recognized as one of the most challenging tasks for a mesoscale model, because it requires reproducing the fine interaction between the surface-atmosphere scheme, the Planetary Boundary Layer scheme, and numerical schemes (like horizontal filters, for example).

The most studied case is certainly Utah's Salt Lake Valley both with measurements (Lareau et al. 2013), and models (Lu and Zhong, 2014, Wei et al. 2013, Sun et al. 2020), but similar phenomena have been studied and modeled in alpine valleys (Burns and Chemel, 2014, Chemel and Burns, 2015, Chemel et al. 2016, Largeron and Staquet, 2016), and in other regions of the world, like Santiago de Chile (Saide et al. 2011).

Even if Salt Lake City has been home to one of the first experiments dedicated to the urban atmosphere (URBAN 2000, Alwine et al. 2002), to our knowledge there are no studies with mesoscale models with urban canopy parameterizations (UCP) published over that region, given that most of the studies used standard mesoscale models, and focused mainly on the dispersion (Darby et al. 2006). Extending the analysis to numerical studies for other cities in similar situations, we can conclude that mesoscale simulations of persistent wintertime inversions over cities performed by models with UCP are lacking. This is probably because, historically, mesoscale models with UCP have been used and evaluated for summer periods, aiming to understand the causes of urban overheating and develop tools that can be used to evaluate heat mitigation strategies, and little attention has been paid to winter situations, except Salamanca and Mahalov (2019), and Pappaccogli et al. (2021).

Therefore, the research questions that motivate this study are:

- 1) *Are today's mesoscale models able to simulate stable boundary layers over cities? Is the inclusion of a UCP beneficial?*
- 2) *What is the role of the city in modulating the strength of the inversion?*

To address these questions, the mesoscale model WRF (Skamarock et al. 2008), coupled with the multilayer urban canopy parameterization BEP-BEM (Martilli et al. 2002, Salamanca et al. 2010), has been used to simulate a multi-day episode of thermal inversion that occurred over the region of Madrid (Spain) at the end of December 2016 (Borge et al. 2018). Madrid is a region that frequently hosts these types of episodes in winter and therefore is an interesting case study. However, compared to cases like Salt Lake City, Santiago de Chile, or cities in alpine valleys, mentioned above, which are all locations bounded by two mountain chains, the case of Madrid is different having only one mountain chain (the Sierra de Guadarrama) on one side. Moreover, the city is on undulating terrain, with small topographic depressions where localized cold air pools can form. As it will be shown below, these elements are very important and determine the evolution of the PBL structure over Madrid during persistent wintertime inversions.

This work is presented in two parts: this first part focuses on meteorology, while the second part (Martilli et al. 2021) analyzes the pollutant dispersion.

In section 2 the case study, available measurements and model set-ups are presented. Section 3 is dedicated to the validation of the atmospheric variables. Section 4 presents a new methodology to assess the model's capability to represent spatial distribution of air temperatures. Section 5 focus on the analysis of the three dimensional meteorological fields produced by the model, and section 6 is where the remote sense surface temperature produced by the model are compared with the satellite data. Conclusions are in section 7.

2. Case Study

2.1. Site description

Madrid is located in the center of the Iberian Peninsula, on a high-plateau at approximately 600 m a.s.l. (Fig. 1), with a mountain ridge (Sierra de Guadarrama) at NW, that reaches 2428 m a.s.l.. The core of the "old city" is bounded by two N-S small valleys, crossed by the rivers Manzanares and Jarama, still urbanized. The difference in elevation between the bottom of the valley and the hill where the city's old core is, is about 80-100m. In the municipality of Madrid, approximately 3 million people live (with a population density of 5400 hab/km²), while in the entire region of Madrid the population is 6 million with a density of 2900 hab/km². The rural areas surrounding Madrid are typically Mediterranean, although modified by the altitude and isolation from the sea with a mix of farmlands and shrublands, and evergreen needle leaf trees in particular in the Sierra. In the IGBP-MODIS classification, used by WRF, the dominant land use surrounding Madrid is "savanna". Note that, given the longitude of Madrid, 3.7 W, UTC corresponds to Local Solar Time.

2.2. Observational data

Several sources of observational data are used to characterize the episode, and evaluate the model results :

- 1) vertical soundings at 00UTC and 1200 UTC from Barajas airport.
- 2) Satellite-derived surface temperatures from MODIS (1km resolution) for the 26th of December.
- 3) hourly data from ground meteorological stations, including a meteorological tower (CIEMAT), eight stations of the Spanish met office (AEMET), and five of the air quality network of Madrid's municipality, all of these located in the urban area or close by. Fig. 1 shows the location of the ground meteorological stations used in the study, and more information about those stations is in Table 1.

Station	Urban fraction	Building width (m)	Street width (m)	Mean Building Height (m)	H/W (of the urban fraction of the grid cell)	Height of measurements above ground level (m)	Height above sea level of the grid cell where the station is.
U1 Ciudad Universitaria	0.34	7.2	33.4	19	0.56	2	673.
U2 Retiro	0.86	23.7	17.1	19.7	1.15	2	662.
R1 Barajas	non-urban	Landuse class Savannas	-	-		2	589.
U3 Cuatro Vientos	0.67	7.2	62.1	5.7	0.09	2	687.
U4 Getafe	0.65	10.5	12.4	10.	0.80	2	617.
U5 Torrejon	0.85	28.8	32.5	8.3	0.25	2	599.
R2 Arganda del Rey	non-urban	Landuse class Croplands				2	538.
R3 El Goloso	non urban	Landuse class Savannas				2	736.
U6 CIEMAT	0.87	8.8	36.1	12.6	0.35	4 and 52	681.
U7 Moratalaz	0.31	4.1	21.1	7.7	0.37	12	685.
U8 Villaverde	0.73	9.9	32.8	10.2	0.31	9	605.
U9 La China	0.91	6.0	49.3	19.	0.38	7	589.
R4 Acustica	Non-urban	Landuse class Savannas	-	-		3	609.
U10 Hortaleza	0.99	10.0	44.8	14.4	0.32	9	710.

Table 1. Ground meteorological stations used in the study, and the characteristics of the grid points where the stations are located. The urban fraction and morphological parameters are those used by WRF.

The meteorological tower of CIEMAT (U6) is in the NW part of the city of Madrid in a vegetated area in the CIEMAT's campus (see the picture in the SI). It provides measurements of wind speed, wind direction, and air temperature at 52 m above ground level, and air temperature at 4m above ground.

Among the stations of the Spanish Met Office (AEMET), four of them (R1 - Barajas, U3 - Cuatro Vientos, U4 – Getafe, and U5 - Torrejon de Ardoz) are stations located in airports. However, the surrounding areas are different (see pictures in the SI), and for this reason, R1 is in a grid cell classified as rural, while the other three stations (U3, U4, and U5) fall in an urban grid cell. U1 - Ciudad Universitaria is in the North of the city, in an urban grid cell, with a low urban fraction, close to the CIEMAT's tower (U6), while U2 -Retiro is the oldest meteorological station in Madrid, located in a park in the center of the city. R2 – Arganda del Rey and R3 – El Goloso are rural stations. The comparison between U2 (urban) and R1 (rural) have been often used for Urban Heat Island studies in Madrid (Yagüe et al. 1991, Rasilla et al. 2019, Maqueda et al. 2020). All the temperatures measured at these stations are at 2m a.g.l.. Standard T2 WRF output, in rural areas, is the result of a logarithmic interpolation, modified to account for atmospheric stability, based on Monin-Obukhov Similarity Theory (MOST), while, in urban areas, giving that MOST is not valid, T2 is just the temperature of the lowest model level (when BEP-BEM is used). To be consistent, instead of T2, the temperature of the

lowest model level is used also for R1, R2, R3, based also on the consideration that the thickness of the lowest model level is about 4-5m, with temperature defined at half level (2-2.5m).

For stations U6, U7, U8, U9, U10, and R4, which belong to the city council of Madrid, the measurement point is higher than 2m a.g.l., and it varies from one station to the other (Table 1). Model results are then linearly interpolated at the height of the measurements.

Finally, to be consistent with the measurements, the results are time-averaged over one hour, and the average is done internally in WRF based on the values of each timestep.

2.3. The episode.

The episode of persistent inversion analyzed in this study lasted from the 25th until the 30th of December 2016. Its time evolution can be seen in the vertical soundings obtained at Barajas's airport, located 12 km NE of the city area (Fig. 2a). Starting from the 25th of December a mass of warm air (up to 14° Celsius) is present few hundred meters above the surface. This strong inversion prevents vertical mixing. In this period, the wind is weak close to the surface (no more than 3m/s) and from East (Fig. 2c,e). This situation lasts until the night between the 28th and the 29th of December when a wind from the North of about 4m/s brings cold air and breaks the thermal inversion. The second part of the episode starts on the 29th and lasts until the 30th. During this second period, the mass of warm air is located higher up. Moreover, the air mass is slightly colder (1-2° C less) than the one in the first part of the episode. This, and the very low wind, is likely what allows the stronger cooling in the lower layers observed in this second part of the episode. During both periods, the base of the capped inversion during the daytime assumed to be an indication of the mixing layer (Borge et al. 2018), is very low, around or less than 100m above ground level. As observed in Borge et al. 2018, where a more thorough description of the episode can be found, the days when the base of the inversion in the sounding is lower, are also the days when the pollutant concentration (NO₂ and PM_{2.5}) is higher.

2.4. Model set-up.

WRF version 4.1 has been used to perform the simulations, with the multilayer BEP_BEM scheme to represent the urban areas (Martilli et al. 2002, Salamanca et al. 2010). Four nested domains have been used, with resolutions 27km, 9km, 3km, and 1km (Fig. 3). In the vertical, a constant resolution of approximately 5m is used for the lowest 10 vertical levels (e. g. up to 50 m a.g.l.), and then the resolution is progressively stretched for a total of 57 numerical grid levels (33 of them in the lowest 1km), until 20km. Such high resolution is considered necessary to resolve part of the urban canopy and the shallow inversion layer. The analysis is focused on the results of the inner domain, centered on the city of Madrid and including most of the close mountain ridge (Sierra de Guadarrama). This domain has 111 grid points in the E-W direction and 123 in the N-S. To characterize the city, high-resolution (1m) LIDAR's derived data are used to estimate plan area building density (fraction of horizontal surface occupied by buildings, λ_p), total surface building density (the building envelope area, divided by the total horizontal area, λ_b), distribution of building heights, and impervious fraction (urb_frac) for the most urbanized area of the domain (Fig. 3). The region with higher values of λ_p indicates the old urban core, which conserves the ancient structure, while the region of high building heights in the northern part represents the new financial district. Outside this region, the Local Climate Zones (LCZs, Stewart and Oke, 2012) classes derived by Brousse et al. (2016) have been used to define the morphological parameters used by the model. The building materials have been defined based on the Spanish Building Code, including a layer of insulation in walls and roofs. Building heating is considered by BEM, with a comfort range between 17.5 and 24.5°C. Furthermore, the dependency of the drag coefficient with the building packing density of Santiago and Martilli (2010) has been introduced in the model.

For the non-urban areas, the NOAA scheme (Chen and Dudhia, 2001) has been used. The PBL scheme is Bougeault and Lacarrere (1989). First-order Horizontal Smagorinsky diffusion is used, and it is computed along the true horizontal direction (not terrain-following) in the 4th domain, following the suggestion of Zangl (2002).

Boundary and initial conditions are derived from the ECMWF ERA5 reanalysis. The simulation starts at 00UTC on the 22nd of December, to allow 24hrs of spin-up and simulate the beginning of the inversion period, and lasts until 00UTC on the 31st of December 2016.

Compared to the standard version of WRF, several modifications have been made :

- *Soil temperature initialization.* The standard procedure in WRF consists of initializing the soil temperature using data from a global model, in our case ECMWF ERA5 data. To account for the difference between the coarse resolution topography used in the global model and the fine resolution used in the nested domains, standard WRF performs an extrapolation using a default gradient of -6.5° K/km, which results in hotter soil temperatures for the grid cells with a height below the terrain elevation of the global model and colder temperatures for the points above such height. In the case of strong atmospheric inversion, this is not realistic and may lead to anomalous warmer soil at low altitudes. For this reason, in this study, a gradient derived from the difference in air temperature measured at midnight (time at which the model has been initialized) between the meteorological station of Navacerrada in the local Sierra de Guadarrama (1894m a.s.l.) and the meteorological station of Barajas airport (R1, approx. 600m a.s.l.) was used instead of the -6.5° K/km gradient for the extrapolation. In this specific case, for the midnight of 22nd of December 2016, this gradient is null, so the soil temperature is initialized based on the soil temperature of the ECMWF model, but constant with the topographical altitude.
- *Soil moisture initialization.* Another key initialization parameter is the soil moisture. Soil moisture has a strong impact on the energy partitioning between storage, sensible and latent heat fluxes because it directly affects the heat capacity and diffusivity of the soil, and the amount of latent heat flux. High values of soil moisture are expected to favor latent over sensible heat flux, and to increase the storage, reducing the nocturnal cooling, and daytime heating. Usually, there is great uncertainty about this variable, giving the scarcity of observational data and its strong spatial variability. Some studies propose methodologies based on several months of simulations, using models like HRDLAS (Chen et al. 2007), where the soil-vegetation model is run offline, while others use a self-cycling spin up (Angevine et al. 2014). A different approach has been taken in this study. Four sensitivity runs have been performed with 100% (**100%SMOIS**), 75% (**75%SMOIS**), 50% (**50%SMOIS**), and 25% (**25%SMOIS**) of the soil moisture provided by ERA5, which is around 0.38 m³/m³ in the plateau where Madrid is located. This value is greater than those observed during moist periods in the closest meteorological station that measures soil moistures, located in the mountain 50km N-W of the city (C. Yagüe, personal communication – note that the precipitation was low in November and December). It is important to stress that ERA5's soil moisture is not based on measurements, but it is the output of another soil and vegetation model. Then, the time evolutions of the air temperature at the lowest model level at the four meteorological stations located in rural areas that are in the area of interest (R1, R2, R3, and R4, see the location in Fig. 1, and Table 1 for more information) have been analyzed in Figure S1 and the statistics reported in Fig. 4. In three stations (R4, R1, and R3) **25%SMOIS** shows the lowest RMSE, and in all four stations, this simulation is also the closest to the maximum temperatures (Fig S1). During the nighttime, **25%SMOIS** is also the closest to the measurements at R4, R3, and in particular, at R1, the meteorological station of Barajas Airport, often chosen as representative of rural areas in the region in UHI studies. Different behavior is present at R2, which has a different land use (croplands) than the others, where **25%SMOIS** significantly underestimates the nighttime temperatures, in particular towards the end of the episode. This may be an indication that the initial soil moisture field provided by ERA5 does not capture the spatial

variability of this variable, probably due to the coarse resolution of ERA5. On the other hand, the analysis of the Relative Humidity at R1 (the only rural station where this variable is available, Fig S1), shows completely different behavior. In all the simulations the RH drops faster than the measurements during the nine days simulated, indicating a general dry bias in the model, and **25%SMOIS** is the worst, and **100%SMOIS** the best. This contrasting behavior is a clear indication that the model is missing some relevant physical mechanisms, either in the heat and moisture exchanges between the surface and the atmosphere or in reproducing the structure of the Planetary Boundary Layer. Tests with a different PBL scheme (MYJ, Janic, 2012), a different surface scheme (Jimenez et al. 2012), and a different vegetation scheme (NOAH-MP), did not change the results significantly (not shown). Solving this problem is beyond the scope of the present article, therefore in the following analysis, we will just focus on the **25%SMOIS** and **100%SMOIS** simulations to assess the sensitivity of the model to this parameter. It is also worth mentioning that soil moisture of 0.1 m³/m³ is typical for Spain in December (Roman-Cascon et al. 2020), so the value chosen for **25%SMOIS** is not unrealistic. Another thing to consider is that the wind speed at R1 is not very sensitive to the value of soil moisture, which is an important consideration for pollutant dispersion modeling.

- **6th order filter.** The 6th order horizontal filter of Knierel et al. (2007) has been employed in the simulations. This scheme is very effective in removing undesired 2-4 Δx features from the kinematical fields. In particular in urban areas, where a strong spatial variability in building-induced drag can easily trigger 2 Δx noise, the application of this filter is essential. However, in the standard version of WRF, this filter is applied not only to the three velocity components but also to the potential temperature. Given that this filter is applied on terrain-following coordinates (not in the true horizontal direction), it may generate artificial vertical mixing of potential temperature over sloping terrain, something that artificially reduces the inversion strength. This is the reason why the possibility to avoid the application of this filter in regions of steep topography has been added in the last releases of WRF. This option is activated in our simulations, and also, the filter is applied only to the three components of the velocity, and **not** to the potential temperature (by commenting out the call to the routine). Thanks to this, the model was able to better reproduce the nocturnal cooling observed in some stations (Fig. 5), not only those on or close to the sloping ground (like stations U8, U9, and R4, in the topographic depression of the Manzanares river) but also stations U7, and U10, not close to slopes. This is because the 6th order filter, when applied to temperature, mixes in the horizontal, dumping the local effects. We consider this setup appropriate given that the main objective of the filter is to avoid the accumulation of kinetic energy at a 2-4 Δx scale. So, it is important to eliminate features at this scale from the wind field, but not from the temperature fields. Applying this filter to the temperature would result in a loss of the model's capability to reproduce the impact of the surface spatial heterogeneity at the scale of the model resolution (in this case 1km). Indeed, 2 Δx features in the temperature field can generate perturbations at the same scale in the pressure field, and therefore in the wind field – this is why it is even more necessary to apply the filter to the three velocity components. Note also that the sensitive runs to soil moisture described above were done with this configuration activated.

Therefore, we chose to analyze the following simulations to answer the research questions:

Name	Soil moisture	landuse	Specific objectives	
No City	25%	Urban areas replaced by savanna	Evaluate the impact of the city on the PBL structure	
25%SMOIS	25%	True urban area		Analyze the sensitivity of the

100%SMOIS	100%	True urban area		model to soil moisture
------------------	------	-----------------	--	------------------------

Table 2. Simulations considered in the study

3. Validation of atmospheric variables.

3.1. Vertical soundings.

In Fig.2, b, d, f, (right panels) the vertical profiles of air temperature produced by **25%SMOIS**, following the path of the sounding started from Barajas airport, have been plotted, and in Fig. S2 the same graphs are presented for **No City** and **100%SMOIS**. The three simulations can reproduce the development of the elevated inversion layer on the days 26th to 28th, the breaking of the inversion on the 29th, and the development of the second part of the episode on the 30th. **25%SMOIS** and **No City** produce very similar results, as it can be expected because Barajas airport is a rural area, while **100%SMOIS** has a colder temperature at around 300-400 m a.g.l, implying that soil moisture and then rural surface heat fluxes, may affect the strength of the inversion. Following Borge et al. 2018, in Table 3, the base of the capped inversion layer at midday – identified where the vertical gradient of air temperature changes sign from negative to positive and considered an indicator of mixing layer depth – calculated by the three simulations is compared to the one obtained from the sounding.

Elevated inversion base at 12UTC in meters	23rd	24th	25th	26th	27th	28th	29th	30th
Sounding	184	142	199	98	92	144	135	100
25%SMOIS	238	122	140	123	108	140	107	94
100%SMOIS	313	313	160	140	108	139	107	93
No city	208	140	123	108	95	123	83	83

Table 3. Height of the base of the inversion layer at 1200 UTC in the three simulations and the sounding.

The three simulations can follow reasonably well this height, with **100%SMOIS** slightly overestimating the inversion height in particular during the first days of the episode, and **No city** underestimating it during the last three days. The range of low wind speed is also reproduced by the three runs, as well as the prevalent wind direction from East at 300-400m above ground and the Northern winds close to the surface during nighttime.

3.2. Ground stations.

The meteorological stations listed in Table 1 provided hourly data of temperature, and in some cases wind speed and direction. Their location is indicated in Fig. 1, and an aerial picture of each station is provided at the end of the SI.

3.2.1 Air temperature.

A summary of the statistics computed for the air temperature, without the first 24 hours considered as spin-up period, for all the stations is presented in Fig. 6 (Root Mean Square Error, RMSE) and Fig. 7 (BIAS), while in the SI there are figures of the time series of each station.

Let focus on the rural stations first, and in particular R1, Barajas Airport. There, **No city** has the largest RMSE and a negative BIAS and **25%SMOIS** the smallest. This is an indication that the temperature close to the surface at that station is likely influenced by the city, and therefore does not represent the temperature that

would be there without the city. This influence is particularly important during the night when **No city** strongly underestimates the temperatures (Fig. S3). On the other hand, **100%SMOIS** overestimates the temperature during the night and underestimates it during the day. Similar behavior is present at R3, while at R2, the negative bias of **25%SMOIS** and **No city** is due to an underestimation during the night, not present in **100%SMOIS** that, on the other hand, underestimates the maxima. The biggest differences between the three runs are present at R4 (Fig. S4), located in the topographic depression of the Manzanarre's river West of the city, in particular during night-time when **100%SMOIS** strongly overestimates temperature, **No city** underestimates them, and **25%SMOIS** can capture the trend relatively well. Moreover, **100%SMOIS** also underestimates the maxima.

Regarding the urban stations, at U1 the large value of the RMSE and positive BIAS in **25%SMOIS**, are mainly due to the overestimation of air temperature during the night, in particular the days 26th, 27th, 28th, and 30th. Subgrid effects may be present for U2, which is located in a park (see the aerial picture in SI), despite the grid cell where it falls has an urban fraction of 0.86. If we move just one grid point in the direction of the park, in a grid cell with an urban fraction of 0.48, probably more representative of the value measured, the cooling in **25%SMOIS** is stronger during nighttime and the agreement improves (not shown). The better RMSE and BIAS for **100%SMOIS** at U2 could also be explained by the fact that the soil moisture for an urban park may behave differently, due to irrigation, than for a rural area outside the city. In U3, U4, and U5, located in airports, the best agreement is for **25%SMOIS** and **100%SMOIS**, with bias below 1 °C, while **No city** shows a large negative bias, mainly due to underestimation of air temperature during nighttime (Fig. S3). At U7, U8, and U9 the largest discrepancies are during the night with **25%SMOIS** and **100%SMOIS**, the two simulations that represent the city and uses BEP-BEM, overestimating the nighttime temperature (Fig. S4). **No city** has smaller discrepancies at U7 and U8, despite these stations are urban, and underestimates the temperature at U9, another station that falls in an urban grid cell with a high urban fraction, but which is likely more representative of rural land-use (see the aerial picture in SI). Moreover, **100%SMOIS** underestimates the maxima at the three stations. At all these stations (U7, U8, and U9), localized on the S-E part of the city, the observed minimum temperatures go to zero and below, while **25%SMOIS** and **100%SMOIS** stay 2-3 °C warmer during the night. However, **25%SMOIS** and **100%SMOIS** are capable to capture the warmer observed nighttime temperatures (above 0 C) at station U10, a densely urbanized area (0.99 of urban fraction) at the northern part of the city (Fig. S4).

At both heights of the CIEMAT's tower (U6), **25%SMOIS** is the simulation with the smaller RMSE and BIAS (Fig. 6, and 7), and the negative bias at 52m is mainly due to an underestimation the 23rd and 24th (Fig. S5). Overall, at 4m, **No City** has lower nighttime temperatures, in some days in better agreement with the measurements than **25%SMOIS**, while **100%SMOIS** underestimates the maxima, and, in some days, overestimates the minima. At 52m, the strong negative BIAS for **100%SMOIS** indicates that the strength of the inversion is underestimated because the gradient with the values at 4m is reduced, while the negative BIAS of **No city** at both heights can be seen as the result of the impact of the city on air temperature.

3.2.2 Wind speed and direction

The analysis of wind direction (Fig. S5) at U6, shows a predominant wind from 60 Degrees (NE) during the whole period, interrupted, mainly during afternoon hours, by periods of wind swinging between South and West, likely a weak signal of anabatic winds. The model is quite successful in reproducing this behavior except for the afternoon of the 27th and 28th when no change in direction is modeled, differently than what was observed in the measurements. Wind speed (Fig. 8, and S5) is better reproduced by **25%SMOIS** and **100%SMOIS**, the two simulations that take into account the presence of the city with BEP-BEM, with a slight negative bias, while **No City** overestimates the wind speed.

For some of the AEMET ground stations, wind speed and direction are also available. Statistics for wind speed are presented in Fig. 8, while the time series of wind direction and speed are in Fig. S6 and S7. Overall, the two simulations that consider the presence of the city with BEP-BEM (**25%SMOIS** and **100%SMOIS**) behave very similarly and have the best RMSE, below 1 m/s at most built-up sites, while **No city**, as it could

be expected, has worse statistics. However, **25%SMOIS** and **100%SMOIS** underestimate the wind speed in the stations located NE of the city (R1 and U5), where **No city** has better RMSE and BIAS. For the rest of the stations, **25%SMOIS** and **100%SMOIS** have better RMSE and BIAS than **No city**. Regarding wind direction (Fig. S6) the dominant direction ranges between North (R1), and East (R3), with the majority of stations having a wind from 60 Degrees. This direction is broken only during few hours in the afternoon when the wind shifts from the South. There are small differences between the three runs, that, in general, reproduce well this behavior.

4. Spatial distribution.

An important skill expected from the model is the capability to reproduce the spatial distribution of air temperatures. To quantify this, the *differences* for the daily minimum and maximum temperatures between the ground meteorological stations produced by the model are plotted against those deduced from the observations. All the possible couples of stations have been formed, and the differences between the minimum (maximum) temperatures of two stations have been calculated, or

$$d_{min}(i, j) = Tmin(i) - Tmin(j)$$

$$d_{max}(i, j) = Tmax(i) - Tmax(j)$$

where $d_{min(max)}(i, j)$ is the difference in minimum (maximum) daily temperature between stations i and j . The values of the differences observed vs. modeled are plotted in Fig. 9 (minimum), and 10 (maximum) for the 24th, representative of the period before the episode, the 26th and 27th representing the first part of the episode, and the 30th for the second part of the episode. In such plots, the points that are in the lower left (LL) and upper right (UR) quadrants indicate that the difference in temperatures has the same sign in the model and the observations, so the model correctly reproduces the observed trend in temperature difference between two stations. For example, if the observations show that station A is warmer than station B, the model shows the same. Moreover, points that are above (below) the diagonal in the LL (UR) quadrant indicate that the model underestimates the difference in temperature observed. Points out of these quadrants indicate that the model does not estimate the sign of the difference correctly. To these categories, we added the points with a difference of less than 1° C (within the square at the center of the figure), a value considered within the measurement error. For those points, the model basically reproduces correctly the differences observed, even if the point lies in the Upper Left or Lower Right quadrant. On the graph, a linear fit for the three simulations (dashed lines) is also plotted, indicating the general tendency of the simulations. The percentage of points in the UR and LL quadrants and within the central box are reported in Tables 4 and 5.

Spatial variability for minimum temperatures is greater than for maximum temperatures. This tendency is present both in the observations and the three runs and is due to the stronger turbulent mixing during the daytime that homogenizes the temperatures, compared to the nocturnal situation when the strong atmospheric stability and the topography determine the large spatial variability. This is why the nocturnal variability is also larger during the days of the episode than before it (24th). In fact, during the 24th, **25%SMOIS** and **No city** reproduce the observed trend in minimum temperature differences in about 75% of the couples of stations, while only 50% is obtained from **100%SMOIS** (Table 4). All three simulations tend to underestimate the differences, in particular **100%SMOIS**. During the days of the episode (26th, 27th, 30th), on the other hand, **25%SMOIS** has slightly more points in the UR and LL quadrants than the other two simulations, but **100%SMOIS** has a clear tendency to underestimate the differences, while **25%SMOIS** and **No city** to overestimate them (Tab. 4). This, therefore, indicates that a good model performance of the thermal behavior in rural areas is an important factor to capture the spatial distribution of air temperature across Madrid under these synoptic conditions.

On the other hand, all simulations underestimate the observed spatial variability for maximum temperatures (Fig. 10). However, the general trend, indicated by the linear fit, indicates that in particular the 27th and 30th, **100%SMOIS** has an opposite behavior than the measurements (linear fit in the LR and UL quadrants).

		24th	26th	27th	30th
<u>Minimum temperatures</u>					
% in Upper Right and Lower left (sign of differences correctly simulated)	25%SMOIS	73.1	89.7	79.5	88.5
	NO city	75.6	75.6	74.4	73.1
	100%SMOIS	51.3	75.6	76.9	79.5
Overestimations	25%SMOIS	19.2	53.9	41.0	61.5
	NO city	23.1	48.7	44.9	48.7
	100%SMOIS	12.8	10.3	20.5	29.5
Underestimations	25%SMOIS	48.7	30.1	34.6	21.8
	NO city	43.6	25.6	26.9	21.8
	100%SMOIS	32.1	52.6	52.6	47.5
Within 1° C	25%SMOIS	5.2	5.7	3.9	5.2
	NO city	8.9	1.3	2.6	2.6
	100%SMOIS	6.4	12.7	3.8	2.5

Table 4. Statistics for the couples of minimum temperature.

		24th	26th	27th	30th
<u>Maximum temperatures</u>					
% in Upper Right and Lower left	25%SMOIS	76.9	69.2	91.0	64.1
	NO city	80.8	61.5	84.6	79.5
	100%SMOIS	75.7	60.3	68.0	42.3
Overestimations	25%SMOIS	2.6	10.3	20.5	9.0
	NO city	7.7	14.6	36.4	18.2
	100%SMOIS	12.8	20.5	26.9	10.3
Underestimations	25%SMOIS	47.4	51.3	51.3	30.8
	NO city	53.9	34.6	43.6	43.6

	100%SMOIS	39.7	29.5	24.4	15.4
Within 1C	25%SMOIS	26.9	38.4	19.2	24.3
	NO city	19.2	20.0	4.6	17.7
	100%SMOIS	23.2	10.3	16.7	16.6

Table 5. Statistics for the couples of maximum temperature

5. Meteorological fields

To have a general picture of the most important physical mechanisms that shape the urban boundary layer structure over Madrid in these conditions we focus on three days: the 24th (representative of the situation before the episode when the inversion was present but weaker than during the episode), the 27th (first part of the episode), and the 30th (second part). For each day we analyze the situation at 0700 UTC, the time of the minimum temperature, and at 1500 UTC, the time of the maximum.

4.1 - - At 0700 UTC

Figure 11 shows horizontal fields of modeled air temperature at the lowest model level (approx. 2.5 m above ground level), and wind field interpolated at 10m a.g.l. (the typical height of the measurements) for the three simulations at 0700 UTC on 24th.

The structure of the temperature field for **25%SMOIS** shows a region of higher temperatures over the densest part of the city, surrounded by cooler areas in correspondence of the two topographic depressions of the Manzanares and Jarama rivers. It is interesting to see the sharp gradient of temperature close to U1, U6, and R4, in the NW part of the city, present both in the model and the measurements (represented by the colored boxes). Overall, for the other stations, the agreement with the simulation is good (the color of the boxes is similar to the color of the temperature field), except U8 and U9, where the model overestimates the temperature. The wind field shows a clear pattern with a flow coming from NE, induced by a larger scale sloping terrain, and a mountain northerly flow in the two topographic depressions of the Manzanares and Jarama. Comparing **25%SMOIS** with **No city**, we can see the city's contribution to air temperature, as well as the effect on the northerly flow in the two depressions which is consistently slowed down by the presence of the city. Combining the underestimation of air temperature in U8 and U9 of **No city**, with the overestimation of **25%SMOIS**, we can hypothesize that BEP-BEM is reducing too much the cooling and the cold air advection in that part of the city. **100%SMOIS**, on the other hand, does not have enough cooling in the two depressions. Likely, the heat capacity of the moister soil is too high to generate enough cooling. As a consequence, there is no northerly flow in the depressions.

To have a more complete picture of the situation, a vertical section from SW to NE passing through R1 (Barajas Aiport) is presented in Fig. 12. Through this section, there are also U7, U8, and U9, and U4 is just at few grid points (see Fig. 1). The temperature sections of both **25%SMOIS** and **No city** clearly show the strong stratification in the two depressions, more intense in **No city**, but no stratification is present in **100%SMOIS**, due to the insufficient surface cooling, as already mentioned. On the other hand over the hill between the two depressions, where the densest part of the city is, the stratification is less strong, also due to a significant downward vertical turbulent transport of heat. This is likely due to a turbulence shear generated above the hill. As shown in Fig. S8, in fact, the flow from NE “jumps” over the two depressions, that are decoupled from that flow due to the strong atmospheric stability, and therefore generates vertical shear above the hill. Comparing the flow produced by the three simulations in Fig. S8, we can see how the city, being rougher and warmer than the *savannas* (the landuse used to replace the urban area in the **No city** simulation) reduces the wind speed over the hill.

The contrast between the cooler topographical depressions, and the warmer hill in between, is accentuated on 27th during the first part of the episode compared to 24th both for **25%SMOIS** and **No city** (Fig. 13), likely due to a weaker wind. The most interesting thing is that also **No city** shows similar behavior to **25%SMOIS**. Giving that in **No city** the landuse is uniform in the region, these differences are only due to the interactions between the stability and the topography. The consequence is that, in these conditions, the difference in temperature between an urban station (e.g. U2), and a rural station (e.g. R1), which are often used to evaluate the UHI intensity, are not only due to the difference in land use, but also to other factors. As shown in Fig. 14 for **No City**, this is mainly due to the downward turbulent flux of heat, linked to the presence of a mass of warm air above. This flux is stronger in **No city** because it has to compensate for a stronger surface cooling compared to **25%SMOIS**, but it is over a shallower layer. In **25%SMOIS**, probably due to the turbulence generated by the city, this downward transport of heat happens over a thicker layer. Note that the magnitude of these fluxes, between 50 W/m² (**25%SMOIS**) and 100 W/m² (**No city**) is significant and comparable with the upward flux during daytime driven by solar radiation in this period of the year. Again, **100%SMOIS** does not show enough cooling, in particular in the depression. For all simulations, the general structure of the wind field is similar to the one described for the 24th, with a general flow from NE, and decoupled flow from the North in the two depressions (Fig. S9). While weaker than for the 24th, the vertical section still shows a maximum of wind speed above the city, likely the result of gravity waves excited by the hill in such a very stable atmosphere. This maximum is stronger and closer to the surface in **No city** than in **25%SMOIS** and **100%SMOIS**, likely because **No city** has less surface friction than the other two simulations.

Temperatures are cooler on the 30th than the previous days, both in the measurements and in the three model results (Fig. 15). This time **25%SMOIS** overestimates the minimum temperatures, in particular in the part of the urban area on the hill, and to less extent in the topographical depressions. On the other hand, **No city** reproduces better the temperatures on the hill, but too cold in the depressions. Also based on the results of Fig.16, we can hypothesize that the overestimation of the minimum temperatures on the top of the hill in **25%SMOIS** might be linked with a combination of not enough surface cooling, and too much vertical mixing that brings warm air to the surface. **100%SMOIS** shows an overestimation of temperatures everywhere, and in particular in the depressions. The general structure of the wind field, both near the surface and in the vertical, is very similar to the one of the previous days analyzed, but with a weaker intensity (Fig. S10). We can then conclude that, despite the differences in the temperature fields between the three days, the winds are very similar, in all the cases dominated by a combination of the NE wind generated by the drainage over larger-scale sloping terrain, and the Northerly mountain flow that forms in the two topographic depressions.

4.2 At 1500 UTC

During the daytime, at 1500 UTC of the 24th, the spatial variability of temperature is much smaller than at 0700 UTC (Fig. 17), as shown in the previous section, and something typical for Madrid, as it has been seen also for summer situations (Salamanca et al. 2012). **25%SMOIS** and **No city** show very similar fields, with temperature values in the range of the measurements, while **100%SMOIS** underestimates the temperatures, due to the larger thermal inertia of the moist soil, and the larger proportion of latent heat fluxes. This underestimation of temperature in **100%SMOIS** is present also in urban areas, even though the change in moisture affects only the rural areas and the non-urban fraction of the urban grid cells. This shows how important is the thermal behavior of the rural surroundings for the city's atmosphere itself. The vertical structure of the PBL over the city (Fig. 18) is quite similar for the three runs, with a PBL height that reaches about 350-400 m ag.l. above the city. The wind speed is in general very low in the simulations within the PBL, without a clear wind direction (Fig. S11), while the ground measurements show slightly stronger winds from NE. No anabatic wind is present in the simulations and the measurements, something relatively common for winter periods in the region.

The temperature fields at 1500 UTC of the 27th for **25%SMOIS** and **No city**, show again very small spatial variability in both runs (Fig. 19), and overall, a higher temperature compared to 24th. There is a very small difference between the rural and the urban areas, and between the hill and the topographical depressions. Measurements confirm this, with an even higher temperature than those simulated. On the other hand, **100%SMOIS** significantly underestimates the temperatures, in particular in the rural areas. Again, the urban area of **100%SMOIS**, even though it is simulated in the same way as **25%SMOIS** (i. e., with BEP-BEM), is colder than **25%SMOIS** and the measurements.

This is likely due to the combination of the influence of the surrounding rural areas, and the thermal behavior of the non-urban fraction of the urban grid cells. This can be seen in the vertical sections of Fig. 20, where the potential temperature and the vertical kinematic heat flux are plotted. **25%SMOIS** and **No city** show clear heating in the rural region NE of the city (right side in the plots). This heat is slowly pushed by the wind (Fig. S12) over the city and joins the mass of hot air present aloft. Then **25%SMOIS** develops a PBL that mixes down to the surface part of this heat. This entrainment of warm air is weaker in **No city**, due to a weaker surface sensible heat flux and a smoother surface that generate less turbulence and a shallower PBL. In other words, on the 27th, the PBL is heated from below (surface) but also from above (entrainment), and this explains why it is warmer than the 24th. In **100%SMOIS**, where there is little heating in the rural areas, the PBL over the city does not grow enough to develop significant entrainment. The role of the rural areas is then crucial on the 27th to vertically mix the heat, even over the city.

A completely different picture is given on the 30th. This day the measurements show that the city is colder than the rural surroundings (Fig. 21). **25%SMOIS** and **No city** reproduce the trend, but still overestimate temperature in the urban area, while **100%SMOIS** gets a better estimate of the urban temperatures, but underestimates significantly the rural ones, missing then the pattern (see also the analysis done in section 4). Moreover, the wind, even if still weak, has changed the direction completely and now blows from the South (Fig. S13). The vertical sections (Fig. 22) show a very strong stratification that starts at about 200-300m a.g.l. that stops the growth of the PBL, but also reduces the entrainment of warm air in the PBL, and, combined with synoptic changes, likely explain why the temperature in the PBL is colder than the 27th.

6. Satellite surface temperatures.

Remotely sensed surface temperatures have been, and are, extensively used to assess the thermal impact of a city focusing mainly on the Surface Urban Heat Island (Peng et al. 2012). However, they have been comparatively less used to quantitatively validate mesoscale models.

Here we compare the surface temperature recorded by MODIS at 1100UTC and at 2300 UTC of the 26th of December, the only available clear data for the period of simulation. The spatial resolution is approximately 1km. The nocturnal image (Fig. 23, at 2300 UTC) from MODIS shows a warmer region of(3-4 °C) located at the densest urban area, surrounded by regions with surface temperatures around 0 °C in less densely urbanized areas, and -2 °C in rural areas. To compare the model results to remotely sensed satellite measurements, in the urban grid points, the following equivalent surface temperature has been derived:

$$Tsk_{sat} = \left(\frac{(1 - f_{urb})Rl_{up_{rur}} + f_{urb}(wRl_{up_{road}} + bRl_{up_{roof}})}{\sigma \epsilon} \right)^{\frac{1}{4}}$$

Where f_{urb} is the urban fraction, $Rl_{up_{rur}}$, $Rl_{up_{road}}$, $Rl_{up_{roof}}$ are the upward longwave radiation from the non-urban fraction (rural) of the grid cell, the road, and the roof, respectively, estimated as follows

$$\begin{aligned} Rl_{up_{rur}} &= \sigma \epsilon_{rur} Tsk_{rur}^4 + (1 - \epsilon_{rur}) Rl_{down} \\ Rl_{up_{road}} &= \sigma \epsilon_{road} Tsk_{road}^4 + (1 - \epsilon_{road}) Rl_{road} \\ Rl_{up_{roof}} &= \sigma \epsilon_{roof} \sum_{k=1}^{nu} s_k Tsk_{roof_k}^4 + (1 - \epsilon_{roof}) Rl_{down} \end{aligned}$$

With σ Stefan Boltzmann constant, ε_{rur} , ε_{road} , ε_{roof} emissivities of vegetation fraction, road and roof, Tsk_{rur} , Tsk_{road} surface temperatures of the rural and road, and Tsk_{roof_k} surface temperature of the roof of a building with k floors. Rl_{down} is the longwave radiation downward from the sky reaching the roofs and the rural fraction, and Rl_{road} is the longwave radiation reaching the road, including the amount of radiation reflected/emitted by the walls. Finally, the emissivity ε is estimated as follows

$$\varepsilon = (1 - f_{urb})\varepsilon_{rur} + f_{urb}\varepsilon_{urb}$$

Where ε_{urb} is an average emissivity of the urban surfaces. This approach neglects the contribution of the walls (which are instead included in the calculation of the TSK variable of the standard output of WRF), and it has been adopted because the view angle of the satellite, in particular, if close to Nadir as it is this case, can "see" only roofs and roads, and not walls. In non-urban grid cells, the standard skin temperature (TSK) value has been used.

At 2300 UTC, the map produced by **25%SMOIS** shows an overestimation of the surface temperature in urban areas, in particular in the zones at intermediate building density, and a clear underestimation in the rural areas. As a consequence, the contrast between the urban and rural areas is much sharper in the model than in the measurements. The BIAS for **25%SMOIS** is 3 °C in the urban areas, and -0.5 °C in the rural ones at 2300 UTC (Table 6), even if NW of the city there is an area with much colder temperatures than those observed. The map from **100%SMOIS** has approximately the same surface temperature in the urban areas of **25%SMOIS**, but in the rural areas, the RMSE is decreased from 2.6 °C to 1.5 °C, with a positive BIAS. There is, then, a contrasting behavior of air and surface temperatures between the two simulations, being **25%SMOIS** closer to the observed air temperature, but **100%SMOIS** closer to the measured surface temperatures. In the map produced by **No city**, the surface temperature closely follows the topography, with higher temperatures (3-4 °C) in the regions above 600m, and lower in the two topographic depressions generated by the rivers Manzanares and Jarama, with surface temperatures that go down to -5 °C or even -6 °C. This behavior is the result of the strong temperature inversion and the dry soil of this configuration. There is, indeed no way to evaluate if this simulation is correct. However, this is another indication that the air temperature contrasts between urban and rural areas in Madrid in these conditions are not due only to the differences in the surface fluxes induced by the different land use, but also to the interaction between the atmospheric stability and the topography. Interestingly, this simulation that does not use BEP-BEM, being cooler than the other two, is also the one that has the smallest RMSE in the urban area.

During the daytime, the contrasting tendency between **25%SMOIS** and **100%SMOIS** is confirmed, with **25%SMOIS** strongly overestimating the temperatures in the rural areas, while **100%SMOIS** being closer to the MODIS values (Fig. 24). So, **25%SMOIS** heats too much the rural surface during the day and cools it too much during the night. **No city**, sharing the same soil moisture with **25%SMOIS**, is also overheating the surface.

1100 UTC	RMSE Urban	RMSE Rural	BIAS Urban	BIAS Rural
25%SMOIS	1.6	3.8	0.1	3.1
100%SMOIS	1.7	1.9	-1.0	-0.5
No City	2.8	4.0	2.1	3.4
2300 UTC	RMSE Urban	RMSE Rural	BIAS Urban	BIAS Rural
25%SMOIS	3.8	2.6	3.1	-0.5
100%SMOIS	3.6	1.5	3.1	0.7

No City	2.7	3.2	0.1	0.5
---------	-----	-----	-----	-----

Table 6 Statistics for the surface temperature for the three simulations.

7. Summary and conclusions.

To assess the ability of WRF to simulate a period of persistent wintertime inversion over Madrid, as well as the interactions between the city and the PBL structure in such conditions, three simulations with the mesoscale model WRF have been carried out. Two of them represent the city with the multilayer urban canopy parameterization BEP-BEM but have different soil moisture, one with the original value provided by ERA5 reanalysis (**100%SMOIS**) and the other with 25% of such value (**25%SMOIS**) at all soil levels. The third simulation shares the same soil moisture with **25%SMOIS** but replaces the city with savanna, the dominant rural land use in the region (**No city**).

From the analysis of the results, the following conclusions have been reached about the structure of the PBL, over the area in these conditions:

- The horizontal and vertical structure of the PBL over Madrid during strong atmospheric inversion is the result of the interaction between the urban and rural surface heat and momentum fluxes, the topography, and the strength of the downward turbulent heat flux (entrainment), which is affected by the level of turbulence developed over the city and the thermal stratification.
- During the nighttime, the temperature fields are characterized by an accumulation of cold air in the two topographic depressions caused by the rivers Jarama and Manzanares, and relatively warmer air in the hill in between the depressions, where the densest part of the city is located. Over the hill, there is entrainment of warm air from above, that is absent in the depressions. Note that, while the temperature pattern is the same before and during the episode, the intensity of the spatial temperature gradients changes. In particular during the first part of the episode (e.g., 27th December) when there is a mass of warm air at a relatively low altitude above the city, the entrainment plays a very relevant role in warming the hill where the city is, but not the depressions. This is different during the second part of the episode (e.g., 30th December) when the mass of warm air is higher up, leaving space for a stronger surface cooling, even on the hill.
- The fact that during nighttime the temperature field is not only linked to the surface heat fluxes (so to the land use) but also the topography and the strength of the entrainment, has an important consequence for UHI studies. For example, this implies that the difference in air temperature between an urban site (for example, St. U2-Retiro) and a rural station (for example St. R1, Barajas, located in one of the two depressions), does not represent only the contribution of the city to the temperature field. Moreover, the comparison between **25%SMOIS** and **No city** indicates that the temperature at R1, despite its rural land use, is influenced by the city. In conclusion, by no mean, we can consider the temperature measured at a rural station like R1 as representative of the temperature that there would be at the urban station U2 (or in another point of the city), if the city was not there, which is the basic assumption behind any UHI study (Martilli et al. 2020).
- In addition, this situation causes that the nocturnal thermally-driven wind from NE (induced by larger-scale sloping terrain) to "jumps" over the two depressions, which are decoupled from the air above and where a northerly, valley wind is present. The same wind structure is observed every day before and during the episode, but the strength of the winds changes.
- During the daytime, the temperature field shows much less spatial variability than during the night, and no (or very weak) thermally-driven wind develops. The dominating factors that determine the temperature are the heat fluxes from the rural areas and the vertical structure of the atmosphere. For example, during the first part of the episode (27th) when the mass of warm air is at a lower altitude above the city, the strong vertical mixing determines greater heating than during the second part (30th) that has stronger stability and lower temperatures.

From the modeling point of view the following conclusions can be derived:

- While BEP-BEM is useful to reproduce the spatial variability of temperature and in particular wind speed in the city, it is even more important to simulate the correct heating/cooling in the surrounding rural areas, to reproduce the spatial distribution of the temperature fields. Failing to do so, results in an incorrect representation of the spatial distribution of the temperature both during daytime and, in particular, during nighttime.
- To reach the goal of reproducing correctly the thermal behavior, in particular during nighttime, our results indicate that it is essential to :
 - Avoid the use of the 6th order horizontal filter for the temperature. This filter acting on terrain-following coordinates, in case of sloping terrain artificially mixes temperature vertically and prevents the formation of the stability. Moreover, due to its horizontal mixing, it reduces the capability of the model to represent correctly the impact of the variability in urban structure within the city. It is important to keep the filter for velocities, to avoid $2\Delta x$ structures, but it is not necessary for temperature.
 - Adapt the soil moisture to obtain the right temperature behavior in rural areas. In our case, this meant reducing the soil moisture by a factor of 4 compared to the value of ERA5.
- While the soil moisture value used (0.1 m³/m³) is in a realistic range for Spain in this period of the year, the analysis of the relative humidity and, in particular, surface temperature as derived from the satellite, indicates that it causes a dry bias in the atmosphere, and too cold surface temperatures during the night and too hot during daytime in the rural areas compared with the MODIS values. This contradictory behavior of the model likely means that the true soil moisture was higher than 0.1 m³/m³, but the surface layer/pbl scheme used by WRF needs such a low value to produce the correct atmospheric stability close to the surface, which means that there is a possible error compensation in the model. A reason could be an underestimation of the surface exchange coefficients used in the surface layer scheme, which therefore requires a stronger difference between surface and air temperatures to produce the right fluxes, as hypothesized by Sun et al. (2020), or/and a problem in the PBL scheme to simulate stable conditions as indicated by Maroneze et al. (2020). An essential future work that needs to be done is to improve the capability of the PBL scheme, in connection with the urban canopy parameterization, to represent stable conditions.
- Regarding the urban areas, results show that BEP-BEM tends to overestimate both surface and air temperatures during the night and underestimate them during the day, with an overall slight negative bias for the wind speed. This may point to the need to adapt the drag coefficient, and the heat exchange coefficients, to these stability conditions. So far BEP-BEM has been tested mainly for summer periods in strong convective, or neutral conditions, but less in winter and stable conditions, so that some adjustments may be necessary.

Acknowledgments.

This work has been done in the framework of the project "Extreme Climate Events in Urban Areas," No.CGL2016-80154-R funded by the Spanish Ministry of Science, Innovation, and Universities." We also thank M. Pujadas (CIEMAT) for providing the data of the CIEMAT's tower, Madrid's city council for ground stations data, and AEMET for ground stations and soundings.

References

Allwine, K.J., Shinn, J.H., Streit, G.E., Clawson, K.L. and Brown, M., 2002. Overview of URBAN 2000: A multiscale field study of dispersion through an urban environment. Bulletin of the American Meteorological Society, 83(4), pp.521-536.

- Angevine, W.M., Bazile, E., Legain, D. and Pino, D., 2014. Land surface spinup for episodic modeling. *Atmospheric Chemistry and Physics*, 14(15), pp.8165-8172.
- Borge, R., Artíñano, B., Yagüe, C., Gomez-Moreno, F.J., Saiz-Lopez, A., Sastre, M., Narros, A., García-Nieto, D., Benavent, N., Maqueda, G. and Barreiro, M., 2018. Application of a short term air quality action plan in Madrid (Spain) under a high-pollution episode-Part I: Diagnostic and analysis from observations. *Science of the Total Environment*, 635, pp.1561-1573.
- Bougeault, P. and Lacarrere, P., 1989. Parameterization of orography-induced turbulence in a mesobeta--scale model. *Monthly weather review*, 117(8), pp.1872-1890.
- Brousse, O., Martilli, A., Foley, M., Mills, G. and Bechtel, B., 2016. WUDAPT, an efficient land use producing data tool for mesoscale models? Integration of urban LCZ in WRF over Madrid. *Urban Climate*, 17, pp.116-134.
- Burns P and Chemel Ch (2014): Evolution of Cold-Air-Pooling Processes in Complex Terrain. *Boundary-Layer Meteorology*. 150. 423-447. 10.1007/s10546-013-9885-z.
- Chemel Ch and Burns P (2015): Pollutant Dispersion in a Developing Valley Cold-Air Pool. *Boundary-Layer Meteorology*. 154. 391-408. 10.1007/s10546-014-9984-5.
- Chemel Ch, Arduini G, Staquet Ch, Largeron Y, Legain D, Tzanos D and Paci A (2016): Valley heat deficit as a bulk measure of wintertime particulate air pollution in the Arve River Valley. *Atmospheric Environment*. 128. 208-215. doi: 10.1016/j.atmosenv.2015.12.058.
- Chen, F. and 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(4), pp.569-585.
- Chen F., Kevin W. Manning, Margaret A. LeMone, Stanley B. Trier, Joseph G. Alfieri, Rita Roberts, Mukul Tewari, Dev Niyogi, Thomas W. Horst, Steven P. Oncley, Jeffrey B. Basara, and Peter D. Blanken, 2007: Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System. *J. Appl. Meteor. Climatol.*, 46, 694–713.
- Darby, L.S., Allwine, K.J. and Banta, R.M., 2006. Nocturnal low-level jet in a mountain basin complex. Part II: Transport and diffusion of tracer under stable conditions. *Journal of applied meteorology and climatology*, 45(5), pp.740-753.
- Fernando, H.J.S. and Weil, J.C., 2010. Whither the stable boundary layer? A shift in the research agenda. *Bulletin of the American Meteorological Society*, 91(11), pp.1475-1484.
- Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P. and García-Bustamante, E., 2012. A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, 140(3), pp.898-918.
- Lareau, N.P., Crosman, E., Whiteman, C.D., Horel, J.D., Hoch, S.W., Brown, W.O. and Horst, T.W., 2013. The persistent cold-air pool study. *Bulletin of the American Meteorological Society*, 94(1), pp.51-63.
- Largeron, Y. and Staquet, C., 2016. The atmospheric boundary layer during wintertime persistent inversions in the Grenoble valleys. *Frontiers in Earth Science*, 4, p.70.
- Lu, W. and Zhong, S., 2014. A numerical study of a persistent cold air pool episode in the Salt Lake Valley, Utah. *Journal of Geophysical Research: Atmospheres*, 119(4), pp.1733-1752.
- Janjic, Z.I., 2002. Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. *NCEP Off. Note* 2002, 437, 61.
- Jiménez, P.A., Dudhia, J., González-Rouco, J.F., Navarro, J., Montávez, J.P. and García-Bustamante, E., 2012. A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, 140(3), pp.898-918.
- Knievel, J.C., Bryan, G.H. and Hacker, J.P., 2007. Explicit numerical diffusion in the WRF model. *Monthly Weather Review*, 135(11), pp.3808-3824.

- Maqueda, G., Yagüe, C., Román-Cascón, C., Serrano, E., and Arrillaga, J. A.: Analysis of temperature trends and urban heat island in Madrid, EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-10969, <https://doi.org/10.5194/egusphere-egu2020-10969>, 2020.
- Maroneze, R., Acevedo, O.C., Costa, F.D., Puhales, F.S., Anabor, V., Lemes, D.N. and Mortarini, L., 2020. How is the Two-Regime Stable Boundary Layer Reproduced by the Different Turbulence Parametrizations in the Weather Research and Forecasting Model?. *Boundary-Layer Meteorology*, pp.1-31.
- Martilli, A., Clappier, A. and Rotach, M.W., 2002. An urban surface exchange parameterization for mesoscale models. *Boundary-layer meteorology*, 104(2), pp.261-304.
- Martilli, A., Krayenhoff, E.S. and Nazarian, N., 2020. Is the urban heat island intensity relevant for heat mitigation studies?. *Urban Climate*, 31, p.100541.
- Martilli, A., Sanchez B., Santiago J. L., Rasilla D., Pappaccogli G., Allende F., Martin F., Roman C., Yagüe C., Fernandez, F., 2021, Persistent Wintertime Thermal Inversion in the City. The case of Madrid. Part II. Dispersion. Submitted to *Atmospheric Research*.
- Pappaccogli G., L. Giovannini, D. Zardi, A. Martilli., 2021, Assessing the ability of WRF-BEP+BEM in reproducing the wintertime building energy consumption of an Italian Alpine city, Accepted for publication in *Journal of Geophysical Research*.
- Peng, S., Piao, S., Ciais, P., Friedlingstein, P., Ottle, C., Bréon, F.M., Nan, H., Zhou, L. and Myneni, R.B., 2012. Surface urban heat island across 419 global big cities. *Environmental science & technology*, 46(2), pp.696-703.
- Rasilla, D., Allende, F., Martilli, A. and Fernández, F., 2019. Heat waves and human well-being in Madrid (Spain). *Atmosphere*, 10(5), p.288.
- Román-Cascón, C., Lothon, M., Lohou, F., Ojha, N., Merlin, O., Aragonés, D., González-Dugo, M.P., Andreu, A., Pellarin, T., Brut, A. and Soriguer, R.C., 2020. Can We Use Satellite-Based Soil-Moisture Products at High Resolution to Investigate Land-Use Differences and Land-Atmosphere Interactions? A Case Study in the Savanna. *Remote Sensing*, 12(11), p.1701.
- Saide, P.E., Carmichael, G.R., Spak, S.N., Gallardo, L., Osses, A.E., Mena-Carrasco, M.A. and Pagowski, M., 2011. Forecasting urban PM10 and PM2.5 pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model. *Atmospheric Environment*, 45(16), pp.2769-2780.
- Salamanca, F., Krpo, A., Martilli, A. and Clappier, A., 2010. A new building energy model coupled with an urban canopy parameterization for urban climate simulations—part I. formulation, verification, and sensitivity analysis of the model. *Theoretical and applied climatology*, 99(3), pp.331-344.
- Salamanca F. and Mahalov, A., 2019. Summer-and Wintertime Variations of the Surface and Near-Surface Urban Heat Island in a Semi-arid Environment. *Weather and Forecasting*, 34(6), pp.1849-1865.
- Santiago, J.L. and Martilli, A., 2010. A dynamic urban canopy parameterization for mesoscale models based on computational fluid dynamics Reynolds-averaged Navier–Stokes microscale simulations. *Boundary-layer meteorology*, 137(3), pp.417-439.
- Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W. and Powers, J.G., 2008. A description of the Advanced Research WRF version 3. NCAR Technical note-475+ STR.
- Stewart, I.D., and Oke, T.R., 2012. Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), pp.1879-1900.
- Sun, X., Holmes, H.A. and Xiao, H., 2020. Surface Turbulent Fluxes during Persistent Cold-Air Pool Events in the Salt Lake Valley, Utah. Part II: Simulations. *Journal of Applied Meteorology and Climatology*, 59(6), pp.1029-1050.
- Yagüe, C., Zurita, E. and Martinez, A., 1991. Statistical analysis of the Madrid urban heat island. *Atmospheric Environment. Part B. Urban Atmosphere*, 25(3), pp.327-332.

Wei, L., Pu, Z. and Wang, S., 2013. Numerical simulation of the life cycle of a persistent wintertime inversion over Salt Lake City. *Boundary-layer meteorology*, 148(2), pp.399-418.

Zängl, G., 2002. An improved method for computing horizontal diffusion in a sigma-coordinate model and its application to simulations over mountainous topography. *Monthly Weather Review*, 130(5), pp.1423-1432.

Figure 1. Topography of the region of Madrid and position in the Iberian Peninsula. The symbols indicate the stations, see text.

Figure 2 Time evolution of the vertical profile of temperature (a,b), wind speed (c,d) and wind direction (e,f) observed (left column) and modelled by the 25%SMOIS simulation(right column) at Barajas airport at 00 and 12 UTC from the 22nd until the 30th of December 2016.

Figure 3. Top nested domains used in the simulations. Mid panel left: urban fraction used by the model in the central part of the domain. Solid line indicates city boundaries. Mid panel right: plan area building density. Bottom: mean building height in m. These data have been derived from LIDAR measurements.

Figure 4. RMSE (top 4 panels) and bias (bottom 4) for 2m air temperature at four rural stations in the domain, for the four sensitivity runs with different initial soil moisture (100%, 75%, 50% and 25% of the value provided by the ERA5 reanalysis).

Figure 5. Impact of the 6th-order filter on time evolution of air temperature at some stations. Diamonds are the measurements, orange solid line is the simulation with 6th order filter applied to the velocity components and the temperature, and black line is the simulation where the filter has been applied only to the velocity components and **not** the temperature.

Figure 6. RMSE for air temperature (Celsius) for all the stations and the three simulations.

Figure 7. Mean BIAS for temperature (Celsius) for all the stations and all the simulations.

Figure 8. RMSE (top two rows) and BIAS (bottom two rows) for wind speed (m/s) for all the stations and the three simulations.

Minimum Temperatures

Fig. 9 Plots of the differences $d(i,j)$ in minimum temperature between couple of stations modelled (vertical axis) and observed (horizontal). Given that $d(i,j)=-d(j,i)$ only half of the couples have been plotted. Black diamonds are **25%SMOIS**, green triangles **No city**, and orange squares **100%SMOIS**. The dashed line is a linear fit obtained for each simulation and indicates the general trend.

Maximum Temperatures

Fig. 10. Same as Fig. 12 but for maximum.

Figure 11. Simulated temperature fields at about 2.5m a.g.l. (half of the lowest model level) in Celsius (left column), and wind field at 10m a.g.l. (right column, the red vector on the lower left of the figures is 2m/s) for the 24th at 0700 UTC. Solid black line is the limit of the urban areas. Red arrows are measurements (all at 10m a.g.l, except U6. which is at 52m. a. g.l.). Temperature measurements are represented in the squares using the same color scale. The stations can be identified in Fig. 1. Top panels are for 25%SMOIS, middle for No city, and bottom for 100%SMOIS.

Figure 12. SW-NE vertical section, along the solid blue line of Fig. 1, of air temperature (Celsius), left column, and vertical turbulent flux of potential temperature (K m/s), right column. Top panel is for **25%SMOIS**, middle panel for **No city**, and bottom for **100%SMOIS**. Squares indicate the position of stations, from left to right. Square's colors indicate the measured values, in the same color scale. Black bar indicate the position of the city.

Fig. 14 Same as Fig. 10 but for the 27th

Fig. 16 Same as Fig. 10 but for the 30th

Fig. 17. Same as Fig. 9 but for the 24th at 1500 UTC.

Fig. 18 Same as Fig. 10 but for the 24th at 1500 UTC. Here on the left panels the potential temperature is plotted, giving that during daytime is more representative of the thermal structure of the PBL. The temperature at the stations has been transformed to potential using the pressure of the model.

Fig. 19. Same as Fig. 15 but for the 24th at 1500 UTC.

Fig. 20 Same as Fig. 16 but for the 24th at 1500 UTC

Fig. 21. Same as Fig. 15 but for the 24th at 1500 UTC.

Fig. 22 Same as Fig. 16 but for the 24th at 1500 UTC

Fig. 23. Surface temperature at 23UTC of the 26th, for **MODIS** (top left), **25%SMOIS** (top right), **No city** (bottom left), and **100%SMOIS** (bottom right).

Fig. 24. Surface temperature at 23UTC of the 26th for **MODIS** (top left), **25%SMOIS** (top right), **No city** (bottom left), and **100%SMOIS** (bottom right).

Supplemental Information

Fig. S1. Time series measured (diamonds), and simulated by the four simulations with different soil moisture (see legend) for air temperature, relative humidity and wind velocity (see title of the graphs) at the rural stations.

*Fig. S2. Same as Fig. 2, but for **No city** (left) and **100%SMOIS** (right)*

Fig. S3. Time series measured (diamonds), and simulated, 25%SMOIS (black line), No city (Green), and 100%SMOIS (orange), for temperature at AEMET stations (see title to identify the station).

Fig. S4. Same as Fig. 3 but for City Council stations (see title to identify the station).

Fig. S5. Time series measured (diamonds), and simulated, 25%SMOIS (black line), No city (Green), and 100%SMOIS (orange), for temperature at 4m and 52m, and wind velocity and direction, both at 52m at st. U6- CIEMAT

Fig. S6. Time series measured (diamonds), and simulated, 25%SMOIS (black line), No city (Green), and 100%SMOIS (orange), for wind direction at AEMET stations (see title).

Fig. S7. Time series measured (diamonds), and simulated, 25%SMOIS (black line), No city (Green), and 100%SMOIS (orange), for wind speed at AEMET stations (see title).

Fig. S8. Vertical section of wind speed (left) and wind direction (right), for the 0700 UTC of the 24th through the SE-NW section indicated by solid line in Fig. 1.

0700 UTC 27/12

Wind Speed (m/s)

Wind Direction (Deg)

Fig. S9. Same as Fig. S8, but for the 0700 UTC of the 27th

0700 UTC 30/12

Wind Speed (m/s)

Wind Direction (Deg)

Fig. S10. Same as Fig. S8, but for the 0700 UTC of the 30th

Fig. S11. Same as Fig. S8, but for the 1500 UTC of the 24th

1500 UTC 27/12

Wind Speed (m/s)

Wind Direction (Deg)

Fig. S12. Same as Fig. S8, but for the 1500 UTC of the 27th

1500 UTC 30/12

Wind Speed (m/s)

Wind Direction (Deg)

Fig. S13. Same as Fig. S8, but for the 1500 UTC of the 30th

Aerial photos of the ground stations from Google Earth