Contents lists available at ScienceDirect

Atmospheric Research

On the short-term simulation of heat waves in the Southeast Mediterranean: Sensitivity of the WRF model to various physics schemes

Christos Giannaros[a],[*], Dimitrios Melas[a], Theodore M. Giannaros[b]

[a] Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki, PO Box 149, Thessaloniki 54124, Greece
[b] National Observatory of Athens, Institute for Environmental Research and Sustainable Development, Vas. Pavlou & I. Metaxa, Penteli 15236, Greece

ABSTRACT

The current study investigated the impact of the Weather Research and Forecasting (WRF) (a) surface layer (SL), (b) land surface and (c) planetary boundary layer (PBL) physics on the short-term simulation of the meteorological and human thermal comfort conditions during 15 heat waves identified between 2004 and 2013 in the southeast Mediterranean and Balkan Peninsula. Four widely used PBL-SL schemes (YSU-MM5, MYJ-Eta, ACM2-Revised MM5, YSU-Revised MM5) and land surface models (LSMs; Noah, Noah-MP, CLM4, RUC) were tested. Comparisons with ground-based observations in 60 measuring sites showed that the WRF model is characterized by a cold bias leading to the underestimation of the extreme heat stress conditions over the entire study region. The MYJ-Eta experiment simulated greater sensible heat fluxes compared to the other PBL-SL scenarios due to the high surface heat exchange coefficient provided by the Eta SL scheme. This contributed to the high MYJ-Eta-modeled daytime 2-m air temperature (T2) and physiologically equivalent temperature (PET) values that were close to the observations. The ACM2 PBL algorithm produced significantly higher PBL heights compared to rest of the tested PBL parameterizations reflecting the strong vertical mixing generated in this scheme. This contributed to the most accurate ACM2-Revised MM5-modeled T2 and PET results in overall. The use of the RUC LSM strengthened significantly the modeled sensible heat fluxes contributing directly to the satisfactory replication of the observed T2 and PET values. However, the RUC-induced enhanced sensible heating relative to the latent heating is may related to a possible underestimation of the soil moisture, as indicated by the notable dry biases during the RUC experiment. Thus, the most accurate T2 and PET results provided by the RUC model may arise from the wrong reason.
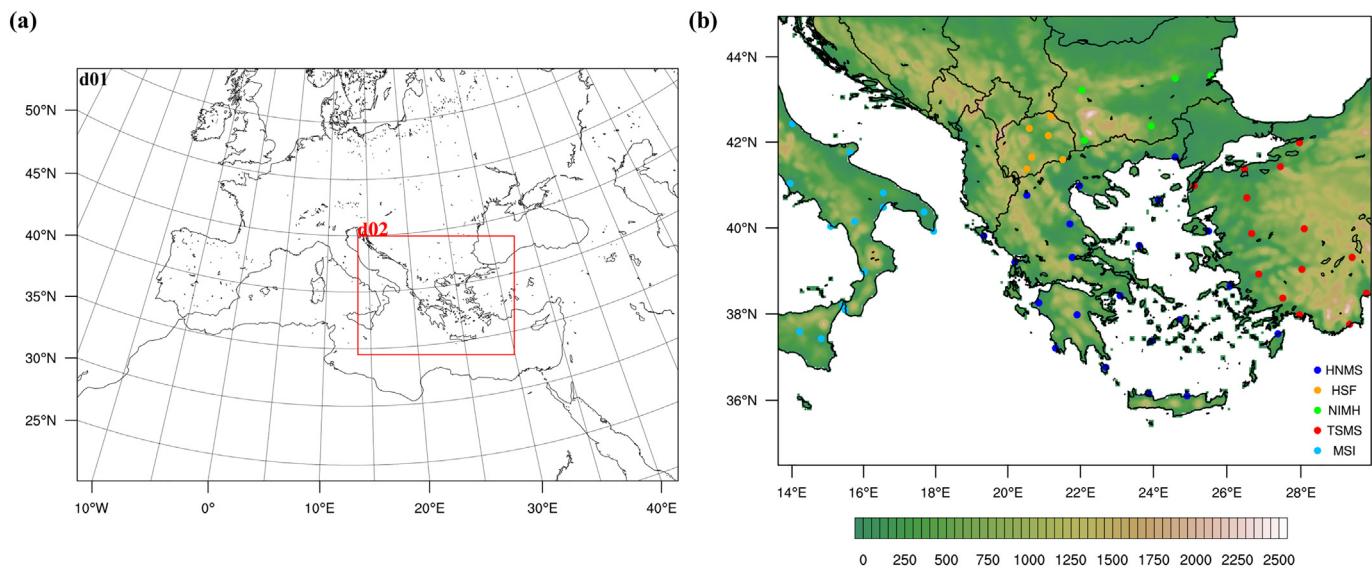
1. Introduction

Heat waves (HWs) affect severely the natural and human environment. Their most critical impact is the deterioration of the human thermal comfort conditions, which can lead to a significant increase in morbidity and mortality. For instance, during the June 2007 HW episode over the National Health's Operational Center (NHOC) of Greece reported over 140 emergency department visits for heat exhaustion and heatstroke, as well as six deaths due to excessive heat exposure in the Attica region (Athens; Theoharatos et al., 2010). Future projections of HWs show that the intensity, frequency and duration of such extreme hot weather events will increase in the twenty-first century (Beniston, 2004; Meehl and Tebaldi, 2004), especially in the SE Europe (Founda and Giannakopoulos, 2009; Nastos and Kapsomenakis, 2015). Under these future climate conditions, numerical weather prediction (NWP) models can serve as a valuable tool in developing short-term HW

forecasting systems that can be used operationally considering human-biometeorological factors to assist the public health protection services (Giannaros et al., 2018a; Giannaros et al., 2015; Matzarakis and Nastos, 2011; Ramamurthy et al., 2017).

In this direction, a variety of studies implemented with the use of the mesoscale meteorological Weather Research and Forecasting (WRF) model highlighted the substantial role that land surface processes play in the causation and evolution of extreme hot weather conditions (Akta, 2011; Chiriaco et al., 2014; Dasari et al., 2014; Stéfanon et al., 2014; Vautard et al., 2013; Zeng et al., 2014). These physical mechanisms are parametrized in the WRF modeling system by the land surface models (LSMs). Zeng et al. (2011, 2015) and Ma et al. (2017) and Attada et al. (2018) demonstrated the significant LSM-induced differences in short- and medium-range WRF simulations of HWs. These differences are associated with the initialization of the surface parameters and the representation of the surface sensible heating in the LSMs. Also, they are

---

* Correspondence author.
E-mail address: cgiannar@auth.gr (C. Giannaros).

(a)

(b)

Fig. 1. (a) Configuration of the two 2-way nested WRF modeling domains. (b) Topography of the SE Mediterranean and Balkans Peninsula with identification of the locations of the ground-based WMO weather stations.

strongly related to the meteorological and climate characteristics of each simulated domain. For this, more research is necessary over different geographical regions. Beyond the land surface processes, the planetary boundary layer (PBL) dynamics significantly affect the high-temperature weather episodes (Miralles et al., 2014) while the surface-layer (SL) parameterizations, which in some cases are tied to the choice of the PBL scheme, provide the interconnection between the LSM and the first atmospheric model level (Shaffer et al., 2015). Thus, the reproduction of HWs is also depended on the ability of the PBL-SL schemes to attain the linkage between the surface and the boundary layer and predict the PBL processes adequately (Sathyanadh et al., 2017). Many researchers have investigated the performance of various PBL-SL schemes in the WRF model over several regions and during numerous atmospheric conditions (Banks et al., 2016; Cohen et al., 2015; García-Díez et al., 2013; Hu et al., 2010; Shin and Hong, 2011; Xie et al., 2012; Zhang et al., 2013). However, studies focused on extreme hot weather conditions are relatively rare (Kotroni et al., 2011; Sathyanadh et al., 2017).

The principal objective of the present study is to investigate the impact of various WRF (a) surface layer, (b) land surface, and (c) planetary boundary layer parameterizations on the short-term simulation of HW events over the SE Mediterranean and BP. Four different PBL-SL physics (YSU-MM5, MYJ-Eta, ACM2-Revised MM5, YSU-Revised MM5) and LSM options (Noah, Noah-MP, CLM4, RUC) were tested during 15 HWs (57 HW days) identified between 2004 and 2013. Observational data for the same period were used to evaluate the model performance under the various sensitivity experiments, focusing on near-surface atmospheric variables and particularly on the 2-m air temperature. Additionally, a thermal bioclimate index, i.e., the physiologically equivalent temperature (PET), was computed through the application of the RayMan model (Matzarakis et al., 2007, 2010) to examine the capability of the WRF model under the examined physics to replicate the observed human thermal comfort conditions. The factors contributing to the modeled differences were thoroughly investigated for a severe case study event (June 2007).

2. Methodology

2.1. Identification of HW events

In the present study, a temperature criterion based on the IPCC's definition of extreme temperature index, TX90p (Hartmann et al.,

2013), was applied for the identification of the examined heat waves. According to this index, a HW day is determined when the daily maximum temperature exceeds the 90th percentile of a late-20th century reference period (1961–1990). Thus, HW days were identified by applying the TX90p indicator to the air temperature data of three ground-based weather stations over Greece: (a) Helliniko (Athens), (b) Larissa, and (c) Thessaloniki. When at least three consecutive HW days were found at the same time over all the stations mentioned above, then a HW episode was defined – except from two events (July 2010 and July 2013) that defined as 2-day HWs due to the lack of identifying a third continual HW day in these episodes. In total, 15 HW events (57 HW days) were identified between 2004 and 2013 (see supplementary material Table S1). It is worth mentioning that even though the identification criterion was applied only over Greece, the selected HW episodes affected the whole southeast Mediterranean and Balkan Peninsula. This is illustrated in the 0.5°x0.5° climate forecast system (CFS) reanalysis upper-level maps (850 hPa geopotential height and temperature) for the 24–27 June 2007 HW event (see supplementary material Figs. S1a, c and S2a, c), which is a characteristic case concerning the synoptic conditions and the geographical coverage (more details are presented in the Sec. 3.3).

2.2. Meteorological data sources

For the evaluation process, ground-based observational data reported at World Meteorological Organization (WMO) stations were retrieved from five monitoring networks operated by the (i) Hellenic National Meteorological Service (HNMS), (ii) Hydrometeorological Service of the former Yugoslav Republic of Macedonia (FYROM; HSF), (iii) National Institute of Meteorology and Hydrology in Bulgaria (NIMH), (iv) Turkish State Meteorological Service (TSMS), and (v) Meteorological Service of Italy (MSI). Based on the data availability, 60 measuring sites were selected in the study region, which is divided into three sub-regions based on the orography (Gilliam et al., 2006): (a) coastal (altitude > 25 m), (b) inland (25 m < altitude < 350 m), and (c) mountain (altitude > 350 m). The key featres of the weather stations are summarized in Table S2 (see supplementary material), while Fig. 1b illustrates the locations of each measuring site.

Table 1
Summary of the sensitivity experiments.

	PBL scheme			Surface layer scheme			Land surface model			
Experiment	YSU	MYJ	ACM2	MM5	Revised MM5	Eta	Noah	Noah -MP	CLM4	RUC
YSU-MM5 (Noah)	X			X			X			
MYJ-Eta		X				X	X			
ACM2-Revised MM5			X		X		X			
YSU-Revised MM5	X				X		X			
Noah- MP	X			X				X		
CLM4	X			X					X	
RUC	X			X						X

2.3. Implemented models

2.3.1. WRF modeling system

The meteorological model employed in the present study is the WRF model, version 3.5.1 (Skamarock et al., 2008). The numerical simulations were performed over two 2-way nested modeling domains (Fig. 1a). The coarse domain (d01) covers most of Europe and North Africa with a spatial resolution of 18 km (mesh size of 320 × 220) to capture the HWs synoptic-scale conditions. The innermost domain (d02) focuses on the study area of the SE Mediterranean and BP with a horizontal grid resolution of 6 km (mesh size of 250 × 190). Each domain has the same vertical structure that includes 28 unevenly spaced full sigma layers from the lowest layer ($\sim$ 15 m above ground level) to the model top that defined at 100 hPa. The WRF simulations were driven by initial and boundary conditions based on the operational surface and upper-level atmospheric analysis data of the European Center for Medium-range Weather Forecasts (ECMWF), which provided at 0.25°x0.25° spatial and 6 h temporal resolution (Trenberth, 1992). The land use/land cover (LULC) for both domains was represented using the modified MODIS/IGBP (moderate resolution imaging spectroradiometer/international geosphere-biosphere project) global dataset in 30-arc-*sec* spatial resolution. One 84-h numerical simulation was implemented for each 2- and 3-day HW event, while two 84-h simulations were performed in the case of 4- and 5-day HWs. All simulations were initialized at 1200 UTC and provided output every hour. The first 12 h of each simulation were treated as spin-up period, while the remaining hours that corresponded to the observed HW days were used for evaluating the performance of the model.

The model physics used in all domains and during all simulations include the WRF single-moment six-class scheme for parametrizing the microphysics processes (Hong and Lim, 2006), the Dudhia shortwave (SW) scheme (Dudhia, 1989) and the Eta geophysical fluid dynamics laboratory (GFDL) longwave (LW) parameterization (Schwarzkopf and Fels, 1991) to represent the SW and LW radiation processes, respectively, and the Kain-Fritch scheme (Kain, 2004) for the convective parameterization. For the planetary boundary layer, surface layer and land surface processes, a series of seven numerical experiments were carried out by changing the PBL-SL and LSM options. An analytical description of the experimental design and the physics parameterizations tested is presented in Sec. 2.4.1.

2.3.2. RayMan model

RayMan is a numerical model that simulates short- and long-wave radiation fluxes absorbed by the human body in simple and complex environments. The model requires only primary meteorological variables (e.g., 2-m air temperature) for the simulation of these fluxes, which in turn enable the calculation of the mean radiant temperature (Tmrt) that parameterize the heat impact of the fluxes on the human energy balance. Tmrt is then used in the computation of human thermal comfort indices, such as the physiologically equivalent temperature (PET; Matzarakis et al., 2007, 2010). PET was introduced by Mayer and Höppe (1987) and its calculation is based on the human energy balance

model MEMI (Munich energy-balance model for individuals; Höppe, 1984). It is a universal index that used for assessing the thermal conditions of the human body (Höppe, 1999) and has been widely applied in several past human-biometeorological studies (Amengual et al., 2014; Daneshvar et al., 2013; Giannaros et al., 2015; Giannaros et al., 2018b; Matzarakis and Endler, 2010; Muthers et al., 2010). Table S3 (see supplementary material) shows the matching between the PET values and the nine levels of human thermal stress (Matzarakis and Mayer, 1996). In the present work, the RayMan model was driven by the ground-based observations and WRF-simulated data to calculate the PET at the locations of the weather stations during the examined HW episodes over SE Mediterranean and BP. The meteorological data used include the air temperature, wind speed, vapor pressure and cloud cover.

2.4. Experimental design and evaluation methodology

2.4.1. Sensitivity experiments

Table 1 summarizes the seven numerical experiments performed in the present study. The YSU-MM5, MYJ-Eta, ACM2-Revised MM5 and YSU-Revised MM5 scenarios refer to the sensitivity analysis of the simulated HWs to the PBL-SL physics. The Yonsei University (YSU) PBL scheme (Hong et al., 2006) is classified as a first-order non-local closure model based on a parabolic K-profile for the convective boundary layer and with an explicit treatment of the entrainment at the top of the PBL. The YSU algorithm uses non-local eddy diffusivity coefficients to calculate the turbulent fluxes, while the PBL height is determined using the Richardson bulk number (Ri_b) method beginning from the surface. The Mellor-Yamada-Janjić (MYJ) PBL parameterization (Janjić, 1994) is a local closure scheme that computes eddy diffusion coefficients from an additional (1.5-order) prognostic equation for the turbulent kinetic energy (TKE). The local vertical mixing is applied from the lowest to the highest vertical model level for both day and night boundary layer conditions and the PBL height is estimated at the level where the TKE decreases below a critical value. The Asymmetrical Convective Model, version 2 (ACM2) PBL option (Pleim, 2007a, 2007b) is a first-order closure scheme that combines local downward mixing (eddy diffusion) and non-local upward transport (large scale convection). This consolidation is weighted with a parameter depending on the stability with the non-local approach switching off during stable or neutral conditions. The PBL height is calculated using a critical value for the computed above the level of neutral buoyancy Ri_b. Concerning the surface-layer parameterizations, both MM5 and Eta scheme are based on the Monin-Obukhov similarity theory (MOST; Monin and Obukhov, 1954). In the MM5 algorithm, the surface exchange coefficients for heat, moisture, and momentum are calculated based on the empirical stability functions of Paulson (1970), Dyer and Hicks (1970), and Webb (1970). Four stability regimes are considered based on Zhang and Anthes (1982) and a convective velocity following Beljaars (1995) is utilized to strengthen the surface heat and moisture fluxes, while no thermal roughness length parameterization is included. Jiménez et al. (2012) introduced a revised version (Revised-MM5) of this scheme that

leads to a more consistent surface-layer formulation for various atmospheric stability conditions. The Revised-MM5 formulation utilizes more sufficient similarity functions to replicate the surface-layer during strong stable conditions (Dimitrova et al., 2014). The Eta SL scheme employs an iterative method to compute the surface fluxes and includes parameterizations of a viscous sub-layer (Zilitinkevich, 1995; Janjić, 1994). Moreover, the Beljaars (1995) correction is used for avoiding singularities in the case of an unstable surface-layer and vanishing wind speed.

The Noah, Noah-MP, CLM4 and RUC experiments refer to the sensitivity analysis of the modeled HW episodes to the land surface parameterizations. The Noah LSM (Tewari et al., 2004; Chen and Dudhia, 2001) predicts the temperature and moisture contents in four soil layers by applying the force-restore method. The thickness of layers are: 10, 30, 60, and 100 cm from top to bottom, respectively. The model also computes the skin temperature, canopy water content and energy flux terms of the surface energy budget (SEB). It includes detailed descriptions of the vegetation and hydrological processes, such as the evapotranspiration, soil drainage and run-off. The Noah-MP LSM (Niu et al., 2011; Yang et al., 2011) is based on the Noah model and introduces various augmentations in the initial model (e.g., a vegetation canopy layer to calculate the canopy and ground surface temperatures separately). It also uses a framework for multiple options to parameterize the key land surface processes (e.g., dynamic vegetation option). In the current study, the additional options for the Noah-MP LSM were set to their default values. The Rapid Update Cycle (RUC) LSM (Benjamin et al., 2004) incorporates evapotranspiration, soil moisture diffusion, precipitation and run-off processes. It treats the vegetation processes similarly as the Noah LSM and computes the soil moisture and temperature at six layers of 0, 5, 20, 40, 160 and 300 cm thickness from top to bottom, respectively. A special feature of the RUC model is that the first layer (0 cm) includes half of the first atmospheric layer and half of the first soil layer. The Community Land Model, version 4 (CLM4, Lawrence et al., 2011) has a vertical structure of ten unevenly spaced soil layers. It contains sophisticated treatment of biogeophysics, hydrology, biogeochemistry and dynamic vegetation. A limitation in the coupling of the CLM4 LSM with the WRF model is that the WRF input LULC data must be taken from the 30 arc-*sec* horizontal resolution US geological survey (USGS) global dataset.

2.4.2. Evaluation methods

The evaluation of each sensitivity experiment focused on the: (a) 2-m air temperature (T2, °C), (b) 10-m wind speed (WS10, m/s), (c) 2-m vapor pressure (VP2, hPa), derived from the 2-m relative humidity (%) and T2 values, and (d) PET (°C).The WRF-simulated data and near-surface observations were paired in time and space using the "nearest neighbor" technique. According to this approach, the model grid point nearest to the location of the measuring site was selected for extracting the WRF data. Using the model-observation hourly pairs for identified HW days, key statistical measures were calculated including the mean bias (MB), mean absolute error (MAE), and index of agreement (IOA; Tables 2-5). The geographical dependence of the model performance was investigated through the spatial distribution of the MB error for each experiment (Figs. 2-7). A thermal bioclimate diagram showing the occurrence frequency of PET classes was constructed for each experiment by employing the observational and modeled data at total-network scale (Figs. 8 and 11).The analysis of the daytime (1200 UTC) and nighttime (0000 UTC) T2 and PET values is presented through time series of of the variables during the 57 studied HW days (Figs. 9-10 and 12-13).

Table 2
Aggregated model performance statistics for the PBL-SL experiments concerning the T2, WS10, and VP2.

	Geographical sub-regions	YSU-MM5	MYJ –Eta	ACM2-Revised MM5	YSU-Revised MM5
MB					
T2 (°C)	Total	−0.98	−1.06	−0.70	−0.82
	Coastal	−0.27	−0.32	0.07	−0.09
	Inland	−1.24	−1.34	−0.96	−1.09
	Mountain	−1.30	−1.35	−1.05	−1.14
WS10 (m/s)	Total	−0.38	0.17	−0.11	−0.28
	Coastal	−0.42	0.16	−0.16	−0.31
	Inland	−0.59	−0.1	−0.34	−0.51
	Mountain	−0.01	0.59	0.29	0.09
VP2 (hPa)	Total	0.04	0.52	−0.36	0.13
	Coastal	−3.41	−3.20	−3.95	−3.31
	Inland	−0.83	−0.38	−1.19	−0.72
	Mountain	4.72	5.50	4.40	4.79
MAE					
T2 (°C)	Total	2.31	2.37	2.27	2.30
	Coastal	2.02	2.14	2.04	2.04
	Inland	2.36	2.39	2.29	2.33
	Mountain	2.51	2.57	2.46	2.50
WS10 (m/s)	Total	1.64	1.71	1.64	1.63
	Coastal	1.58	1.56	1.54	1.55
	Inland	1.67	1.68	1.65	1.66
	Mountain	1.66	1.91	1.73	1.66
VP2 (hPa)	Total	4.77	5.03	4.81	4.79
	Coastal	4.82	4.71	5.12	4.80
	Inland	4.04	4.26	4.09	4.05
	Mountain	5.83	6.53	5.61	5.89
IOA					
T2 (°C)	Total	0.91	0.91	0.91	0.91
	Coastal	0.91	0.91	0.91	0.91
	Inland	0.90	0.90	0.90	0.90
	Mountain	0.91	0.91	0.91	0.91
WS10 (m/s)	Total	0.67	0.68	0.68	0.67
	Coastal	0.68	0.70	0.69	0.67
	Inland	0.68	0.71	0.69	0.67
	Mountain	0.65	0.62	0.64	0.64
VP2(hPa)	Total	0.64	0.60	0.64	0.64
	Coastal	0.64	0.63	0.62	0.64
	Inland	0.72	0.69	0.72	0.73
	Mountain	0.53	0.49	0.55	0.53

3. Results and discussion

3.1. Statistical evaluation

3.1.1. PBL-SL schemes

As can be seen in Table 2, all experiments exhibit a satisfactory correlation (IOA = 0.91 overall) between the modeled and observed 2-m air temperature. The model is biased cold underestimating T2 in terms of total MB (MAE) by 0.70 °C (2.27 °C) in the ACM2-Revised MM5 experiment to 1.06 °C (2.37 °C) in the MYJ-Eta scenario. The largest cold bias produced by the MYJ-Eta experiment may be associated with the weakened vertical mixing and entrainment in the MYJ PBL algorithm caused by its inability to simulate large-scale eddies (García-Díez et al., 2013). All scenarios produce higher errors as the altitude increases from the coastal to mountain sites. Akylas et al. (2007) showed that the forecast accuracy for temperature rises when higher horizontal grid resolution is applied because of the more accurate replication of the terrain elevation. Thus, the geographical dependence of the T2 errors in the present study could be partly attributed to possible inadequate representation of the complex terrain in the high-altitude regions due to the coarse spatial resolution.

The 10-m wind speed is overall slightly underpredicted in all sensitivity scenarios, except in MYJ-Eta. The most significant total MAE (1.71 m/s) is found for this experiment, while the ACM2-Revised MM5

Table 3
Aggregated model performance statistics for the LSM experiments concerning the T2, WS10 and VP2.

	Geographical sub-regions	Noah	Noah-MP	CLM4	RUC
MB					
T2 (°C)	Total	−0.98	−1.54	−1.24	−0.39
	Coastal	−0.27	−0.03	−0.54	−0.04
	Inland	−1.24	−1.78	−1.37	−0.58
	Mountain	−1.30	−1.78	−1.72	−0.44
WS10 (m/s)	Total	−0.38	−0.40	0.06	−0.29
	Coastal	−0.42	−0.44	0.00	−0.33
	Inland	−0.59	−0.64	−0.13	−0.54
	Mountain	−0.01	0.01	0.41	0.14
VP2 (hPa)	Total	0.04	0.63	1.33	−1.02
	Coastal	−3.41	−2.76	−1.98	−4.80
	Inland	−0.83	−0.29	0.42	−1.86
	Mountain	4.72	5.34	5.94	3.93
MAE					
T2 (°C)	Total	2.31	2.53	2.68	1.97
	Coastal	2.02	2.23	2.32	1.89
	Inland	2.36	2.56	2.57	1.96
	Mountain	2.51	2.78	3.19	2.08
WS10 (m/s)	Total	1.64	1.70	1.78	1.63
	Coastal	1.58	1.61	1.63	1.57
	Inland	1.67	1.71	1.81	1.60
	Mountain	1.66	1.77	1.88	1.71
VP2 (hPa)	Total	4.77	5.05	5.24	5.09
	Coastal	4.82	4.78	4.66	5.65
	Inland	4.04	4.42	4.54	4.36
	Mountain	4.72	6.28	6.89	5.66
IOA					
T2 (°C)	Total	0.91	0.90	0.90	0.93
	Coastal	0.91	0.90	0.90	0.93
	Inland	0.90	0.89	0.90	0.93
	Mountain	0.91	0.90	0.88	0.94
WS10 (m/s)	Total	0.67	0.65	0.64	0.68
	Coastal	0.68	0.66	0.66	0.69
	Inland	0.68	0.66	0.65	0.69
	Mountain	0.65	0.61	0.59	0.66
VP2(hPa)	Total	0.64	0.61	0.61	0.62
	Coastal	0.64	0.63	0.65	0.60
	Inland	0.72	0.68	0.68	0.71
	Mountain	0.53	0.51	0.48	0.51

Table 4
Aggregated model performance statistics for the PBL-SL experiments concerning the PET.

	Geographical sub-regions	YSU-MM5	MYJ –Eta	ACM2-Revised MM5	YSU-Revised MM5
MB					
PET (°C)	Total	−1.22	−1.65	−1.09	−1.10
	Coastal	−0.56	−1.03	−0.40	−0.45
	Inland	−1.37	−1.77	−1.25	−1.25
	Mountain	−2.09	−2.54	−2.04	−1.99
MAE					
PET (°C)	Total	3.42	3.61	3.46	3.43
	Coastal	3.26	3.45	3.32	3.28
	Inland	3.27	3.44	3.31	3.26
	Mountain	3.95	4.20	3.96	3.98
IOA					
PET (°C)	Total	0.94	0.94	0.94	0.94
	Coastal	0.94	0.94	0.94	0.94
	Inland	0.95	0.95	0.95	0.95
	Mountain	0.93	0.93	0.93	0.93

Table 5
Aggregated model performance statistics for the LSM experiments concerning the PET.

	Geographical sub-regions	Noah	Noah-MP	CLM4	RUC
MB					
PET (°C)	Total	−1.22	−1.84	−1.66	−0.62
	Coastal	−0.56	−1.19	−0.94	−0.33
	Inland	−1.37	−2.05	−1.85	−0.59
	Mountain	−2.09	−2.61	−2.61	−1.17
MAE					
PET (°C)	Total	3.42	3.70	3.98	3.04
	Coastal	3.26	3.46	3.64	3.02
	Inland	3.27	3.59	3.75	2.80
	Mountain	3.95	4.34	4.98	3.50
IOA					
PET (°C)	Total	0.94	0.94	0.93	0.95
	Coastal	0.94	0.94	0.94	0.95
	Inland	0.95	0.94	0.94	0.96
	Mountain	0.93	0.92	0.91	0.94

Hence, it is evident that the tendency of biases depends on the terrain geomorphology, as also concluded by Santos-Alamillos et al. (2013). The cause of the wind speed overestimation in the high-altitude locations may be associated with variations in the sub-grid surface roughness inducing local wind patterns that are not simulated accurately because of the low horizontal grid resolution (Avolio et al., 2017; Sathyanadh et al., 2017).

The YSU PBL scheme yields overall the best scores for the 2-m vapor pressure with the ACM2 PBL option close second. The high MAE values for the VP2, ranging from 4.77 hPa (YSU-MM5) to 5.03 hPa (MYJ-Eta), indicate that the model lacks ability in simulating the magnitude of this variable (Table 2). Also, the relatively low VP2 IOA values suggest model errors in phase. The largest total moist bias generated by the MYJ-Eta scenario is in agreement with the highest cold bias also delivered by the same experiment indicating that the MYJ vertical mixing is not sufficient to entrain drier and warmer air from above the PBL (Hu et al., 2010). On the other hand, the local ACM2 PBL algorithm seems to be characterized by stronger vertical mixing, as also reported in other studies (Avolio et al., 2017; García-Díez et al., 2013; Hu et al., 2010), that contributes to the drier and warmer near-surface atmosphere modeled in the ACM2-Revised MM5 experiment. Concerning the geographical sub-regions, the model underestimates the VP2 over the coastal and inland regions. The degree of the underestimation is higher in the coastal sites, with MAE ranging from 4.71 hPa (MYJ-Eta) to 5.12 hPa (ACM2-Revised MM5). In contrast, significant overpredictions reaching up to 6.53 hPa (MYJ-Eta) regarding MAE are found for the VP2 over the mountain areas (Table 2). Moisture in the PBL originates mainly from the evapotranspiration near the surface (García-Díez et al., 2013; Hu et al., 2010). Thus, the moist biases in the high altitudes regions may associated with a possible overestimation of the latent heat fluxes in these sites in all PBL-SL scenarios.

Fig. 2 illustrates that all experiments tend to underestimate the 2-m air temperature in most regions of the southeast Mediterranean and BP. The largest underpredictions occur in the island part of Greece and west inland of Turkey with mean biases as high as −4.5 °C. The use of the Revised MM5 SL scheme reduces significantly the biases compared to the other two SL physics over the majority of the examined sites in the study area (Fig. 2c, d). This is because the T2 variability is not only dependent on the parameterization of the PBL vertical turbulent fluxes, but it is also highly affected by the surface-layer formulations (Shin and Hong, 2011). In particular, the MB ranges from −4.14 °C to 1.61 °C during the YSU-Revised MM5 scenario (Fig. 2d), and it takes values between −4.11 °C and 1.73 °C in the ACM2-Revised MM5 experiment (Fig. 2c).

The spatial distribution pattern of the 10-m wind speed MBs is quite similar for all PBL-SL numerical experiments (Fig. 3). A relatively

scenario produces the lowest deviations from the observations over the coastal (MAE = 1.54 m/s) and inland (MAE = 1.65 m/s) regions (Table 2). Positive biases are found in the mountain areas for all experiments with the MYJ scheme exhibiting the highest MB (0.59 m/s).

Fig. 2. Spatial distribution of T2 MBs (°C) for the (a) YSU-MM5, (b) MYJ-Eta, (c) ACM2-Revised MM5 and (d) YSU-Revised MM5 numerical experiments.

strong overestimation of the WS10 is evident mainly in the island regions of Greece, in most of the high-altitude FYROM sites, and in the west (mountain) part of Bulgaria, whereas the model underpredicts the WS10 over Italy and Turkey. The local PBL scheme (MYJ) produces the highest overestimations with mean biases reaching up to 3.4 m/s (Fig. 3b). On the other hand, the ACM2 PBL algorithm provides the best results for the MB spanning from −1.9 m/s to 2.6 m/s (Fig. 3c). These findings are in agreement with the conclusion of Xie et al. (2012), according to which the local schemes overpredict the WS10 more severely. A factor that may contribute to the WS10 overestimation is a possible underestimation of friction velocity, which is a parameter that calculated by the surface-layer schemes (Kim et al., 2013; Shin and Hong, 2011).
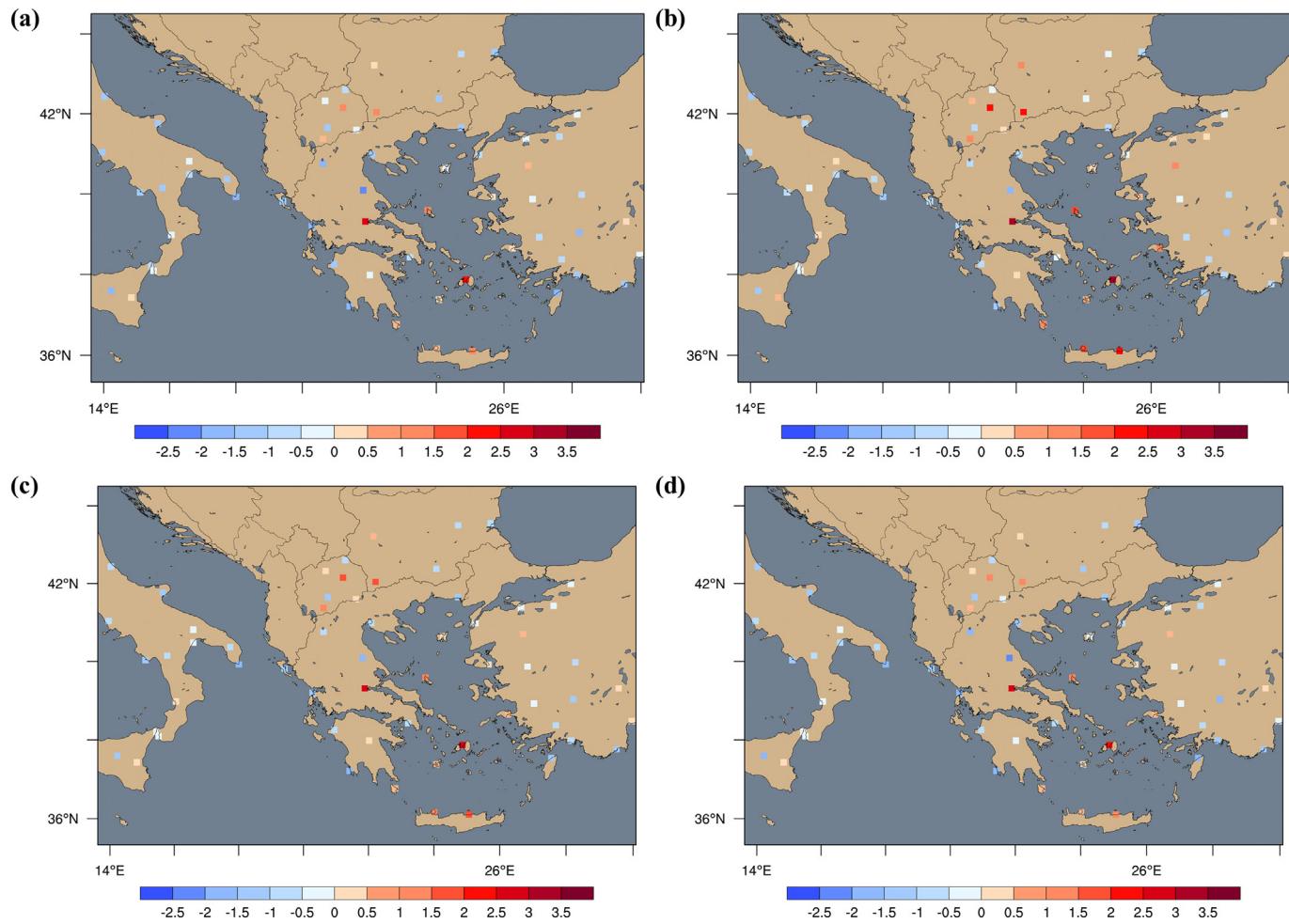
All experiments show a tendency to simulate drier near-surface conditions compared to the observed ones in the majority of the examined locations (Fig. 4). However, the magnitude of the negative biases is notably lower than that of the positive ones in all scenarios, except in ACM2-Revised MM5 (Fig. 4c). The MYJ PBL scheme produces the highest moist MBs, especially over Turkey (Fig. 4b), in contrast with the ACM2 PBL option that reduces the VP2 overestimations significantly (Fig. 4c). The latter PBL scheme underestimates the VP2 by 1.2–5 hPa and by 0.9–7.3 hPa in most regions of Greece and Italy, respectively, and by 0.7 hPa to 4.5 hPa in parts of FYROM and Bulgaria. These underestimations contribute to the overall negative MB reported in Table 2 for the ACM2-Revised MM5 scenario. This finding supports the above-described diagnosis indicated that the ACM2 PBL parameterization is characterized by strong vertical mixing compared to the other PBL schemes. The strengthen ACM2 vertical mixing enhances the

upward surface fluxes and the entrainment of free-tropospheric air into the mixing layer leading to a warmer and drier modeled near-surface atmosphere.

3.1.2. Land surface modes

Table 3 highlights that the use of the RUC LSM reduces the T2 model errors remarkably. The RUC scenario underestimates the observed temperatures regarding overall MB by 0.39 °C, whereas the Noah-MP experiment produces the highest total MB of −1.54 °C. The RUC MAEs are the lowest showing little variation for the different geographical sub-regions, as they span from 1.89 °C (coastal) to 2.08 °C (mountain). The Noah experiment yields the second best scores in terms of MAE, whereas the CLM4 scenario results in the greatest T2 deviations from the observations, which exceed 3 °C over the mountain areas (MAE = 3.19 °C). Small and probably insignificant differences are found between the tested LSMs for the IOA, even though the RUC model exhibits the highest IOA values (0.93 overall, Table 3).

The correlation coefficients for the 10-m wind speed are lower than that for T2 varying in total from 0.64 (CLM4) to 0.68 (RUC; Table 3). This reveals that the observed diurnal variation of the WS10 is not well reproduced by the model. The CLM4 LSM exhibits the lowest MB overall (0.06 m/s), even though it produces the highest total MAE (1.78 m/s). This is because the CLM4 experiment slightly underpredicts the WS10 in the lower terrain heights (coastal and inland), whereas it produces the highest positive MB (MAE) of 0.41 m/s (1.88 m/s) over the mountain areas. A positive MB also occurs in the mountain regions for the RUC experiment. However, the RUC LSM produces the lowest WS10 underestimations over the coastal and inland locations (1.57 m/s

Fig. 3. Spatial distribution of WS10 MBs (m/s) for the (a) YSU-MM5, (b) MYJ-Eta, (c) ACM2-Revised MM5, (d) and YSU-Revised MM5 numerical experiments.

and 1.60 m/s, respectively, regarding MAE). In the same areas, the Noah LSM produces the second lowest MAEs, while it shows the minimum deviations from the observations over the mountain sites (MAE = 1.66 m/s; Table 3).
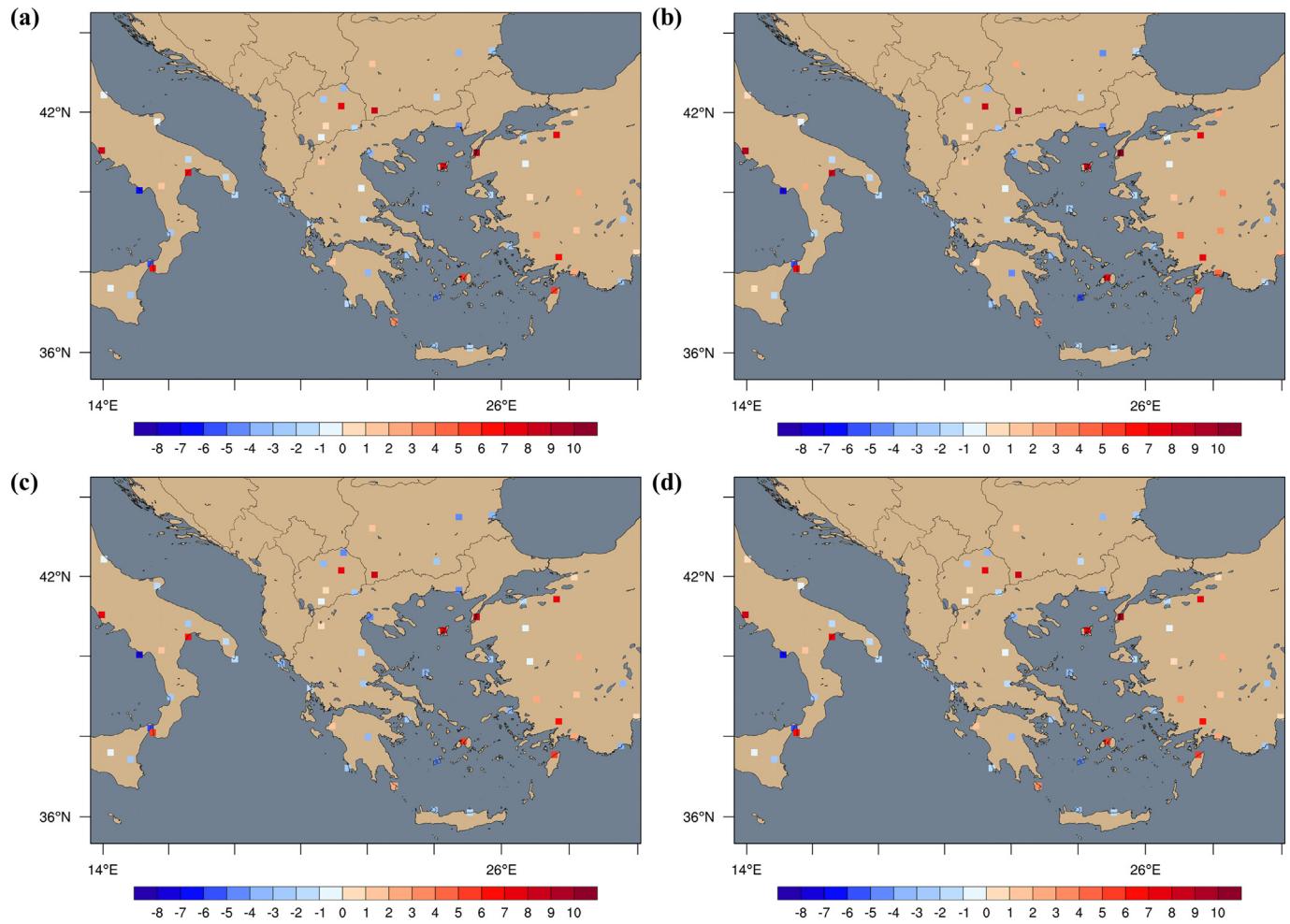
Concerning the 2-m vapor pressure, all experiments lack ability in capturing the magnitude and phase of this variable, as in the case of the PBL-SL scenarios. This is indicated by the total MAEs that are > 4.5 hPa in all scenarios and the rather low indices of agreement ranging from 0.61 (Noah-MP and CLM4) to 0.64 (Noah) in overall (Table 3). Dry biases are evident in the coastal and inland areas for all LSMs, except CLM4 over the inland regions. Contrary, the WRF is biased moist in all sensitivity experiments in the mountain sites. The RUC and CLM4 models exhibit the largest MAEs over the coastal (5.65 hPa) and inland (4.54 hPa) areas, respectively. The highest VP2 overestimations over the mountain regions in terms of MAE are found for the Noah-MP (6.28 hPa) and CLM4 (6.89 hPa) experiments. The Noah LSM exhibits the lowest VP2 deviations from the observations over the inland (MAE = 4.04 hPa) and mountain (MAE = 4.72 hPa) stations (Table 3).

The above-mentioned marked minimization of the T2 model biases (Table 3) by the use of the RUC LSM is also evident in Fig. 5.The maximum cold MBs during the RUC experiment (Fig. 5d) reach up to −3.3 °C, whereas the three other scenarios produce larger under-estimations of T2 as high as −4.5 °C. Additionally, the highest T2 un-derpredictions (> 3 °C) by the RUC LSM are limited mainly in the is-land part of Greece, in contrast with the rest of the experiments that show great negative MBs (> 4 °C) over more areas in the study region. Thus, the RUC LSM is capable of capturing the magnitude and spatial allocation of the abnormal high temperatures over the majority of the

stations in the southeast Mediterranean and BP. The T2 variability between simulations with different land surface parameterizations is dominated by the different LSM-induced sensible heat fluxes (Ma et al., 2017; Zeng et al., 2011, 2015). Hence, the consistentcy between the RUC-modeled and the observed temperatures is expected to arise from the higher RUC-simulated sensible heat fluxes compared to the other three LSM experiments. The examination of this hypothesis is presented in Sec. 3.3.

Concerning the WS10, negative biases are occurred in most of the domain sites, especially over Italy and Turkey (Fig. 6). The MB dis-tribution pattern is quite similar for all experiments. However, a dif-ferent magnitude of errors is evident. The range of the MBs varies from −2.4 and − 2.6 m/s during the Noah scenario (Fig. 6a) and from −2.6 to − 3.2 m/s in the RUC experiment (Fig. 6d). Thus, it seems that the Noah LSM yields a slightly better performance in replicating the WS10, even though the RUC LSM produced the lowest MB and MAE in overall (Table 3). Overestimations of the 10-m wind speed are found primarily in FYROM and in the west part of Bulgaria over locations with high altitude. As already mentioned in Sec. 3.1.1, this could be partially associated with the coarse grid resolution that is may not capable of capturing the sub-grid terrain-induced wind patterns.

The VP2 MBs show a high spatial variation (Fig. 7). The RUC LSM tends to underestimate the VP2 in the majority of the weather stations in the study area, with the mean biases ranging from −8.5 hPa to 8.6 hPa (Fig. 7d). The rest of the LSMs produces mainly positive MBs with CLM4 scenario exhibiting the largest VP2 overestimations that reach up to nearly 10 hPa (Fig. 7c). The simulated VP2 in the lower atmosphere is greatly affected by the vegetation and hydrology

Fig. 4. Spatial distribution of VP2 MBs (hPa) for the (a) YSU-MM5, (b) MYJ-Eta, (c) ACM2-Revised MM5, (d) and YSU-Revised MM5 numerical experiments.

processes parameterized in the LSMs. These procedures, in turn, are highly influenced by the LSM-induced moisture content in the ground (Stéfanon et al., 2014). Hence, the overall dry conditions modeled by the RUC LSM could be attributed to a possible underprediction of the soil moisture. This would be in aggrement with the expected RUC-induced high values of sensible heat fluxes because low soil moisture leads to less evapotranspiration, which in turn results in the enhancement (declination) of the sensible (latent) heating (Stéfanon et al., 2014; Ma et al., 2017).

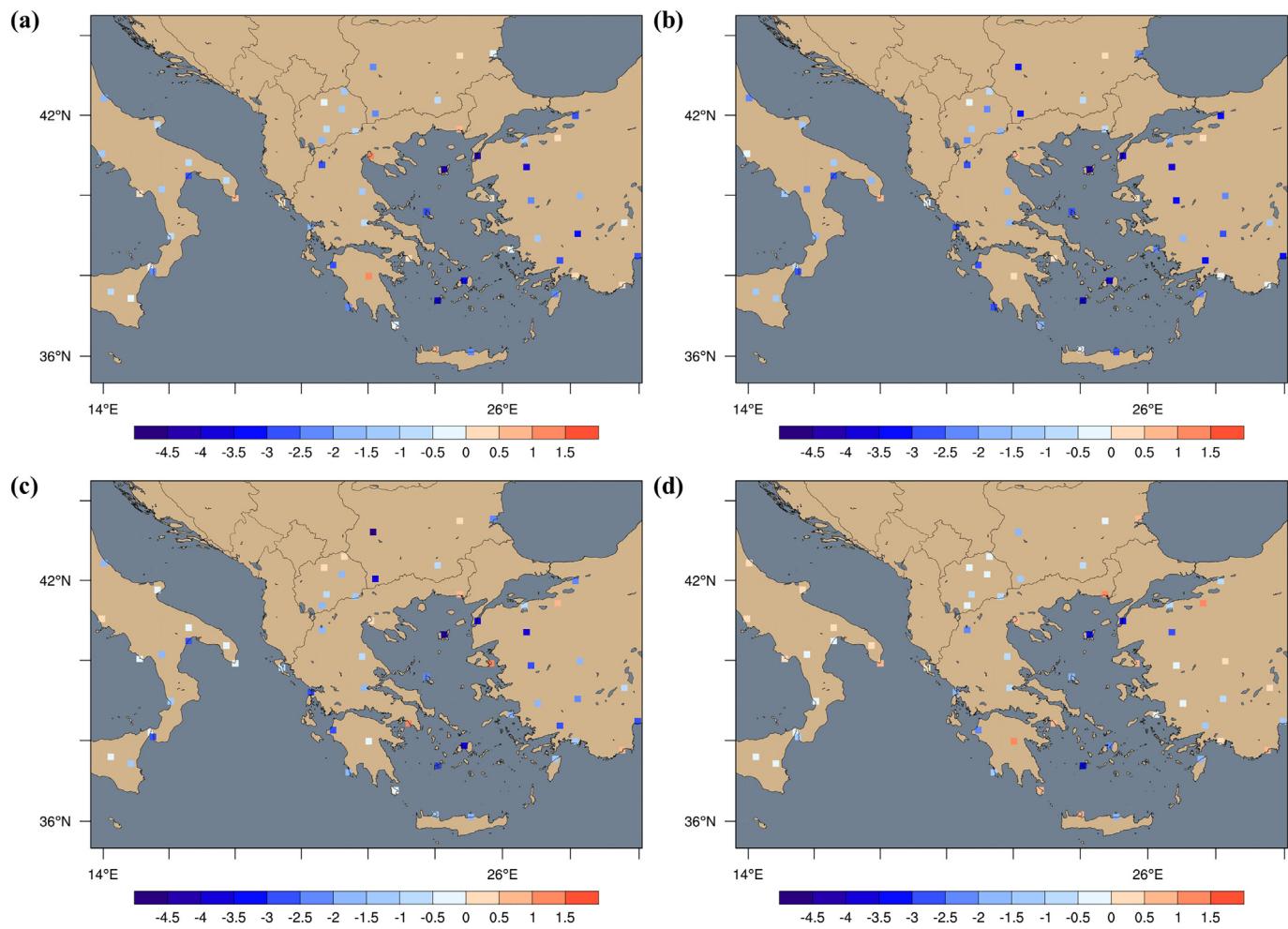### 3.2. Evaluation of thermal bioclimate

3.2.1. PBL schemes

Table 4 shows that the observed PET values are underestimated by the numerical simulations. This can be primarily attributed to the WRF cold bias in simulating the 2-m air temperature (Table 2). The MAE values show little variation between the YSU-MM5, YSU-Revised MM5 and ACM2-Revised MM5 experiments over all geographical sub-regions, while no differences occur between all scenarios for the IOA in overall. The model performs better over the coastal and inland regions compared to the high-altitude locations, where the MAE can be as high as 4.20 °C (MYJ-Eta; Table 4). In overall, all experiments, except MYJ-Eta over the mountain areas, produce deviations from the observations that are lower than the PET reference intervals (see supplementary material Table S3). This means that the simulated PET may only be wrong by one reference scale, as illustrated in the thermal bioclimate diagram for the PBL-SL scenarios (Fig. 8).

As can be seen in Fig. 8, the observed PET values under extreme

heat stress conditions (PET > 41 °C) reach up to 15% highlighting the severe impact of HWs on the human thermal comfort conditions. The model underestimates the class of PET > 41 °C by approximately 4–4.5% during all scenarios. The frequencies of 29 °C < PET < 41 °C are slightly overestimated by the simulations, with the ACM2-Revised MM5 experiment providing the lowest deviations from the observations (~ 0.2% and ~ 1.3% for moderate and strong heat stress, respectively). Slight heat stress conditions (23 °C < PET < 29 °C) are also modeled better in the same scenario underestimating the observed PET values by ~ 3.2%. PET values of "No thermal stress" class (18 °C < PET < 23 °C) are overestimated by ~ 1.6% during the MYJ-Eta scenario to ~ 2.4% in the ACM2-Revised MM5 experiment. Low percentages of cool thermal perception (8 °C < PET < 18 °C) are found, even though the study period refers to HW events. This is associated with the high-altitude areas, where the temperature drops significantly in the nighttime (Fig. 9c) resulting in slight (13 °C < PET < 18 °C) and moderate (8 °C < PET < 13 °C) cold stress (Fig. 9). The model successfully captures the magnitude of the latter PET class in all sensitivity scenarios, while it overestimates the observed frequency of 13 °C < PET < 18 °C with the ACM2-Revised MM5 experiment producing the lowest deviations (~ 3.2%).

Fig. 9a, b show that the observed nighttime temperatures are above 25 °C in the majority of the studied HW days over the coastal and inland regions. In the same areas, most of the observed daytime temperatures are higher than 33 °C reaching up to 39 °C some days (Fig. 10a, b). These features reveal the intensity of the examined HW episodes, which is less pronounced in the mountain locations, especially during the night (Fig. 9c). The high-temperature weather directly affects the mid-

(a)

(b)

(c)

(d)

Fig. 5. Spatial distribution of T2 MBs (°C) for the (a) Noah, (b) Noah-MP, (c) CLM4, and (d) RUC numerical experiments.
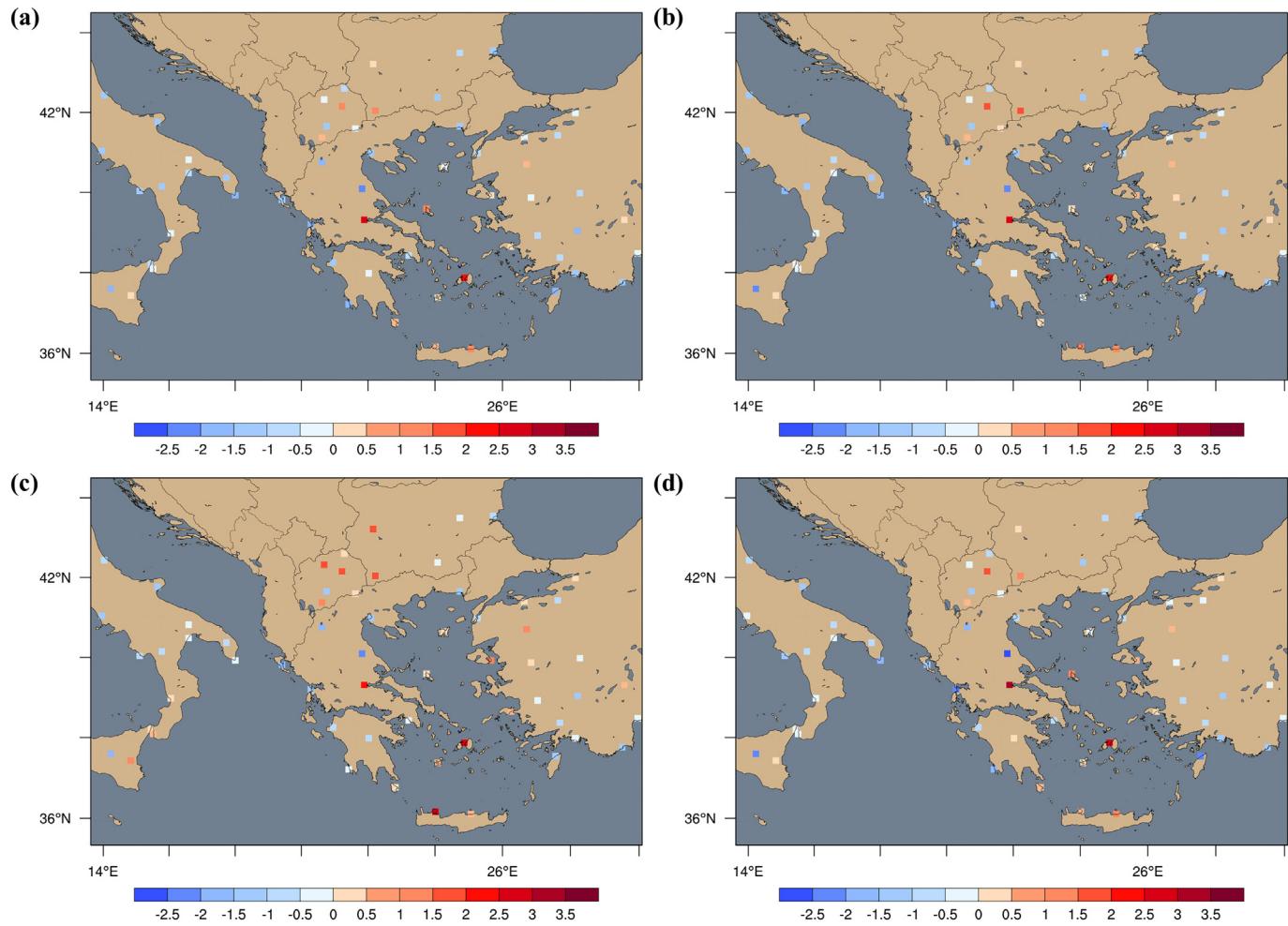
day human thermal comfort conditions, as the PET values exceed 35 °C in most of the HW days over all geographical sub-regions (Fig. 10d-f) denoting a marked heat stress level (i.e., strong and extreme heat stress conditions). The inland areas (Fig. 10e) experience extreme heat stress conditions (PET > 41 °C) more frequently than the coastal ones (Fig. 10d) suggesting that the seaside wind patterns reduce the severe PET values, as also reported in previous studies (Giannaros et al., 2015; Matzarakis and Mayer, 1997). Particularly during HW events, the greater temperatures over the land compared to the sea enhance the sea-breeze circulations in the coastal regions (Stéfanon et al., 2014). Thus, the lower daytime T2 values over the coastal sites (Fig. 10a) that result to less extreme heat stress conditions (Fig. 10d) compared to the inland locations (Fig. 10e) could be attributed to the advection of cool marine air. In the nighttime, the thermal perception is notably lower over all geographical sub-regions (Fig. 9d-f). No thermal stress conditions (18 °C < PET < 23 °C) are prevailing, while PET values between 8 °C and 18 °C (moderate/slight cold stress) are observed in the mountain areas.

The model mostly underpredicts the observed daytime (Fig. 9) and nighttime (Fig. 10) T2 and PET values in all sensitivity experiments. These underestimations are more considerable over the inland and mountain areas at 1200 UTC (Fig. 10b, c), whereas the degree of them is lower in the same regions at the nighttime (Fig. 9b, c) and in the coastal areas during both day and night (Figs. 10a and 9a). It is interesting to see that the MYJ-Eta scenario shows the worst performance under the stable (nocturnal; 0000 UTC) conditions (Fig. 9a-c), whereas it produces the lowest deviations from the observations during the mid-day convective environment (1200 UTC; Fig. 10a-c). This finding

indicates that greater LSM-calculated sensible heat fluxes are produced in the day during the MYJ-Eta runs that contribute directly to the simulation of higher daytime T2 values compared to the rest of the experiments. Minor differences occur among the YSU- and ACM2-based numerical scenarios during the day (Fig. 10), while the ACM2-Revised MM5 experiment yields the closest to the observations results in the night (Fig. 9). The Pearson correlation coefficients in Fig. 10 indicate that the model simulates adequately the variation of the daytime T2 and PET values throughout the examined HW days over all altitudinal classes. The same finding applies to the nighttime T2 and PET variations over the coastal and inland areas (Fig. 9a, b, d, e), whereas a lower correspondence ($R \leq 0.70$) between the modeled and observed T2 and PET values occurs over the mountain regions at 0000 UTC (Fig. 9c, f). In overall, the T2 and PET are correlated well, as the two variables follow the same fluctuations during both night (Fig. 9) and day (Fig. 10). This points out that the near-surface air temperature has a significant influence on the PET, as Giannaros et al. (2015) also denoted.

3.2.2. Land surface models

The previously identified improvement of the model performance in replicating the observed T2 values in the RUC scenario (Sec. 3.1.2) is reflected in the cumulative PET statistical metrics for the LSM experiments (Table 5). The RUC experiment produces the lowest MB (−0.62 °C) and MAE (3.04 °C) in overall, as well as in all geographical sub-regions. It also yields the highest IOA values, even though the differences between the sensitivity experiments for this measure are small and probably insignificant. The Noah LSM yields the second best statistical scores in total and over all examined areas, whereas the
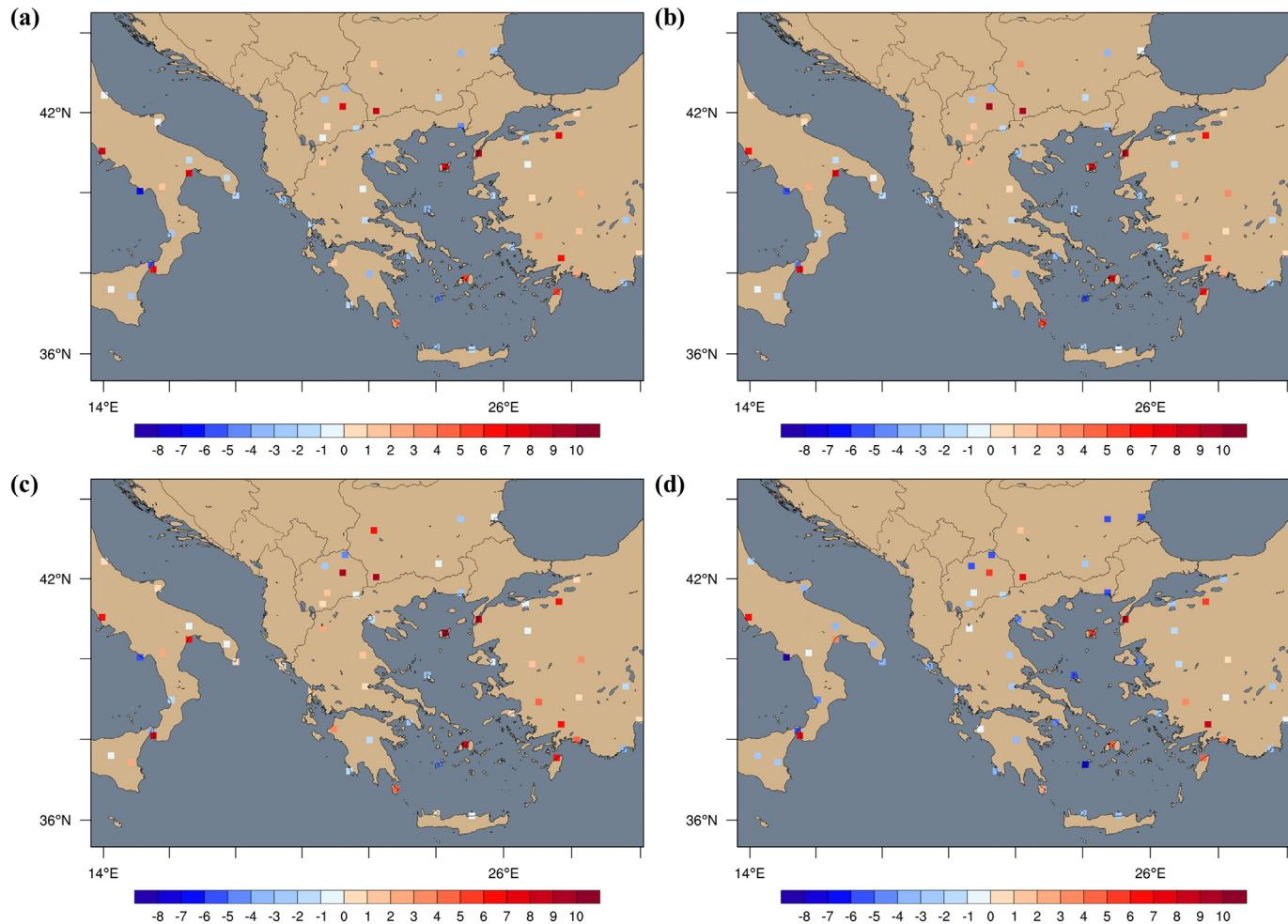
Fig. 6. Spatial distribution of MBs for WS10 (m/s) for (a) Noah, (b) Noah-MP, (c) CLM4, and (d) RUC numerical experiments.

application of the Noah-MP and CLM4 LSMs results in the highest PET underestimations in terms of overall MAE (3.70 °C and 3.98 °C, respectively; Table 5), which surpass the PET reference intervals in the mountain areas (> 4 °C).

Fig. 11 illustrates that the RUC scenario is capable of reproducing successfully the exacerbated, due to the HWs, observed magnitude of the human thermal discomfort. In particular, the RUC experiment slightly underestimates the extreme heat stress conditions (PET > 41 °C) by ∼ 0.9%, while it captures nearly exactly the ∼ 21% of the strong heat stress PET values (35 °C < PET < 41 °C). Moderate (29 °C < PET < 35 °C) and slight (23 °C < PET < 29 °C) heat stress frequencies are underestimated in the RUC experiment by ∼ 0.2% and ∼ 3.4%, respectively. The CLM4 scenario also slightly underestimates the class of PET > 41 °C (∼ 1.1%) and deviates from the observed strong and moderate heat stress conditions by less than ∼ 1%. However, it produces the greatest underestimation (overestimation) of the observed values of 23 °C < PET < 29 °C (8 °C < PET < 18 °C). This indicates that the CLM4 LSM performs quite well in replicating the high daily temperatures, but it fails to capture the HW-induced elevated temperatures in the nighttime. The Noah and Noah-MP LSMs yield somewhat similar results for the PET values between 18 °C and 41 °C (moderate cold stress to strong heat stress conditions), whereas the Noah experiment shows a noticeably better performance in capturing the observed strong cold (8 °C < PET < 18 °C) and extreme heat (PET > 41 °C) stress conditions.

Fig. 12a-c confirm that the CLM4 LSM considerably underestimates the nocturnal 2-m air temperatures, especially in the high-altitude regions. This could be associated with the use of the USGS LULC data in

the CLM4 experiment. The USGS dataset is based on satellite data derived from April 1992 to March 1993. Thus, they may not reflect the actual LULC distribution in the study area leading to possible errors in the simulation of the SEB that in turn result in the poor replication of the nighttime T2 values by the CLM4 scenario. Following the T2 underestimations during the night, the CLM4 experiment fails to capture correctly the observed PET values at 0000 UTC (Fig. 12d-f). The Noah-MP experiment also shows poor performance in simulating the magnitude of the observed nocturnal T2 and PET values. On the other hand, the RUC and Noah LSMs produce the most accurate nighttime T2 and PET values over all altitudinal classes (Fig. 12). The RUC scenario also yields the best T2 and PET results in overall at 1200 UTC, whereas the CLM4 LSM overestimates the daytime SATs during a noticeable number of the examined HW days, especially over the coastal regions. The Noah and Noah-MP LSMs show the largest T2 and PET underestimations during the day (Fig. 13). The lower R values occurred for T2 and PET in the nighttime (Fig. 12) compared to the daytime (Fig. 13) point out a moderate model performance in simulating the variations of both variables during the night. Also, a lower correspondence between the observed and modeled T2 and PET values is evident in the mountain sites (Fig. 12c,f and 13c, f) compared to the coastal and inland stations (Fig. 12a, b, d, e and 13 a, b, d, e). This indicates that the thermal environment conditions are harder to simulated over complex-terrain regions, probably because of the coarse horizontal grid resolution of the model.

(a)

(b)

(c)

(d)
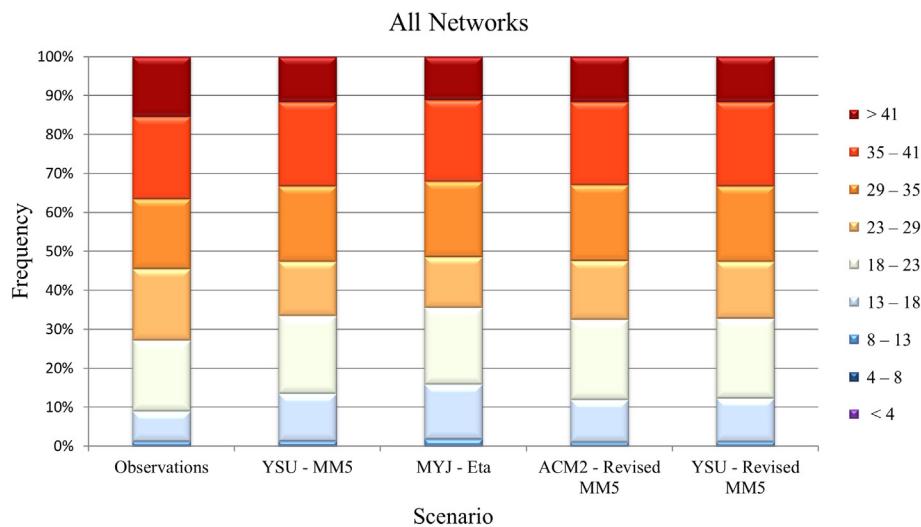
Fig. 7. Spatial distribution of VP2 MBs (hPa) for the (a) Noah, (b) Noah-MP, (c) CLM4, and (d) RUC numerical experiments.
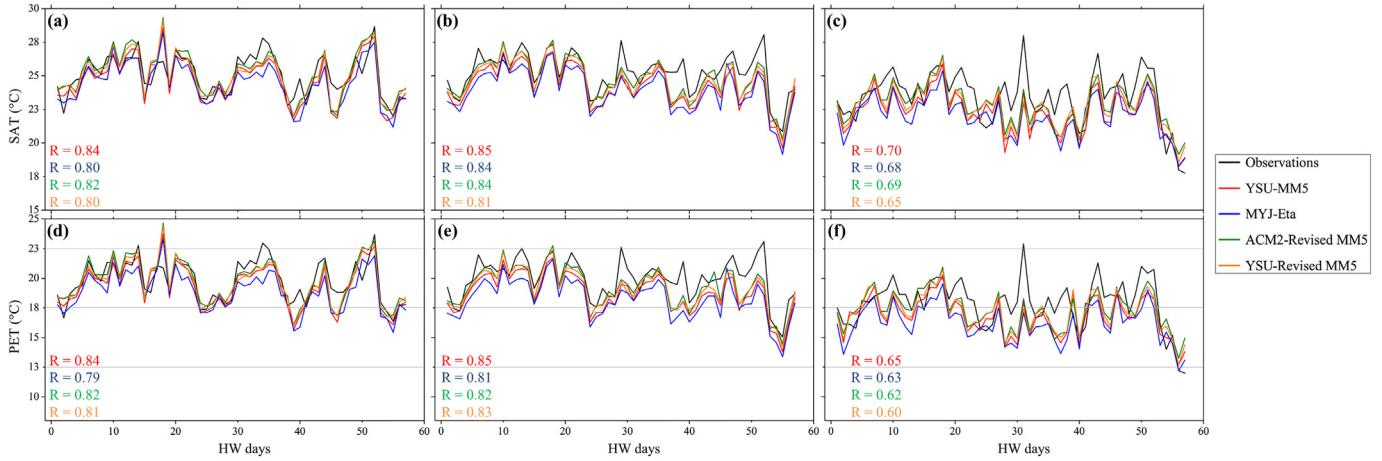
3.3. Interpretation of the modeled differences

The above-discussed results reveal that the examined WRF physics lead to significant differences in the simulation of the HW-associated meteorological and human thermal comfort conditions over the SE Mediterranean and BP. Understanding the sources of these differences

requires investigating the impact of each tested physics on the modeled surface energy fluxes and PBL height. The investigation is conducted for the 24–27 June 2007 HW episode because the synoptic conditions of this event are characteristic and favor the appearance of extreme hot spells over the study area with the highest probability (Katsoulis and Hatzianastassiou, 2005). Figs. S1-S2 (see supplementary material)

Fig. 8. PET thermal bioclimate diagram from observed and WRF-modeled data (PBL-SL experiments).

Fig. 9. Observed and modeled (PBL-SL experiments) daily values of T2 (a-c) and PET (d-f) averaged over the Coastal (a, d), Inland (b, e), and Mountain (c, f) stations at 0000 UTC. The light grey lines denote the classes of slight cold (8 °C < PET < 13 °C), moderate cold (13 °C < PET < 18 °C), no thermal (18 °C < PET < 23 °C), and slight heat (23 °C < PET < 29 °C) stress conditions.

illustrate this synoptic circulation from the CFS reanalysis and WRF-simulated data. The two patterns are in close agreement showing the Azores sub-tropical high to be positioned over the Mediterranean with an extended ridge towards the Balkans. At the same time, two cyclones are present; one over the North Atlantic Ocean (the Icelandic low) and one over eastern Russia. These features created a large-scale advection of hot air from the north Africa to the southeast Mediterranean on the 24th of June (see supplementary material Figs. S1a-b). The next day the low-pressure systems intensified, leading to the strengthening of the anticyclonic circulation over the Mediterranean (see supplementary material Figs. S1c-d), which in turn increased the adiabatic warming of the underlying atmospheric levels. The combined effect of both warming mechanisms persisted in the following days (26–27 June; see supplementary material Figs. S2a-d) leading to positive anomalies that surpassed 12 °C over Greece (Kotroni et al., 2011). As a result, the intensity of the examined HW event was quite strong with persistent high near-surface air temperatures that reached up to 40 °C over some regions of the SE Mediterranean and Balkan Peninsula.
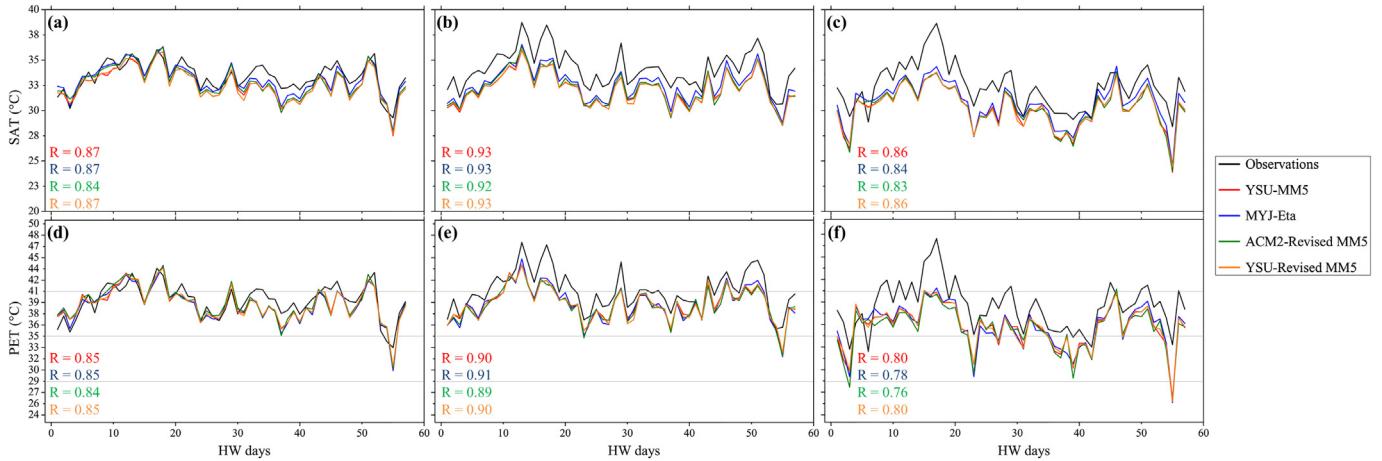
3.3.1. PBL schemes
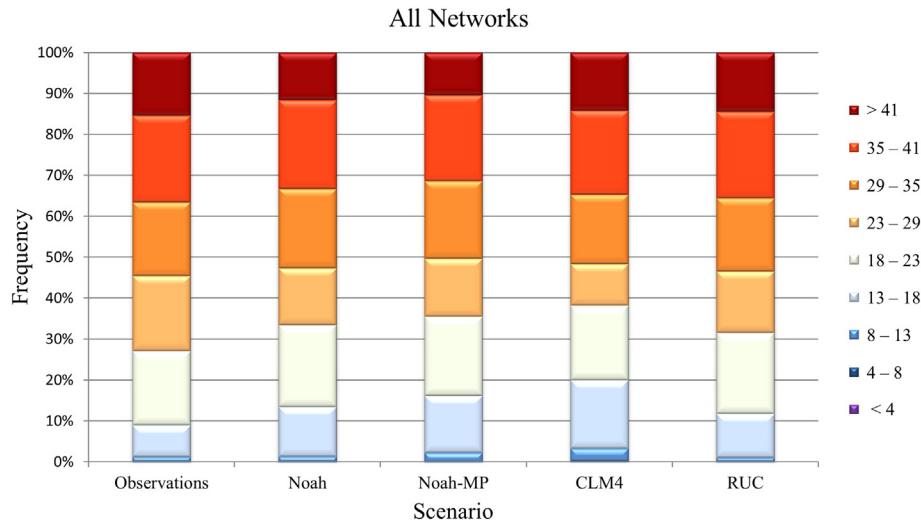
Fig. 14 illustrates the simulated diurnal variation of the sensible heat (SH) flux, latent heat (LH) flux, Bowen ratio (BR; i.e., the ratio of SH to LH flux) and PBL height (PBLH) from the PBL-SL experiments, averaged for 24–27 June 2009, over the coastal, inland and mountain

ground-based stations used in the present study.

As it was expected, the MYJ-Eta experiment simulates higher SH flux values during the day, especially over the coastal and inland regions (Fig. 14a, e) leading to greater heating of the overlaying atmosphere near the surface (i.e., higher modeled daytime T2 values; Fig. 10a-c). This is in agreement with the findings of other similar studies (Hari Prasad et al., 2016; Hariprasad et al., 2014; Sathyanadh et al., 2017; Xie et al., 2012) and it is attributed to the Eta SL scheme, which produces a higher surface heat exchange coefficient than that provided by the MM5 and Revised-MM5 SL parameterizations (Shin and Hong, 2011; Xie et al., 2012). The MYJ-Eta experiment also exhibits slightly greater negative values of SH over all geographical sub-regions in the night (Fig. 14a, e, i). The negative sign means that the SH fluxes are downward transporting heat from the atmosphere to the surface. Thus, it would be expected that the MYJ-Eta scenario would simulate also higher T2 values during the night. However, Fig. 9a-c in Sec. 3.2.1 illustrated that the use of the MYJ-Eta PBL-SL scheme results in the lowest nocturnal T2 values modeled by the WRF. This could be related to the slightly higher LH fluxes generated by the MYJ-Eta experiment in the night (Fig. 14b, f, k) that denote a marginally enhanced nighttime evaporative cooling. This is also supported, by the less negative MYJ-Eta-modeled BR values during most of the nocturnal hours (Fig. 14c, g, f), as small simulated BR values imply a shift to evapotranspiration in the modeled surface energy budget partitioning

Fig. 10. Observed and modeled (PBL-SL experiments) daily values of T2 (a-c) and PET (d-f) averaged over the Coastal (a, d), Inland (b, e), and Mountain (c, f) stations at 1200 UTC. The light grey lines denote the classes of slight heat (23 °C < PET < 29 °C), moderate heat (29 °C < PET < 35 °C), strong heat (35 °C < PET < 41 °C), and extreme heat (PET > 41 °C) stress conditions.

Fig. 11. PET thermal bioclimate diagram from observed and WRF-modeled data (LSM experiments).

(Stéfanon et al., 2014). Beyond the SEB-associated processes, other complex physical mechanisms that affect the simulated T2 changes (e.g., advection) could contribute to the poor MYJ-Eta performance in replicating the observed T2 values in the nighttime (Fig. 9a-c).

For the rest of the sensitivity scenarios, small differences occur for the simulated SH and LH fluxes (Fig. 14a, b, e, f, i, k). In overall, the SH fluxes decrease as the altitude increases from the coastal (Fig. 14a) to inland (Fig. 14e) and mountain areas (Fig. 14i). This is because the LULC in the high-altitude regions is dominated by natural vegetation and forests, where the partitioning of the SEB is directed to the latent heat fluxes (Fig. 14f, k). The larger LH fluxes over the mountain locations suggest the strengthening of evapotranspiration, which in turn enhances the evaporative cooling and results in the lower modeled T2 and PET values in this areas (Fig. 9c) compared to those simulated over the other geographical sub-regions. Further, it contributes to the moist biases reported in the mountain regions for all PBL-SL schemes (Table 2).
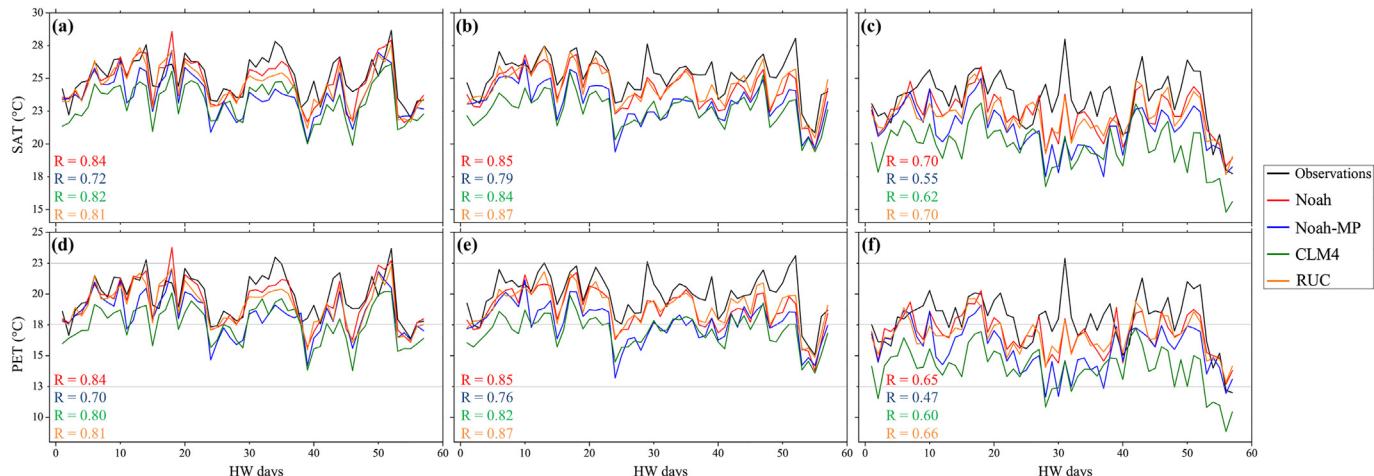
Concerning the PBLH, this variable increases with the terrain elevation because it is proportional to altitude (Fig. 14d, h, m). It would be expected that the higher SH fluxes in the MYJ-Eta experiment will result in higher PBL heights during the day. However, the ACM2-Revised MM5 scenario yields the largest values of PBLHs over all geographical sub-regions (Fig. 14d, h, m). Especially over the coastal areas, the

simulated PBLHs by the ACM2-Revised MM5 scenario are almost double compared to those modeled by the rest of the experiments. This highlights the stronger vertical mixing that speculated to be produced by the ACM2 PBL algorithm (Sec. 3.1.1). The enhanced ACM2 vertical mixing contributes to the efficient transport of warm and moist air from the surface, while it strengthens the dry and hot air entrainment from above the PBL. Hence, a warmer and drier near-surface atmospheric conditions are modeled by the local ACM2-Revised MM5 experiment compared to those simulated by the rest of the PBL-SL scenarios.
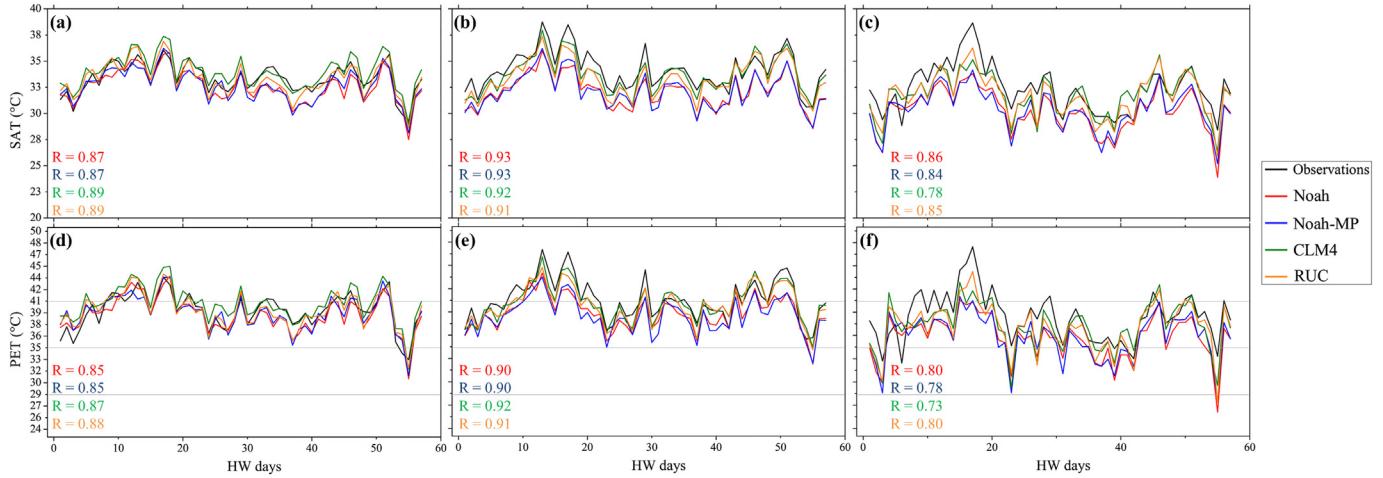
3.3.2. Land surface models

Fig. 15 presents the mean diurnal variation of the same quantities as in Fig. 14, modeled by the LSM numerical experiments over the examined geographical regions.

The RUC LSM produces significantly larger values of SH over all altitudinal classes compared to the other examined land surface parameterizations, which can be higher than ∼ 200 W/m² during the day (Fig. 15a, e, i). This contributes directly to the simulation of higher daytime T2 values by the RUC scenario (Fig. 13a-c) that assist the minimization of the cold T2 model errors during this experiment (Table 3). Concerning the latent heat, the RUC LSM simulates very low values during the day over the coastal areas (Fig. 15b, f, k). The evapotranspiration, expresses as LH, is strongly controlled by the ground

Fig. 12. Observed and modeled (LSM experiments) daily values of T2 (a-c) and PET (d-f) averaged over the Coastal (a, d), Inland (b, e), and Mountain (c, f) stations at 0000 UTC. The light grey lines denote the classes of slight cold (8 °C < PET < 13 °C), moderate cold (13 °C < PET < 18 °C), no thermal (18 °C < PET < 23 °C), and slight heat (23 °C < PET < 29 °C) stress conditions.

Fig. 13. Observed and modeled (LSM experiments) daily values of T2 (a-c) and PET (d-f) averaged over the Coastal (a, d), Inland (b, e), and Mountain (c, f) stations at 1200 UTC. The light grey lines denote the classes of slight heat (23 °C < PET < 29 °C), moderate heat (29 °C < PET < 35 °C), strong heat (35 °C < PET < 41 °C), and extreme heat (PET > 41 °C) stress conditions.

moisture content, while the Bowen ratio can be an indicator of the soil moisture availability (Stéfanon et al., 2014). Fig. 15c illustrates that the BR exceeds the value of five over the coastal locations during the midday and afternoon hours in the RUC scenario. This indicates very dry conditions associated with soil moisture deficit (Stéfanon et al., 2014), as speculated in the Sec. 3.1.2. Thus, this finding explains the high coastal dry bias of the RUC LSM in Table 3. It is also associated with the previous-mentioned enhanced sensible heating in the RUC experiment and disputes the reliability of the RUC-simulated T2 because if the cold T2 model biases are canceled out due to the soil moisture underestimation, then the RUC scenario provides the most accurate T2 results for the wrong reason.
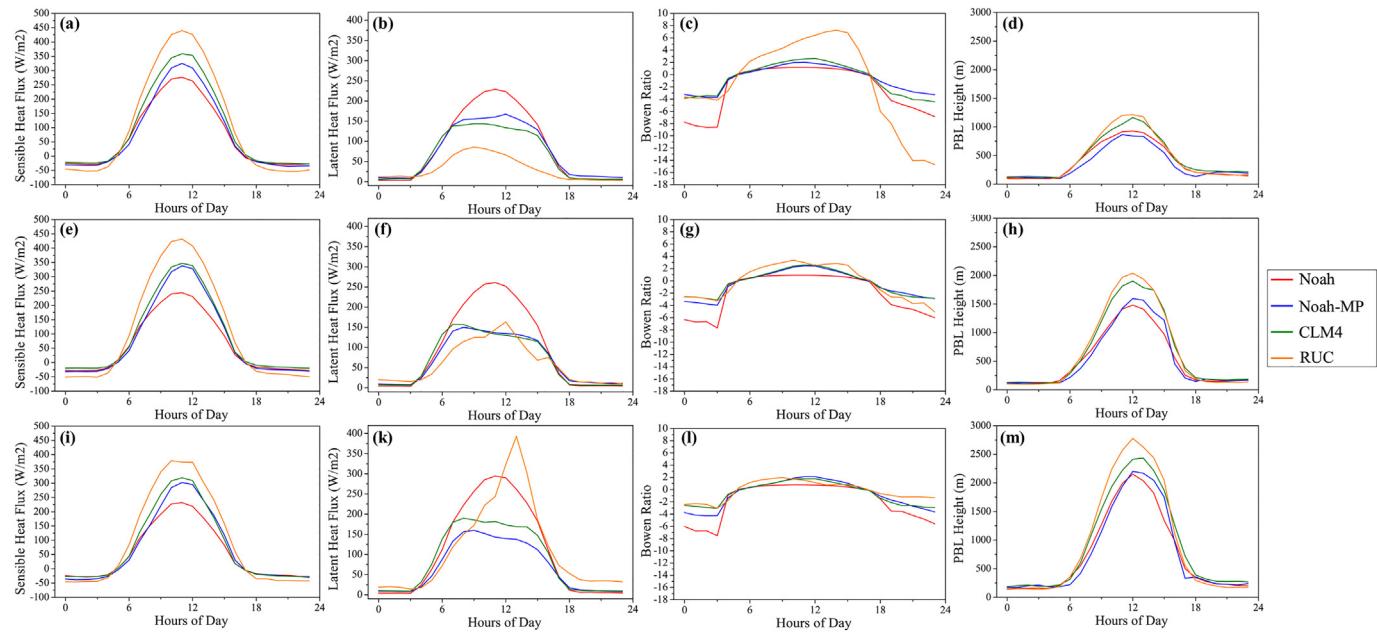
Regarding inland and mountain areas, the RUC BR (LH fluxes) takes lower (higher) values in the day compare to the coastal sites, which are close to those of CLM4 and Noah-MP experiments 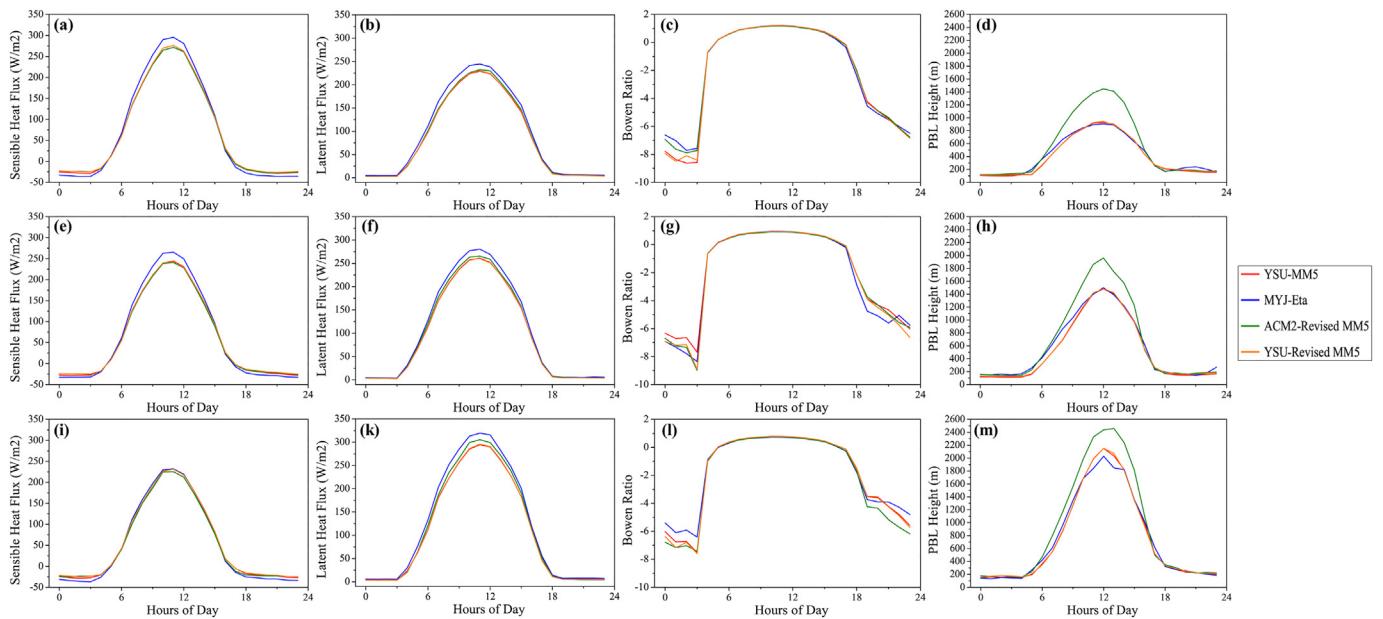(Fig. 15g, l). This is because these areas introduce natural vegetation LULC, as already mentioned in the previous section (3.3.1). Especially for the high-altitude regions, a significant peak in LH fluxes is present at 1300 UTC

reaching close to 400 W/m². This is attributed to the precipitation simulated by the RUC experiment mainly over Bulgaria and Turkey (see supplementary material Fig. S3). The accumulated precipitation between 0000 UTC and 1300 UTC on the last day of the examined HW exceeds the 6 mm over the northeast (west) part of Turkey (Bulgaria). The occurrence of this phenomenon is associated with the synergy between the land surface and PBL processes, highlighting the importance of the coupling of these procedures in the WRF model. In particular, the greater SH fluxes lead to highly convective conditions that support the growth of greater PBLHs in the day (Avolio et al., 2017; Madala et al., 2015), as illustrated in Fig. 15d, h, m for the RUC scenario. The convective atmosphere combined with possible upslope (i.e., anabatic) winds in the complex terrain of the examined areas could trigger the simulation of clouds and precipitation by the RUC LSM, as also highlighted by Stéfanon et al., 2014.

The CLM4 model generates the lowest values of SH in the night over all geographical areas (Fig. 15a, e, i). This contributes to the poor performance of the CLM4 scenario in replicating the observed 2-m air

Fig. 14. Simulated mean diurnal variation of the sensible heat flux, (a, e, i), latent heat flux (b, f, k), Bowen ratio (c, g, l), and PBL height (d, h, m) by the PBL-SL experiments over the Coastal (a-d), Inland (e-h), and Mountain (i-m) regions. Values are averaged for 24–27 June 2007.

Fig. 15. Simulated mean diurnal variation of the sensible heat flux, (a, e, i), latent heat flux (b, f, k), Bowen ratio (c, g, l), and PBL height (d, h, m) by the LSM experiments over the Coastal (a-d), Inland (e-h), and Mountain (i-m) regions. Values are averaged for 24–27 June 2007.

temperatures during the night (Fig. 12a-c) and is may associated with the use of the outdated USGS LULC data in this experiment (see Sec. 3.2.2). On the other hand, the greatest nighttime downward SH fluxes provided by the RUC LSM over all examined areas (Fig. 15a, e, i) contribute to the simulation of more consistent with the observations nighttime T2 values (Fig. 12a-c). Also, another contributing factor in the coastal and inland regions could be the depletion of the nocturnal evaporative cooling in the RUC LSM. This is demonstrated by the very high values of the RUC-modeled BR in these areas that reach up to absolute fifteen over the coastal sites (Fig. 15c). However, such extreme BR values may be unrealistic supporting the possibility that the RUC LSM underestimates significantly the soil moisture. The Noah-simulated BR values also point out a limited evaporative cooling during the night, in contrast with the Noah-MP experiment that yields much less negative BR values denoting more evaporative cooling. This finding could partially explain the higher (lower) nocturnal T2 values simulated by the Noah (Noah-MP) LSM. Both Noah and Noah-MP scenarios exhibit lower PBL heights during the day than that simulated by the RUC and CLM4 LSMs. This indicates a lower turbulent mixing in the lower atmosphere that inhibits the convective heat transfer contributing to the poor performance of the Noah and Noah-MP models in capturing the elevated daytime T2 values (Fig. 13a,c). Another contributing factor for the Noah experiment could be the high (low) LH fluxes (BR) that indicate an enhanced daytime evaporative cooling.

4. Conclusions

The mesoscale NWP model WRF was utilized in the present study to investigate the impact of the (a) surface-layer, (b) land surface and (c) planetary boundary layer parameterizations on the short-term simulations of heat wave episodes over the SE Mediterranean and Balkan Peninsula. Four widely used PBL-SL schemes (i.e., YSU-MM5, MYJ-Eta, ACM2-Revised MM5, YSU-Revised MM5) and LSMs (Noah, Noah-MP, CLM4, RUC) were tested during 15 HW events (57 HW days) that identified between 2004 and 2013. The model performance was evaluated focusing on near-surface atmospheric variables and particulary on the 2-m air temperature. Additionally, a thermal bioclimate index (i.e., PET) was used through the application of the RayMan model to investigate the capability of the WRF model under the examined

parameterizations to reproduce the observed human thermal comfort conditions.

The evaluation of the modeled results against observations showed a cold model bias. The WRF model performed slightly better in reproducing the observed T2 and PET values in the ACM2-Revised MM5 scenario compared to the other experiments. The comparison beteween the modeled and observed nighttime and daytime temperatures revealed that the same experiment produced the most accurate T2 results during the nigth. However, the MYJ-Eta scenario yielded the lowest deviations from the observed daytime temperatures. The RUC LSM minimized remarkably the cold model bias and showed the best performance in replicating the elevated temperatures during the day. This contributed to the sufficient replicaton of the above 35% observed strong and extreme heat stress conditions by the RUC scenario (∼ 0.9% deviation). The best results for the nighttime T2 were also provided by the RUC model with the Noah LSM close second. The observed WS10 was primarily overestimated by the WRF model over the island and mountain sites of the study area, whereas it was mostly underpredicted over Italy and Turkey. In both cases, the application of the ACM2-Revised MM5 PBL-SL scheme and Noah LSM led to the lowest WS10 biases. The WRF model lacked ability in replicating the magnitude and phase of the VP2. In overall, the model simulated drier conditions over the coastal and inland areas than the observed, whereas it showed moist biases in the mountain regions. The moister biases were produced by the local PBL scheme (MYJ) indicating a weaker vertical mixing compared to the rest of the PBL parameterizations. On the other hand, the RUC experiment exhibited the most significant dry biases that may associated with a possible underestimation of the soil moisture in the RUC LSM.

The above results revealed the substantial simulated differences introduced by the examined WRF physics. To understand thoroughly the sources of these differences, the mean diurnal variation of the modeled surface fluxes and PBL height was investigated during the severe HW event of June 2007. The conducted analysis provided valuable insight on the systematic effects of the PBL-SL schemes and LSMs on the WRF-simulated HWs. More precisely, the application of the MYJ-Eta PBL-SL parameterization leads to the simulation of greater sensible heating compared to the other PBL-SL schemes due to the high surface heat exchange coefficient provided by the Eta SL scheme (Xie et al.,

2012; Shin and Hong, 2011). This contributes to the high MYJ-Eta-modeled daytime T2 values that are close to the observations. The ACM2 PBL algorithm produces significantly higher PBL heights compared to rest of the tested PBL parameterizations reflecting the strong vertical mixing generated in this scheme (García-Díez et al., 2013; Hu et al., 2010). This contributes to the simulation of warm and dry near-surface atmospheric conditions. Thus, the ACM2-Revised MM5 PBL-SL scheme yields the most accurate T2 results over the entire study region, whereas it exhibits low (significant) VP2 overestimations (under-estimations) over the mountain (coastal and inland) areas. The CLM4 LSM results in the simulation of the lowest nighttime downward SH fluxes. This contributes to the poor model performance in capturing the observed high nocturnal T2 values and is may associated with the employment of the outdated USGS LULC dataset in the CLM4 experiment. The BR values produced by the Noah (Noah-MP) LSM denote that the modeled evaporative cooling is limited (enhanced) during the night. This contributes to the adequate (poor) representation of the observed nighttime T2 values by the Noah (Noah-MP) scenario. Contrary, the daytime evaporative cooling is strengthened in the Noah LSM contributing to the simulation of low T2 values during the day, which deviate significantly from the observations. The use of the RUC LSM strengthens significantly the modeled sensible heat fluxes and Bowen ratio contributing directly to the satisfactory replication of the observed T2 and PET values. However, the enhanced sensible heating relative to the latent heating is may related to the possible underprediction of soil moisture, as indicated by the significant RUC-induced dry biases, especially over the coastal sites. This would be in agreement with the findings of previous studies over China (Ma et al., 2017; Zeng et al., 2015), while it is worth mentioning that Stegehuis et al. (2015) found significant biases in the RUC-simulated sensible and latent heat fluxes over western Europe during the summer of 2003. Thus, the most accurate T2 results provided by the RUC model in the present study may arise from the wrong reason. A follow-up study could focus on the validation of the WRF-simulated SEB components based on flux tower filed measurements (e.g., FLUXNET; Baldocchi et al., 2001) and satellite-driven methodologies (e.g., Global Land Evaporation Amsterdam Model – GLEAM; Martens et al., 2017; Miralles et al., 2011) to investigate the speculation mentioned above.

To sum up, the physical factors that primarily affect the mesoscale short-term simulation of HWs with the use of the WRF model are the (a) surface heat exchange coefficients provided by the SL schemes, (b) surface energy budget partiotioning modeled by the LSMs and (c) vertical mixing parameterized by the PBL algorithms. The identification of these factors in the current study contributes to understanding better the performance of the tested WRF physics during high-temperature weather events. It is also very meaningful in the efforts to adapt properly and improve the model configuration for forecasting HWs and their impact on the human thermal comfort conditions accurately. This is of great importance for the examined area of the southeast Mediterranean and Balkan Peninsula, which is one of the most vulnerable regions regarding heat-related climate change impacts.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.atmosres.2018.11.015.

References

Akta, D., 2011. Heatwave occurances over Ankara and its simulations by WRF-ARW. Geophys. Res. Abstr. 13, 9393.

Akylas, E., Kotroni, V., Lagouvardos, K., 2007. Sensitivity of high-resolution operational weather forecasts to the choice of the planetary boundary layer scheme. Atmos. Res. 84, 49–57. https://doi.org/10.1016/j.atmosres.2006.06.001.

Amengual, A., Homar, V., Romero, R., Brooks, H.E., Ramis, C., Gordaliza, M., Alonso, S., 2014. Projections of heat waves with high impact on human health in Europe. Glob. Planet. Change 119, 71–84. https://doi.org/10.1016/j.gloplacha.2014.05.006.

Attada, R., Kumar, P., Dasari, H.P., 2018. Assessment of Land Surface Models in a High-Resolution Atmospheric Model during Indian Summer Monsoon. Pure Appl. Geophys. 175, 3671–3696. https://doi.org/10.1007/s00024-018-1868-z.

Avolio, E., Federico, S., Miglietta, M.M., Feudo, T., Calidonna, C.R., Sempreviva, A.M., 2017. Sensitivity analysis of WRF model PBL schemes in simulating boundary-layer variables in southern Italy. Exp. Campaign 192, 58–71. https://doi.org/10.1016/j.atmosres.2017.04.003.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw, U., Pilegaard, K.T., Schmid, K., Valentini, H.P., Verma, R., Vesala, S., Wilson, T., Wofsy, K., 2001. FLUXNET: a New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. Bull. Am. Meteorol. Soc. 82, 2415–2434.

Banks, R.F., Tiana-Alsina, J., Baldasano, J.M., Rocadenbosch, F., Papayannis, A., Solomos, S., Tzanis, C.G., 2016. Sensitivity of boundary-layer variables to PBL schemes in the WRF model based on surface meteorological observations, lidar, and radiosondes during the HygrA-CD campaign. Atmos. Res. 176–177, 185–201. https://doi.org/10.1016/j.atmosres.2016.02.024.

Beljaars, A.C.M., 1995. The parametrization of surface fluxes in large scale models under free convection. Q.J.R Meteorol. Soc. 121, 255–270. https://doi.org/10.1002/qj.49712152203.

Beniston, M., 2004. The 2003 heat wave in Europe: a shape of things to come? An analysis based on Swiss climatological data and model simulations. Geophys. Res. Lett. 31. https://doi.org/10.1029/2003GL018857.

Benjamin, S., Bleck, R., Brown, J., Brundage, K., Devenyi, D., Grell, G., Kim, D., Manikin, G., Schlatter, T., Schwartz, B., Smirnova, T., Weygandt, S., Alamos, L., 2004. Mesoscale Weather Prediction with the RUC Hybrid Isentropic-Sigma Coordinate Model and Data Assimilation System Operational Numerical Weather Prediction. In: Symp. 50th Anniv. Oper. Numer. Weather Predict, pp. 495–518.

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. Mon. Weather Rev 569–585.

Chiriaco, M., Bastin, S., Yiou, P., Haeffelin, M., Dupont, J.C., Stéfanon, M., 2014. European heatwave in July 2006: Observations and modeling showing how local processes amplify conducive large-scale conditions. Geophys. Res. Lett. 41, 5644–5652. https://doi.org/10.1002/2014GL060205.

Cohen, A.E., Cavallo, S.M., Coniglio, M.C., Brooks, H.E., 2015. A Review of Planetary Boundary Layer Parameterization Schemes and their Sensitivity in Simulating Southeastern U.S. Cold season Severe Weather Environments. Weather Forecast. 30, 591–612. https://doi.org/10.1175/WAF-D-14-00105.1.

Daneshvar, M., Bagherzadeh, A., Tavousi, T., 2013. Assessment of bioclimatic comfort conditions based on Physiologically Equivalent Temperature (PET) using the RayMan Model in Iran. Open Geosci. 5, 53–60. https://doi.org/10.2478/s13533-012-0118-7.

Dasari, H.P., Salgado, R., Perdigao, J., Challa, V.S., 2014. A Regional climate simulation Study using WRF-ARW Model over Europe and Evaluation for Extreme Temperature Weather events. Int. J. Atmos. Sci. 2014, 1–22. https://doi.org/10.1155/2014/704079.

Dimitrova, R., Silver, Z., Fernando, H.J.S., Leo, L., Di Sabatino, S., Hocut, C., Zsedrovits, T., 2014. Inter-comparison between different PBL options in WRF model: Modification of two PBL schemes for stable conditions. In: 94th AMS Annual Meeting.

Dudhia, J., 1989. Numerical Study of Convection Observed during the Winter Monsoon Experiment using a Mesoscale Two-Dimensional Model. J. Atmos. Sci. 46 (20), 3077–3107.

Dyer, A.J., Hicks, B.B., 1970. Flux-gradient relationships in the constant flux layer. Q. J. R. Meteorol. Soc. 96, 715–721. https://doi.org/10.1002/qj.49709641012.

Founda, D., Giannakopoulos, C., 2009. The exceptionally hot summer of 2007 in Athens, Greece - a typical summer in the future climate? Glob. Planet. Change 67, 227–236. https://doi.org/10.1016/j.gloplacha.2009.03.013.

García-Díez, M., Fernández, J., Fita, L., Yagüe, C., 2013. Seasonal dependence of WRF model biases and sensitivity to PBL schemes over Europe. Q. J. R. Meteorol. Soc. 139, 501–514. https://doi.org/10.1002/qj.1976.

Giannaros, T.M., Melas, D., Matzarakis, A., 2015. Evaluation of thermal bioclimate based on observational data and numerical simulations: an application to Greece. Int. J. Biometeorol. 59, 151–164. https://doi.org/10.1007/s00484-014-0832-6.

Giannaros, C., Nenes, A., Giannaros, T.M., Kourtidis, K., Melas, D., 2018a. A comprehensive approach for the simulation of the Urban Heat Island effect with the WRF/SLUCM modeling system: the case of Athens (Greece). Atmos. Res. 201, 86–101.

Giannaros, T.M., Kotroni, V., Lagouvardos, K., Matzarakis, A., 2018b. Climatology and trends of the Euro-Mediterranean thermal bioclimate. Int. J. Climatol. 1–19. https://doi.org/10.1002/joc.5501.

Gilliam, R.C., Hogrefe, C., Rao, S.T., 2006. New methods for evaluating meteorological models used in air quality applications. Atmos. Environ. 40, 5073–5086. https://doi.org/10.1016/j.atmosenv.2006.01.023.

Hari Prasad, K.B.R.R., Venkata Srinivas, C., Venkateswara Naidu, C., Baskaran, R., Venkatraman, B., 2016. Assessment of surface layer parameterizations in ARW using micro-meteorological observations from a tropical station. Meteorol. Appl. 23, 191–208. https://doi.org/10.1002/met.1545.

Hariprasad, K.B.R.R., Srinivas, C.V., Singh, A.B., Vijaya Bhaskara Rao, S., Baskaran, R., Venkatraman, B., 2014. Numerical simulation and intercomparison of boundary layer structure with different PBL schemes in WRF using experimental observations at a tropical site. Atmos. Res. 145–146, 27–44. https://doi.org/10.1016/j.atmosres.2014.03.023.

Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., Alexander, L.V., Brönnimann, S., Charabi, Y., Dentener, F.J., Dlugokencky, E.J., Easterling, D.R., Kaplan, A., Soden, B.J., Thorne, P.W., Wild, M., Zhai, P.M., 2013. Observations: Atmosphere and

Surface. In: climate Change 2013: the Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on climate Change. Cambridge University Press, Cambridge and New York.

Hong, S., Lim, J., 2006. The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). J. Korean Meteor. Soc.

Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. Mon. Weather Rev. 134, 2318–2341. https://doi.org/10.1175/MWR3199.1.

Höppe, P., 1984. Die Energiebilanz des Menschen (dissertation). Wiss. Mitt. Meteorol. Isnt. Univ. München 49.

Höppe, P., 1999. The physiological equivalent temperature - a universal index for the biometeorological assessment of the thermal environment. Int. J. Biometeorol. 43, 71–75. https://doi.org/10.1007/s004840050118.

Hu, X.M., Nielsen-Gammon, J.W., Zhang, F., 2010. Evaluation of three planetary boundary layer schemes in the WRF model. J. Appl. Meteorol. Climatol. 49, 1831–1844. https://doi.org/10.1175/2010JAMC2432.1.

Janjić, Z.I., 1994. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. Mon. Weather Rev. 122 (1994), 927–945.

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. Mon. Weather Rev. 140, 898–918. https://doi.org/10.1175/MWR-D-11-00056.1.

Kain, J.S., 2004. The Kain–Fritsch Convective Parameterization: an Update. J. Appl. Meteorol. 43, 170–181.

Katsoulis, B.D., Hatzianastassiou, N., 2005. Analysis of hot spell characteristics in the Greek region. Clim. Res. 28, 229–241. https://doi.org/10.3354/cr028229.

Kim, Y., Sartelet, K., Raut, J.C., Chazette, P., 2013. Evaluation of the Weather Research and Forecast/Urban Model over Greater Paris. Boundary-Layer Meteorol. 149, 105–132. https://doi.org/10.1007/s10546-013-9838-6.

Kotroni, V., Lagouvardos, K., Retalis, A., 2011. The heat wave of June 2007 in Athens, Greece - part 2: Modeling study and sensitivity experiments. Atmos. Res. 100, 1–11. https://doi.org/10.1016/j.atmosres.2010.12.007.

Lawrence, D.M., Oleson, K.W., Flanner, M.G., Thornton, P.E., Swenson, S.C., Lawrence, P.J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G.B., Slater, A.G., 2011. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. J. Adv. Model. Earth Syst 3. https://doi.org/10.1029/2011MS00045.

Ma, Y., Zeng, X.-M., Zhang, Y., Wang, N., Zheng, Y., Wang, G., Chen, C., 2017. Impact of the choice of Land Surface Scheme on a simulated Heatwave Event: 2 the Case of Sichuan-Chongqing Area. China. Adv. Meteorol. 2017. https://doi.org/10.1155/2017/9545896.

Madala, S., Satyanarayana, A.N.V., Srinivas, C.V., Kumar, M., 2015. Mesoscale atmospheric flow-field simulations for air quality modeling over complex terrain region of Ranchi in eastern India using WRF. Atmos. Environ. 107, 315–328. https://doi.org/10.1016/j.atmosenv.2015.02.059.

Martens, B., Miralles, D.G., Lievens, H., Van Der Schalie, R., De Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017.

Matzarakis, A., Endler, C., 2010. Climate change and thermal bioclimate in cities: Impacts and options for adaptation in Freiburg. Germany. Int. J. Biometeorol. 54, 479–483. https://doi.org/10.1007/s00484-009-0296-2.

Matzarakis, A., Mayer, H., 1996. Another kind of environmental stress: thermal stress. WHO Newsletter 18, 7–10.

Matzarakis, A., Mayer, H., 1997. Heat stress in Greece. Int. J. Biometeorol. 41, 34–39. https://doi.org/10.1007/s004840050051.

Matzarakis, A., Nastos, P.T., 2011. Human-biometeorological assessment of heat waves in Athens. Theor. Appl. Climatol. 105, 99–106. https://doi.org/10.1007/s00704-010-0379-3.

Matzarakis, A., Rutz, F., Mayer, H., 2007. Modelling radiation fluxes in simple and complex environments - application of the RayMan model. Int. J. Biometeorol. 51, 323–334. https://doi.org/10.1007/s00484-006-0061-8.

Matzarakis, A., Rutz, F., Mayer, H., 2010. Modelling radiation fluxes in simple and complex environments: Basics of the RayMan model. Int. J. Biometeorol. 54, 131–139. https://doi.org/10.1007/s00484-009-0261-0.

Mayer, H., Höppe, P., 1987. Thermal comfort of man in different urban environments. Theor. Appl. Climatol. 38, 43–49. https://doi.org/10.1007/BF00866252.

Meehl, G.A., Tebaldi, C., 2004. More intense, more frequent, and longer Lasting Heat Waves in the 21st Century. Science 305, 994–997. https://doi.org/10.1126/science.1098704.

Miralles, D.G., Holmes, T.R.H., De Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15, 453–469. https://doi.org/10.5194/hess-15-453-2011.

Miralles, D.G., Teuling, A.J., van Heerwaarden, C.C., Vilà-Guerau De Arellano, J., 2014. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. Nat. Geosci. 7, 345–349. https://doi.org/10.1038/ngeo2141.

Monin, A.S., Obukhov, A.M., 1954. Basic laws of turbulent mixing in the surface layer of the atmosphere. Contrib. Geophys. Inst. Acad. Sci. USSR 24, 163–187.

Muthers, S., Matzarakis, A., Koch, E., 2010. Summer climate and mortality in Vienna - a human-biometeorological approach of heat-related mortality during the heat waves in 2003. Wien. Klin. Wochenschr. 122, 525–531. https://doi.org/10.1007/s00508-010-1424-z.

Nastos, P.T., Kapsomenakis, J., 2015. Regional climate model simulations of extreme air temperature in Greece. Abnormal or common records in the future climate? Atmos. Res. 152, 43–60. https://doi.org/10.1016/j.atmosres.2014.02.005.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1 Model description and evaluation with local-scale measurements. J. Geophys. Res. Atmos. 116, 1–19. https://doi.org/10.1029/2010JD015139.

Paulson, C.A., 1970. The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. J. Appl. Meteorol. 9, 857–861.

Pleim, J.E., 2007a. A combined local and nonlocal closure model for the atmospheric boundary layer Part I: Model description and testing. J. Appl. Meteorol. Climatol. 46, 1383–1395. https://doi.org/10.1175/JAM2539.1.

Pleim, J.E., 2007b. A combined local and nonlocal closure model for the atmospheric boundary layer, Part II: Application and evaluation in a mesoscale meteorological model. J. Appl. Meteorol. Climatol. 46, 1396–1409. https://doi.org/10.1175/JAM2534.1.

Ramamurthy, P., Li, D., Bou-Zeid, E., 2017. High-resolution simulation of heatwave events in New York City. Theor. Appl. Climatol. 128, 89–102. https://doi.org/10.1007/s00704-015-1703-8.

Santos-Alamillos, F.J., Zquez, D.P.V., Ruiz-Arias, J.A., Lara-Fanego, V., Tovar-Pescador, J., 2013. Analysis of WRF model wind estimate sensitivity to physics parameterization choice and terrain representation in Andalusia (Southern Spain). J. Appl. Meteorol. Climatol. 52, 1592–1609. https://doi.org/10.1175/JAMC-D-12-0204.1.

Sathyanadh, A., Prabha, T.V., Balaji, B., Resmi, E.A., Karipot, A., 2017. Evaluation of WRF PBL parameterization schemes against direct observations during a dry event over the Ganges valley. Atmos. Res. 193, 125–141. https://doi.org/10.1016/j.atmosres.2017.02.016.

Schwarzkopf, M.D., Fels, S.B., 1991. The simplified exchange method revisited – an accurate, rapid method for the computation of infrared cooling rates and fluxes. J. Geophys. Res. 96 (D5), 9075–9096. https://doi.org/10.1029/89JD01598.

Shaffer, S.R., Chow, W.T.L., Georgescu, M., Hyde, P., Jenerette, G.D., Mahalov, A., Moustaoui, M., Ruddell, B.L., 2015. Multiscale modeling and evaluation of urban surface energy balance in the Phoenix metropolitan area. J. Appl. Meteorol. Climatol. 54, 322–338. https://doi.org/10.1175/JAMC-D-14-0051.1.

Shin, H.H., Hong, S.Y., 2011. Intercomparison of Planetary Boundary-Layer Parametrizations in the WRF Model for a Single Day from CASES-99. Boundary-Layer Meteorol. 139, 261–281. https://doi.org/10.1007/s10546-010-9583-z.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Backer, D.M., Duda, M.G., Huang, X.Y., Wang, W., Power, J.G., 2008. A Description of the Advanced WRF Version 3. NCAR Technical Note (NCAR/TN-475 + STR), Boulder, Colorado, USA.

Stéfanon, M., Drobinski, P., D'Andrea, F., Lebeaupin-Brossier, C., Bastin, S., 2014. Soil moisture-temperature feedbacks at meso-scale during summer heat waves over Western Europe. Clim. Dyn. 42, 1309–1324. https://doi.org/10.1007/s00382-013-1794-9.

Stegehuis, A.I., Vautard, R., Ciais, P., Teuling, A.J., Miralles, D.G., Wild, M., 2015. An observation-constrained multi-physics WRF ensemble for simulating European mega heat waves. Geosci. Model Dev. 8, 2285–2298. https://doi.org/10.5194/gmd-8-2285-2015.

Tewari, M., Chen, F., Wang, W., Dudhia, J., Lemone, M.A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., Cuenca, R.H., 2004. Implementation and verification of the unified NOAH land surface model in the WRF model. In: 20th conference on weather analysis and forecasting/16th conference on numerical weather prediction, pp. 11–15.

Theoharatos, G., Pantavou, K., Mavrakis, A., Spanou, A., Katavoutas, G., Efstathiou, P., Mpekas, P., Asimakopoulos, D., 2010. Heat waves observed in 2007 in Athens, Greece: Synoptic conditions, bioclimatological assessment, air quality levels and health effects. Environ. Res. 110, 152–161. https://doi.org/10.1016/j.envres.2009.12.002.

Trenberth, K.E., 1992. Global analyses from ECMWF and Atlas of 1000 to 10 mb Circulation Statistics. TN-373 + STR, National Center for Atmospheric Research, pp. 205.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernández, J., García-Díez, M., Goergen, K., Güttler, I., Halenka, T., Karacostas, T., Katragkou, E., Keuler, K., Kotlarski, S., Mayer, S., van Meijgaard, E., Nikulin, G., Patarčić, M., Scinocca, J., Sobolowski, S., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V., Yiou, P., 2013. The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Clim. Dyn. 41, 2555–2575. https://doi.org/10.1007/s00382-013-1714-z.

Webb, E.K., 1970. Profile relationships: the log-linear range, and extension to strong stability. Q. J. R. Meteorol. Soc. 96, 67–90. https://doi.org/10.1002/qj.49709640708.

Xie, B., Fung, J.C.H., Chan, A., Lau, A., 2012. Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model. J. Geophys. Res. Atmos. 117, 1–26. https://doi.org/10.1029/2011JD017080.

Yang, Z.L., Niu, G.Y., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 2 Evaluation over global river basins. J. Geophys. Res. Atmos. 116, 1–16. https://doi.org/10.1029/2010JD015140.

Zeng, X.M., Wu, Z.H., Xiong, S.Y., Song, S., Zheng, Y.Q., Liu, H.Q., 2011. Sensitivity of simulated short-range high-temperature weather to land surface schemes by WRF. Sci. China Earth Sci. 54, 581–590. https://doi.org/10.1007/s11430-011-4181-6.

Zeng, X.M., Wang, B., Zhang, Y., Song, S., Huang, X., Zheng, Y., Chen, C., Wang, G., 2014. Sensitivity of high-temperature weather to initial soil moisture: a case study using the WRF model. Atmos. Chem. Phys. 14, 9623–9639. https://doi.org/10.5194/acp-14-9623-2014.

Zeng, X.M., Wang, N., Wang, Y., Zheng, Y., Zhou, Z., Wang, G., Chen, C., Liu, H., 2015. WRF-simulated sensitivity to land surface schemes in short and medium ranges for a high-temperature event in East China: a comparative study. J. Adv. Model. Earth Syst. 7, 1305–1325. https://doi.org/10.1002/2015MS000440.

Zhang, D., Anthes, R.A., 1982. A High-Resolution Model of the Planetary Boundary Layer—Sensitivity Tests and Comparisons with SESAME-79 Data. J. Appl. Meteorol. 21, 1594–1609. https://doi.org/10.1175/1520-0450(1982) 021 < 1594:AHRMOT > 2.0.CO;2.

Zhang, H., Pu, Z., Zhang, X., 2013. Examination of Errors in Near-Surface Temperature and Wind from WRF Numerical Simulations in Regions of complex Terrain. Weather Forecast. 28, 893–914. https://doi.org/10.1175/WAF-D-12-00109.1.

Zilitinkevich, S.S., 1995. Non-local turbulent transport: pollution dispersion aspects of coherent structure of convective flows. In: Power, H., Moussiopoulos, N., Brebbia, C.A. (Eds.), Air Pollution III, Air Pollution Theory and simulation. Vol 1. Computational Mechanics Publications, Southampton, UK; Boston, MA, pp. 53–60.