Atmospheric Research 176–177 (2016) 185–201

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

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

Robert F. Banks [a,c,*], Jordi Tiana-Alsina [b], José María Baldasano [a,c], Francesc Rocadenbosch [b], Alexandros Papayannis [d], Stavros Solomos [e], Chris G. Tzanis [f]

[a] Earth Sciences Department, Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS), Barcelona, Spain
[b] Remote Sensing Laboratory, Department of Signal Theory and Communications (TSC), Universitat Politècnica de Catalunya, Barcelona, Spain
[c] Environmental Modeling Laboratory, Universitat Politècnica de Catalunya, Barcelona, Spain
[d] Laser Remote Sensing Unit, Physics Department, National Technical University of Athens, Zografou, Greece
[e] Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, Athens, Greece
[f] Climate Research Group, Division of Environmental Physics and Meteorology, Faculty of Physics, University of Athens, Athens, Greece

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form 16 February 2016
Accepted 24 February 2016
Available online 2 March 2016

Keywords:
Backscatter lidar
PBL parameterization scheme
Complex urban terrain
Planetary boundary-layer height
Weather Research and Forecasting (WRF) model
Athens, Greece

ABSTRACT

Air quality forecast systems need reliable and accurate representations of the planetary boundary layer (PBL) to perform well. An important question is how accurately numerical weather prediction models can reproduce conditions in diverse synoptic flow types. Here, observations from the summer 2014 HygrA-CD (Hygroscopic Aerosols to Cloud Droplets) experimental campaign are used to validate simulations from the Weather Research and Forecasting (WRF) model over the complex, urban terrain of the Greater Athens Area. Three typical atmospheric flow types were identified during the 39-day campaign based on 2-day backward trajectories: Continental, Etesians, and Saharan. It is shown that the numerical model simulations differ dramatically depending on the PBL scheme, atmospheric dynamics, and meteorological parameter (e.g., 2-m air temperature). Eight PBL schemes from WRF version 3.4 are tested with daily simulations on an inner domain at 1-km grid spacing. Near-surface observations of 2-m air temperature and relative humidity and 10-m wind speed are collected from multiple meteorological stations. Estimates of the PBL height come from measurements using a multiwavelength Raman lidar, with an adaptive extended Kalman filter technique. Vertical profiles of atmospheric variables are obtained from radiosonde launches, along with PBL heights calculated using bulk Richardson number. Daytime maximum PBL heights ranged from 2.57 km during Etesian flows, to as low as 0.37 km during Saharan flows. The largest differences between model and observations are found with simulated PBL height during Saharan synoptic flows. During the daytime, campaign-averaged near-surface variables show WRF tended to have a cool, moist bias with higher simulated wind speeds than the observations, especially near the coast. It is determined that non-local PBL schemes give the most agreeable solutions when compared with observations.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

HygrA-CD (hygroscopic aerosols to cloud droplets) was an internationally led field campaign performed from 15 May to 19 June 2014 in the Greater Athens Area (GAA), Greece. The main goal of the HygrA-CD campaign was to bring together different instruments and expertise for the purpose of understanding more about the impact of aerosols and clouds on weather and climate on a local scale. It is a novel attempt to strengthen the links between the remote sensing and the in situ observation communities, while making use of established know-how on numerical weather prediction and atmospheric modeling. An overview of the campaign can be found in Papayannis et al. (2015).

During the time period of the HygrA-CD campaign, it is common to observe diverse types of synoptic flows. Saharan dust events are likely over the GAA based on synoptic winds from the south and southwest, advecting dust aerosols into the region. In addition, air masses carrying mixtures of urban/continental and marine aerosols are probable due to the influence of the Etesian winds to the wind circulation in the GAA (i.e., synoptic winds from the northeast).

Model-simulated meteorological processes in the planetary boundary layer (PBL) are critical to an air quality forecast system, as a numerical weather prediction model is used as the atmospheric driver. Important parameters include temperature, moisture, and winds throughout the

* Corresponding author at: Barcelona Supercomputing Centre, Carrer Jordi Girona 29, Barcelona 08034, Spain.
E-mail address: robert.banks@bsc.es (R.F. Banks).

http://dx.doi.org/10.1016/j.atmosres.2016.02.024
0169-8095/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

PBL and the PBL height (PBLH). Most applications to air pollution rely on high-resolution mesoscale meteorological models to provide accurate simulations of PBL parameters.

The major objective of this work is to provide a performance evaluation of boundary-layer variables such as near-surface temperature, relative humidity, and wind speed and PBLH simulated by different WRF PBL parametrization schemes for the application to air pollution modeling. This study aims to contribute to a reduction of one of the major sources of error in top-down estimates of photochemical pollutant modelling, boundary-layer representation of meteorological processes. In the current study, we aim to evaluate the operational definitions of PBLH for each scheme.

It is necessary with model horizontal grid spacing larger than 1×1 km to properly parameterize the vertical diffusion of surface fluxes, as they are connected with sub-grid-scale processes (Chen and Dudhia, 2000). PBL parametrization schemes fall into one of two main classes: local and non-local schemes. Some local closure schemes use turbulent kinetic energy (TKE) prediction, while most non-local schemes have diagnostic components for the K-profile (Troen and Mahrt, 1986) and PBL top. Above the PBL top, both local and non-local schemes also include vertical diffusion due to turbulence.

Previous studies have evaluated the performance of model PBL parametrization schemes in locations known for complex atmospheric situations (Pérez et al., 2006a; Bossioli et al., 2009; Banks et al., 2015). In Pérez et al. (2006a), the influence of three PBL schemes from the legacy Fifth Generation Penn State-NCAR Mesoscale Model (MM5; Dudhia, 1993) on meteorological and air quality simulations over Barcelona was analyzed. The authors found that the MM5 model tended to show a cold bias, with higher model-simulated wind speeds compared with observations, depending on the PBL scheme used.

In addition, Banks et al. (2015) evaluated WRF model-simulated PBLH over Barcelona using eight PBL schemes. Model-simulated PBLH was validated with PBLH estimates from a backscatter lidar during a 7-yr period. The authors determined that a non-local scheme such as the Asymmetrical Convective Model version 2 (ACM2) provide the most accurate simulations of PBLH, even under diverse synoptic flows such as regional recirculations. Banks et al. (2015) was data limited to the evaluation of only PBLH, while the current study compares surface and upper-air variables important for PBL applications.

Over the GAA, Bossioli et al. (2009) investigated the impact of four PBL schemes from the MM5 model on meteorological and air quality simulations. The authors found that the selection of PBL scheme shapes the horizontal and vertical extension of variables in the PBL. It was determined that non-local and semi non-local schemes were far superior to other schemes due to the favor of strong vertical mixing and transport towards the surface. Additionally, other studies have examined the performance of WRF PBL parametrization schemes in northern areas of Europe (Kim et al., 2013; Draxl et al., 2014), and over the continental United States (Hu et al., 2010; Coniglio et al., 2013).

Treatment of PBLH from both measurements and models is quite complex with many methods applied previously. LeMone et al. (2012) subjectively confirmed WRF model-simulated PBLH against modelled virtual potential temperature (Θ_v) profiles using different threshold values of $\delta\Theta_v/\delta z$. The most accurate model-simulated PBLH were compared against field observations. The authors found the best threshold defines PBLH as the lowest model level at which $\delta\Theta_v/\delta z = 2$ K km^{-1}, which works for all four PBL schemes they compared. In this work, we estimate PBLH from radiosoundings and lidar using well-tested methods to ensure quality comparisons to model simulations.

The current study will be presented as follows. Section 2 will describe the configuration of the WRF experiment sets and the various observation tools used to evaluate the model performance. Results of the performance evaluation will be presented in Section 3. Finally, a summary and conclusions will be shown in Section 4.

2. Data and methodology

2.1. WRF model configuration

In this study, we used WRF version 3.4.1 with the Advanced Research WRF (ARW) dynamical solver (Skamarock et al., 2005). Three model domains (Fig. 1) were configured with varying horizontal grid spacing at the parent European level (12×12 km; 481×401 grid points), and two one-way nested domains for the Greece (4×4 km; 202×202 grid points) and GAA (1×1 km; 101×101 grid points) regions. It is assumed that 1×1 km grid spacing is of fine enough detail to resolve most mesoscale features in the complex study area (Jiménez et al., 2013).

Final analysis (FNL) data from the National Centers for Environmental Prediction (NCEP) are used as the WRF initial and lateral boundary conditions, which are operational global analysis data available on $1^\circ \times 1^\circ$ grids at six-hourly time steps. FNL analyses are available from the surface and at 26 mandatory pressure levels from 1000 hPa to 10 hPa.

Daily WRF-ARW simulations were computed with a 36-h forecast cycle, including 12 h allotted for model spin-up time. Each day's simulation was initialized from 1200 UTC the previous day. The spin-up cycle is added to counter instability issues within the simulation, and the first 12 h of each forecast cycle is not included in the evaluation process. An output temporal resolution of 1-h was chosen for comparison with observations. The model was run with 38 terrain-following (ETA) vertical levels, of which 13 are located in the lowest 3 km of the atmosphere, with a model top set at 50 hPa.

The physics options selected include WRF single-moment 3-class microphysics (Hong et al., 2004), Kain–Fritsch cumulus parametrization (Kain, 2004), Dudhia shortwave radiation (Dudhia, 1989), rapid radiative transfer model longwave radiation (Mlawer et al., 1997), and the Noah land-surface model (Tewari and CoAuthors, 2004). No urban parameterization is used. More information about these physics options can be found in Skamarock and Klemp (2008).

The primary objective of this study is to evaluate the performance of boundary-layer variables simulated by WRF using varying PBL parametrization schemes. In version 3.4.1 of WRF-ARW, there is the option to choose from nine PBL schemes. Each PBL scheme is associated with one or more surface-layer schemes which provide the surface fluxes of momentum, moisture, and heat to the PBL scheme. Eight PBL schemes are evaluated here since the MYNN3 (Mellor–Yamada–Nakanishi–Niino Level 3) scheme shares similar characteristics to the MYNN2 (Mellor–Yamada–Nakanishi–Niino Level 2.5) scheme. An overview of the PBL schemes selected in this study is shown in Table 1. Also shown are the associated surface-layer schemes, another important source of error in WRF model simulations.

Five of the eight PBL schemes selected are tied to the MM5 and ETA surface-layer schemes, which are based on the Monin–Obukhov similarity theory. The other three PBL schemes use their own unique surface-layer schemes.

The PBL parametrization schemes selected consist of five local and three non-local closure schemes. The operational definition of PBLH in the individual schemes falls into one of two general classes. The first class calculates the PBLH using the Richardson bulk number (Ri_b) method from some predetermined starting level. The second class determine the PBLH at a level where the turbulent kinetic energy (TKE) profile decreases to some predefined threshold value. A brief description of the schemes follows.

2.1.1. Yonsei University (YSU) scheme

The first and most widely used PBL scheme is the Yonsei University (YSU) scheme (Hong et al., 2006). The YSU scheme is a first-order, non-local scheme with an explicit entrainment layer and a parabolic K-profile in an unstable mixed layer. It is a modified version of the Medium Range Forecast (MRF) scheme (Hong and Pan, 1996) from

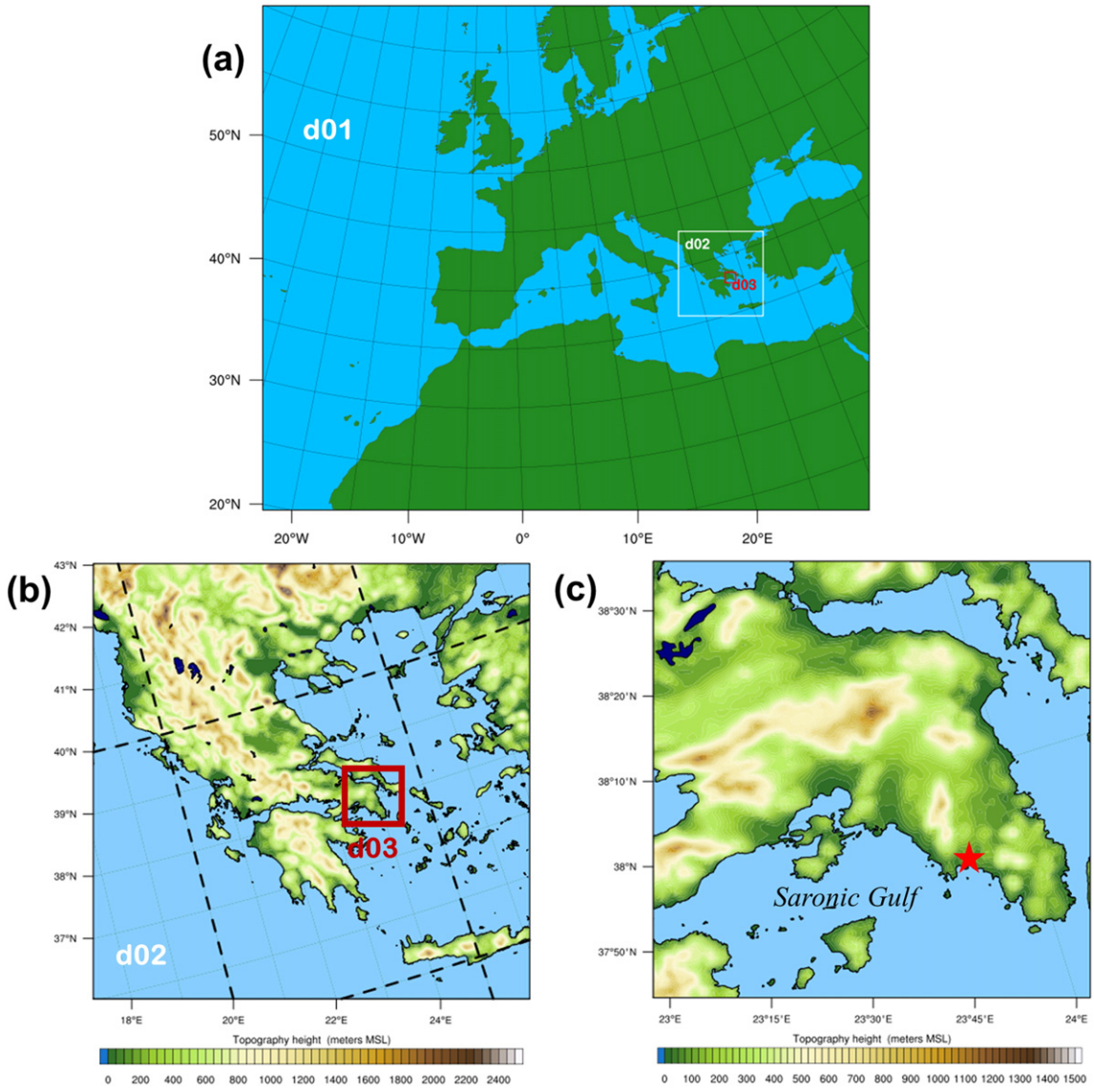

Fig. 1. Model domain configuration for (a) the European-level parent domain (d01, 12 × 12 km), (b) Greece domain (d02, 4 × 4 km), and (c) Greater Athens Area (GAA) domain (d03, 1 × 1 km). The Greece and GAA domains (red box) and associated topography are shown in 1b. The topographic map is further zoomed-in to the GAA domain (1c), with a bold red star denoting the location of the NTUA lidar site.

the MM5 (Dudhia, 1993). The largest improvement to the YSU scheme over the MRF scheme was the addition of an explicit term for the treatment of the entrainment zone. Hong (2010) implemented a modification to the scheme for the stable boundary layer. PBLH in the YSU scheme is determined from the Ri_b method but calculated starting from the surface. A threshold value of zero is used for stable cases, while 0.25 is used for unstable flow.

2.1.2. Mellor–Yamada–Janjic (MYJ) scheme

The next most widely used PBL scheme is the Mellor–Yamada–Janjic (MYJ) scheme (Janjić, 2002). The MYJ scheme is a one-and-a-half order prognostic TKE scheme with local vertical mixing. It is a modified version of the old ETA scheme from the MM5 model (Janjić, 1990). PBLH is determined from the TKE where the PBL top is defined where the profile decreases to a prescribed low value (0.2 $m^2\ s^{-2}$).

Table 1
Summary of WRF v3.4.1 experiment sets with PBL parametrization schemes (PBL SCHEME) and surface-layer schemes (SFC LAYER) used in study.

Set	PBL scheme	Short name	Closure type	SFC layer	PBLH definition
1	Yonsei University	YSU	1.0 non-local	MM5 similarity	Ri_b calculated from sfc
2	Mellor–Yamada–Janjic	MYJ	1.5 local	Eta similarity	TKE-prescribed threshold
3	Quasi-normal scale elimination	QNSE	1.5 local	QNSE	TKE-prescribed threshold
4	Mellor–Yamada–Nakanishi–Niino level 2.5	MYNN2	1.5 local	MYNN	TKE-prescribed threshold
5	Asymmetric Convective Model	ACM2	1.0 non-local	MM5 similarity	Ri_b calculated above neutral buoyancy level
6	Bougeault–Lacarrère	BouLac	1.5 local	MM5 similarity	TKE-prescribed threshold
7	University of Washington	UW	1.5 local	MM5 similarity	Ri_b threshold
8	Total energy–mass flux	TEMF	1.5 non-local	TEMF	Ri_b threshold

2.1.3. Quasi-Normal Scale Elimination (QNSE) scheme

The third scheme is the Quasi-Normal Scale Elimination (QNSE) scheme (Sukoriansky et al., 2005). The QNSE scheme is a one-and-a-half order, local closure scheme and has a TKE prediction option that uses a new theory for stably stratified regions. PBLH is defined as where the TKE profile decreases to a prescribed low value (0.01 $m^2 s^{-2}$), similar to the MYJ scheme.

2.1.4. Mellor–Yamada–Nakanishi–Niino Level 2.5 (MYNN2) scheme

The next scheme is the Mellor–Yamada–Nakanishi–Niino Level 2.5 (MYNN2) scheme (Nakanishi and Niino, 2006). The Mellor–Yamada–Nakanishi–Niino Level 3 (MYNN3) scheme shares similar characteristics to MYNN2, so it will not be evaluated here. The MYNN2 scheme is tuned to a database of large eddy simulations (LES) in order to overcome the typical biases associated with other MY-type schemes, such as insufficient growth of convective boundary layer and under-estimated TKE. The MYNN2 scheme is a one-and-a-half order, local closure scheme and predicts sub-grid TKE terms. PBLH is determined as the height at which the TKE falls below a critical value (1.0×10^{-6} $m^2 s^{-2}$).

2.1.5. Asymmetrical Convective Model version 2 (ACM2) scheme

The fifth scheme is the Asymmetrical Convective Model version 2 (ACM2) scheme (Pleim, 2007). The ACM2 scheme is a first-order, non-local closure scheme and features non-local upward mixing and local downward mixing. It is a modified version of the ACM1 scheme from the MM5 model, which was a derivative of the Blackadar scheme (Blackadar, 1978). The scheme has an eddy-diffusion component in addition to the explicit non-local transport of ACM1. PBLH is determined as the height where the Ri_b calculated above the level of neutral buoyancy exceeds a critical value ($Ri_{bc} = 0.25$). For stable or neutral flows the scheme shuts off non-local transport and uses local closure.

2.1.6. Bougeault–Lacarrère (BouLac) scheme

The following three schemes are lesser used and not as well tested. The first is the Bougeault–Lacarrère (BouLac) scheme (Bougeault and Lacarrere, 1989). The BouLac scheme is a one-and-a-half order, local closure scheme and has a TKE prediction option designed for use with the BEP (Building Environment Parametrization) multi-layer, urban canopy model (Martilli et al., 2002). BouLac diagnoses PBLH as the height where the prognostic TKE reaches a sufficiently small value (in the current version of WRF is 0.005 $m^2 s^{-2}$).

2.1.7. University of Washington (UW) scheme

The University of Washington (UW) scheme (Bretherton and Park, 2009) is a one-and-a-half order, local TKE closure scheme from the Community Earth System Model (CESM), climate model (Gent et al., 2011). PBLH in the UW scheme is defined as the inversion height between grid levels via a Ri_b threshold (0.25).

2.1.8. Total Energy–Mass Flux (TEMF) scheme

The Total Energy–Mass Flux (TEMF) scheme (Angevine et al., 2010) is a one-and-a-half order, non-local closure scheme and has a sub-grid-scale total energy prognostic variable, in addition to mass-flux-type shallow convection. TEMF uses eddy diffusivity and mass flux concepts to determine vertical mixing. PBLH is calculated through a Ri_b method with zero as a threshold value. In this study, there were minor stability issues with five simulation days using the TEMF scheme. All these days were characterised by low-level winds from the southwest. The stability issues are caused by a threshold exceedance of potential temperature over the desert regions in our parent domain. Decreasing the time between calls to the radiation physics scheme improved the stability for two of the five simulation days.

2.2. Identification of primary synoptic situations

The FLEXPART-WRF (FLEXible PARTicle) dispersion model (Brioude et al., 2012; Brioude and CoAuthors, 2013) was used to provide daily backward trajectories (backtrajectories) during the campaign. FLEXPART-WRF is driven by the 12 × 12 km atmospheric fields from the external WRF grid. The model is set up with ten vertical levels (100, 250, 400, 600, 900, 1500, 2000, 3000, 5000, and 10,000 m), and the output grid is at a horizontal grid spacing of 12 × 12 km. A total of 50,000 particles are released for each simulation and the 48-h backtrajectories are computed for the air masses arriving at 0.5, 1.0, 2.0, 3.0, and 4.0 km over Athens.

Based on visual inspection of the backtrajectories and WRF simulations, the atmospheric conditions during the campaign can be classified into one of three synoptic flow types: Continental, Etesians, and Saharan synoptic flows. Fig. 2 shows 2-day backtrajectories (a–c) from the FLEXPART-WRF dispersion model for representative cases of each flow type, along with the associated synoptic weather maps (d–f) simulated by the WRF model.

The predominant synoptic flow (41.7% of campaign) is Continental, which is influenced by winds from the west to northwest. A representative day for this synoptic flow is 25 May 2014 (Fig. 2a, d). The flow is characterized by a stagnant weather pattern with a weak atmospheric pressure gradient over Greece. On this particular day, a sea breeze develops in the afternoon, but this is not a mandatory feature for all cases of this flow type. Frequently, mesoscale processes dominate in this synoptic flow. This type of atmospheric situation can provoke the development of urban pollution episodes.

The second most prevalent synoptic flow during the campaign (36.1% of days) is the Etesians. This situation is influenced by winds coming from the north to northeast. A selected representative day of this pattern is 13 June 2014 (Fig. 2b, e). The Etesians are caused by a gradient between strong high pressure northwest of Greece and a low pressure area over Asia and these days are characterized by increased PBLH over Athens. The weakening of the Etesians synoptic flow allows for the development of local circulation systems (sea and land breezes), sometimes accounting for a decreased PBLH (Melas et al., 1995).

The third major synoptic flow observed during the campaign (22.2% of days) is Saharan type. In this condition, winds are dominant from the southwest with a stagnant atmospheric pressure pattern over Greece. Typically, this wind flow is most associated with dust intrusions from the Saharan source region in Africa. This was confirmed in a previous work by Papayannis et al. (2009) with the use of lidar measurements, satellite images, and a mineral dust transport model. A representative day during the campaign for this flow type is shown as 16 June 2014 (Fig. 2c, f). In this particular case, there is afternoon sea breeze initiation which may be attributable to a lower PBLH.

All data sets (lidar, WRF simulations, radiosoundings, and surface meteorological instruments) are grouped into these synoptic flow types for the analysis. The results are presented as a combination of averages, differences, and representative case studies.

2.3. PBL height from backscatter lidar

A backscatter lidar was operated during daytime hours of the campaign at the National Technical University of Athens (NTUA; 37.98 N, 23.78 E, 212 masl). The instrument at NTUA is a compact six-wavelength Raman lidar system used to perform measurements of suspended aerosol particles in the PBL (Kokkalis et al., 2012). The NTUA lidar station is a member of the European Aerosol Research Lidar Network (EARLINET; Bösenberg and CoAuthors, 2001).

Lidar data were collected for 36 out of 39 days of the month-long campaign (92.3% observation frequency). No lidar data were available on 19 May 2014 (observed thick stratus clouds) and 3–4 June 2014 (observed clouds and stratiform precipitation). The majority of the lidar data were measured from 0800 to 1500 UTC, with the earliest

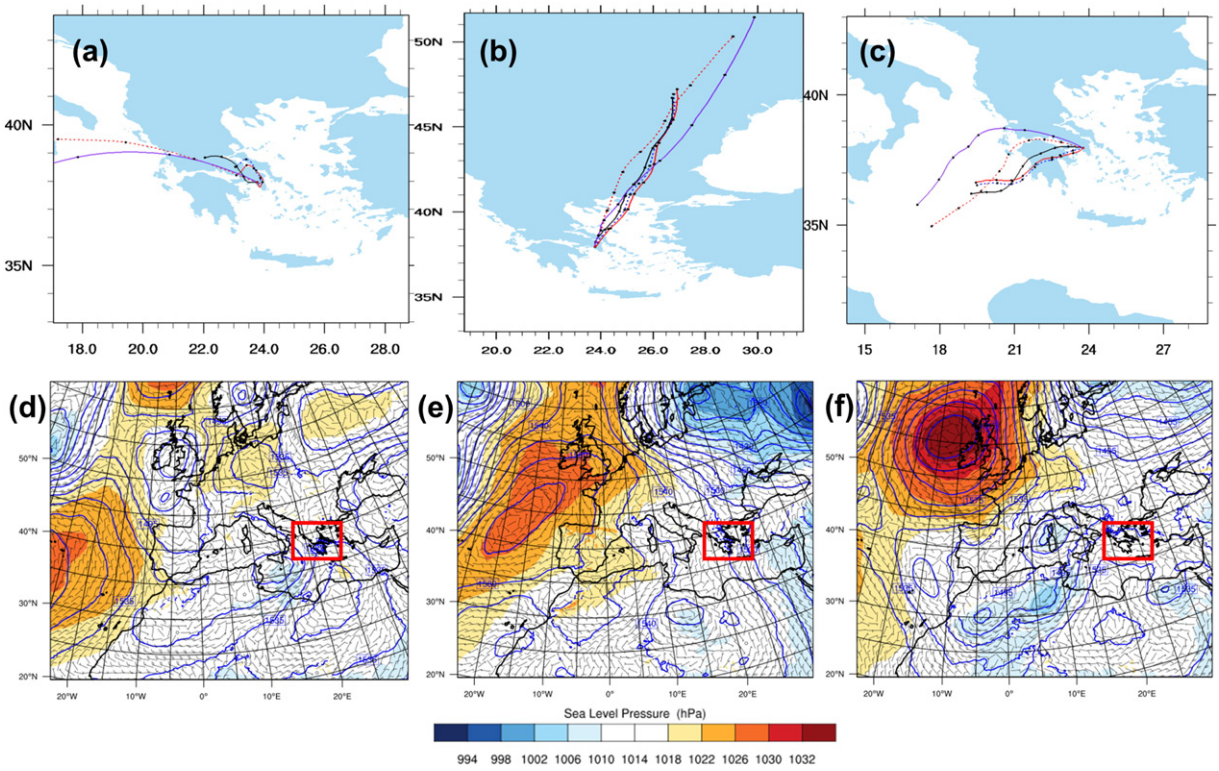

Fig. 2. Two-day FLEXPART-WRF backtrajectories (top) ending at Athens (37.96 N, 23.78 E) and synoptic maps simulated with the WRF model (bottom) representing (from left to right), Continental (25 May 2014), Etesians (13 June 2014), and Saharan (16 June 2014) flow types. Backward trajectories end at 1200 UTC with black dots showing 6-h position and colors denote trajectory height at receptor for 0.5 km (red), 1 km (blue dashed), 2 km (black), 3 km (red dashed), and 4 km (purple). Synoptic maps (valid 1200 UTC) show sea level pressure in hPa (colored contours), geopotential height at 850-hPa in metres (blue lines), and 850-hPa winds (barbs).

start at 0600 UTC and the latest finish around 1800 UTC. Resultant lidar data were visually quality-controlled for clouds in the PBL.

In this study, we use vertical profiles of the range-corrected backscattered power at 1064-nm using a temporal resolution of 100-s as a basis for the PBLH estimation. Tsaknakis et al. (2011) showed this wavelength provides sufficient transition between the suspended aerosol load in the PBL and the cleaner free troposphere (FT) above, as the molecular contribution in the near infrared becomes much smaller than that of the particles. The range resolution of the NTUA lidar is 7.5 m, with a full overlap window around 700 m.

PBLH is estimated from the lidar observations using an adaptive approach utilizing an extended Kalman filter (EKF) (Brown and Hwang, 1982). The technique has been developed and tested with simulated lidar retrievals (Rocadenbosch et al., 1998, 1999; Lange et al., 2014, 2015), and under various atmospheric conditions over Barcelona, Spain (Banks et al., 2014). Banks et al. (2015) showed the EKF technique is suitable for well-mixed convective boundary layers, after an extensive validation effort against classic methods of estimating PBLH from lidar and radiosondes.

The developed and tested EKF approach is based on estimating four time-adaptive coefficients of a highly simplified erf-like curve model, representing the PBL transition in terms of the RCS backscatter lidar signal. The erf-like model, $h(R)$, is formulated as follows:

$$h(R;R_{bl},a,A,c)=\frac{A}{2}\left\{1-erf\left[\frac{a}{\sqrt{2}}(R-R_{bl}\right]\right\}+c \quad (1)$$

where R_{bl} is an initial guess of the PBLH, a is the entrainment zone (EZ) scaling factor, A is the amplitude of the erf transition, and c is the average molecular background at the bottom of the FT. EKF state vector initialization also requires statistical covariance information from the user's side; this is accomplished by providing state vector noise and a priori error covariance matrices (in turn, related to the initial state vector). If the state vector $[R_{bl}\ a\ A\ c]$ is not initialized correctly, one can expect not so reliable estimates of PBLH.

Lange et al. (2015) explained why the present implementation of the EKF uses both inner (R'_1, R'_2) and outer (R_1, R_2) range boundaries of the erf-like model. These ranges are allowed to change adaptively with the estimated PBLH, a computational advantage in cases where the PBLH changes sharply between adjacent time steps. The only requirement is that the inner and outer range intervals defined contain erf transition and erf plateau characteristics.

Also, the EKF method benefits from the Kalman gain, which corrects the projection trajectory of the PBL atmospheric variables and improves its estimation of the PBL parameters via a new state vector. The statistical covariance information, along with the state vector, observation-noise covariance, and Kalman gain, is updated recursively at each iteration of the filter.

Fig. 3 shows the first 100-s lidar profile from 25 May 2014 started at 0701 UTC, which is used to initialize the EKF state vector. In this case, the signal-to-noise ratio is >5 up to 3 km altitude. For better numerical conditioning and physical significance, the observation vector presented to the filter is a molecular-normalized version. The initial state vector parameters are subjectively selected from visual inspection of the first profile. For this case, we annotated on the figure as $R_{bl} = 0.60$ km, $a = 9.23\ \text{km}^{-1}$, $A' = 4.0 \times 10^4$, and $c' = 2.0 \times 10^4$. Additionally, $R_1 = 0.26$ km and $R_2 = 0.75$ km, while $R'_1 = 0.375$ km and $R'_2 = 0.675$ km. We have to note that it is only necessary to provide the initial EKF state vector parameters in general terms. The convergence time to a reliable solution will depend on the complexity of the lidar scene, the initial state vector estimates, and the state-noise and error covariance matrices.

Individual daily (24-h) WRF model simulations were run for the same 36 days lidar-EKF estimates were calculated. This data set is

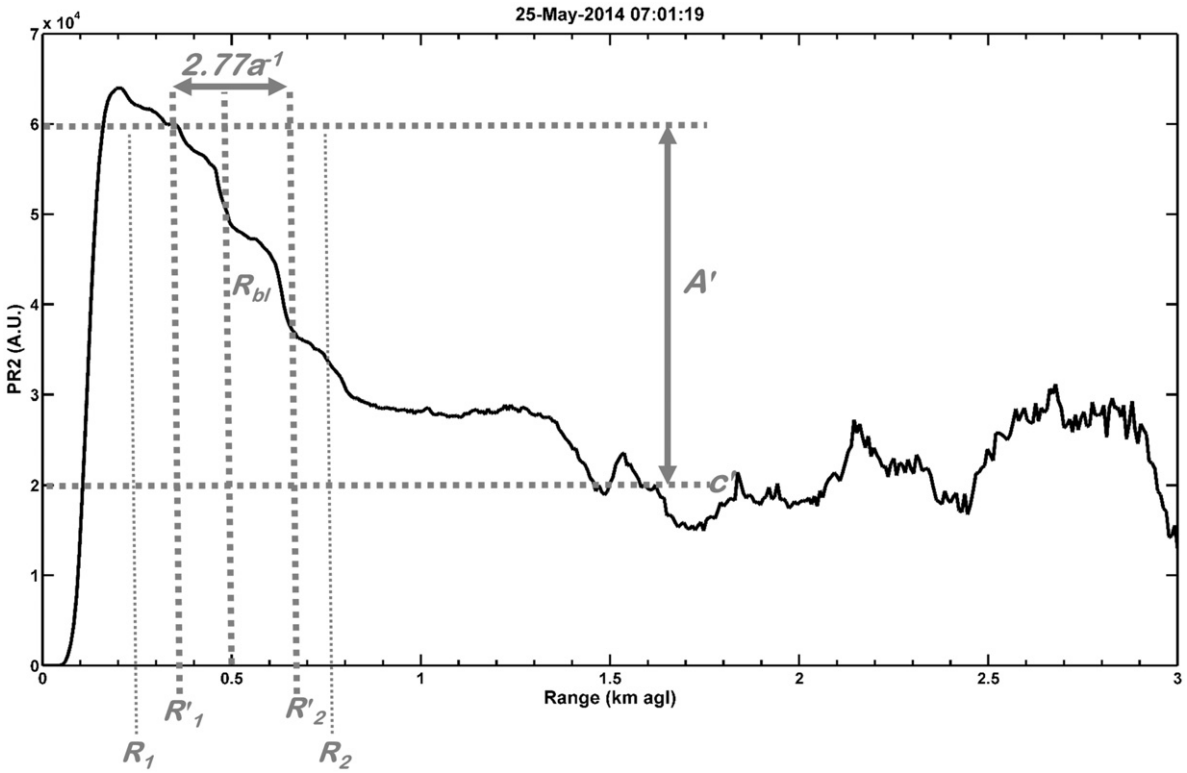

Fig. 3. The 100-s profile of molecular-normalized range-corrected power from the NTUA lidar at 0701 UTC on 25 May 2014. Initial state vector [R_{bl} a A' c'] with lower- and upper-range limits [R_1 R_2] for the extended Kalman filter (EKF) PBL height technique are annotated in gray dashed lines.

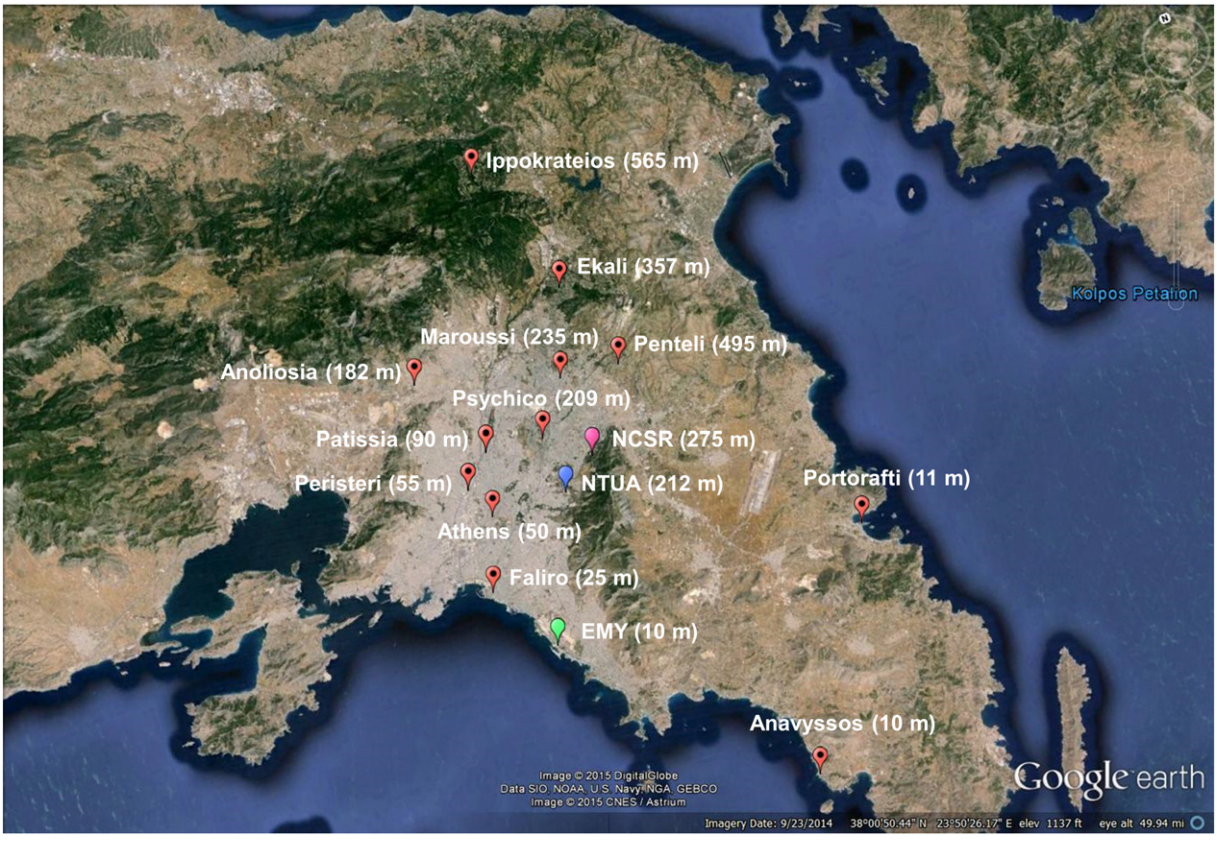

Fig. 4. Location and elevation (meters above sea level) of the 14 surface meteorological stations (12 from the National Observatory of Athens, one from NTUA and one from NCSR) and site of radiosonde launches (EMY). Spatial variations range from 500 m to 12 km apart and in elevation from 10 to 565 m asl.

used for the evaluation of simulated PBLH from the PBL parametrization schemes. For the comparison with WRF model-simulated values, a 15-min average (nine 100-s estimates) of lidar-EKF PBLH is applied, centred on the hourly model output times. PBLHs estimated by lidar and simulated with WRF are shown in km asl throughout this study.

2.4. Surface meteorological observations and radiosoundings

Near-surface meteorological variables are important for the evaluation of PBL schemes as they represent lower boundary-layer processes. We collected 2-m air temperature (T2) and relative humidity (RH2), and 10-m wind speed (WS10) from 14 surface meteorological stations. It should be noted that wind speed measurements are extrapolated to 10-m from the individual station height using the wind power law. Twelve of the stations are associated with the National Observatory of Athens (NOA) network, in addition to one station situated at the NTUA Physics Department (lidar site), and one station located at the National Center for Scientific Research, Demokritos (NCSR). Fig. 4 shows the geographical location and elevation of the measurement stations. The locations were selected for a diversified mix of geographical influences, ranging from 10 m asl at Anavyssos to 565 m asl at Ippokrateios. The mean difference between WRF model-grid height and real terrain height was 3.5 m asl, indicating reliable topographic influences from the WRF model.

In order to evaluate vertical profiles, 13 radiosondes were launched from the Hellenic National Meteorological Service (EMY) during the campaign. EMY is located near the sea (37.88 N, 23.73 E, 10 m asl) approximately 12 km away from the lidar site. The launches were scheduled around 1100 UTC. The Vaisala RS92 radiosonde instrument records atmospheric variables of temperature (°C), dew point temperature (°C), relative humidity (%), wind speed (m s^{-1}) and direction (°), and barometric pressure (hPa).

In this study, PBLH is calculated from the radiosounding data using the bulk Richardson number (Ri_b) method (Holtslag et al., 1990). It is the same method used in many of the WRF PBL schemes (Section 2.1) to diagnose the PBLH. The Ri_b approach requires wind speed and direction, barometric pressure, and temperature as input variables at each altitude (m). The Ri_b method is a proxy of where the wind transitions from turbulent to laminar, possibly indicating the top of the PBL. PBLH is calculated at the altitude where Ri_b exceeds a so-called critical Richardson number (Ri_{bc}).

From many previous studies, the Ri_{bc} is selected as a universal constant anywhere between 0.1 and 1.0 (Richardson et al., 2013). Typically, higher critical values are selected in areas where the flow transition from turbulent to laminar is larger. In this work, many critical values were tested by visually inspecting vertical profiles of potential temperature. We found that a critical value of 0.25 provides reasonable PBL height estimates, which is similar to the critical value used in past works.

2.5. Metrics used for model validation

Statistical measures include the mean bias (MB), root mean squared error (RMSE), coefficient of determination (R^2), and standard deviation (STDEV). Results are organized according to the campaign average, and averages for the three synoptic flows are observed during the campaign. The innermost domain from the WRF model runs is subjected to the evaluation.

Significance of the aggregated statistical metrics presented in Tables 2 and 3 are performed through two-sample *t* tests (Wilks and Department of E. and A.S.C.U., 2006a, 2006b) for each WRF PBL scheme and variable evaluated. For example, eight separate *t* tests were executed for the bias (WRF model—observed) of 2-m air temperature, one for each WRF PBL scheme. The assumption is the data are approximately Gaussian, with the tests using a two-tailed distribution. The significance for each individual metric is determined from the *t* test with a *p*-value, which is the two-tailed decimal probability of the hypothesis being false. A low *p*-value indicates high confidence in rejecting the null hypothesis, which here is hypothesized that the mean (μ) of a particular metric is equal to zero. The outcomes of the *t* tests show *p*-values well below 0.001 for all tests, which is clearly significant at greater than 95% confidence interval.

It should be noted that some assumptions of independence could be violated due to the aggregation methodology of the data. Simulations from the WRF model and the observational data are aggregated over the full diurnal cycle, and over multiple observing stations, which may possibly introduce additional bias. In addition, the presence of gross outliers could have large effects on the computations of standard deviation.

Table 2
Statistics of 2-m air temperature (T2) and relative humidity (RH2), and 10-m wind speed (WS10) between WRF PBL schemes and observations ($N = 12{,}096$). Performance indicators: mean bias (MB) and standard deviation (STDEV), coefficient of determination (R^2), and root mean squared error (RMSE) for the campaign average.

Variable	Scheme	MB (STDEV)	R^2	RMSE
T2 (°C)		(°C)		(°C)
	YSU	−0.56 (1.18)	0.79	2.24
	MYJ	−0.23 (1.30)	0.78	2.25
	QNSE	−0.68 (1.57)	0.77	2.55
	MYNN2	−1.14 (1.04)	0.79	2.48
	ACM2	−0.32 (1.19)	0.79	2.17
	BOULAC	−0.32 (1.12)	0.8	2.09
	UW	−0.5 (1.15)	0.79	2.18
	TEMF	−0.87 (1.26)	0.73	2.67
RH2 (%)		(%)		(%)
	YSU	0.57 (3.77)	0.36	12.47
	MYJ	1.99 (6.74)	0.38	13.08
	QNSE	0.57 (4.82)	0.38	12.64
	MYNN2	−0.61 (6.62)	0.39	13.65
	ACM2	−2.31 (3.90)	0.4	12.1
	BOULAC	−0.41 (3.76)	0.35	12.24
	UW	0.05 (3.72)	0.36	12.3
	TEMF	4.65 (4.50)	0.28	14.83
WS10 (m s^{-1})		(m s^{-1})		(m s^{-1})
	YSU	2.33 (1.21)	0.39	3.1
	MYJ	3.1 (1.01)	0.42	3.86
	QNSE	3.17 (1.02)	0.41	3.98
	MYNN2	2.69 (1.23)	0.44	3.4
	ACM2	2.58 (1.17)	0.41	3.35
	BOULAC	2.5 (1.07)	0.35	3.32
	UW	2.34 (1.13)	0.39	3.15
	TEMF	2.42 (1.19)	0.29	3.23

3. Results and discussion

3.1. Near-surface meteorological variables (T2, RH2, and WS10)

First, WRF model-simulated near-surface meteorological variables are validated against measurements from the GAA surface network. Results are presented for representative coastal and inland locations, then as a statistical set of the 14 total stations.

Fig. 5 shows the campaign-averaged MB (WRF model—observed) for the surface meteorological variables at two locations: Anavyssos (37.73 N, 23.91 E, 10 m asl) and Peristeri (38.00 N, 23.70 E, 55 m asl), representing coastal and inland locations, respectively. WRF model-simulated T2 (Fig. 5a) and RH2 (Fig. 5c) at Anavyssos during daytime (defined as: 0800–1800 UTC) shows a systematic cold (≈1–3 °C), moist (up to 23%) bias among PBL schemes, while at Peristeri (Fig. 5b, d) the errors are much smaller (T2: ≈±1 °C and RH2: ≈ −9% to 5%). The largest spread between the PBL schemes is observed with RH2 at Anavyssos during daytime, with the QNSE and MYNN2 schemes fitting most closely to the observations between 1000 and 1700 UTC. The WRF model shows a consistent over-representation of WS10 (Fig. 5e, f) at both locations, as high as 6 m s^{-1} at Anavyssos. The error in WS10 is fairly constant throughout the diurnal cycle.

Table 3
Similar to Table 2, except calculated according to Continental synoptic flow ($N = 5040$), Etesians synoptic flow ($N = 4368$), and Saharan synoptic flow ($N = 2688$) types.
Shown on the following page.

Variable	Scheme	Continental			Etesians			Saharan		
		MB	R^2	RMSE	MB	R^2	RMSE	MB	R^2	RMSE
T2 (°C)		(°C)		(°C)	(°C)		(°C)	(°C)		(°C)
	YSU	−0.71	0.82	2.15	−0.76	0.73	2.25	0.06	0.74	2.33
	MYJ	−0.38	0.82	2.22	−0.46	0.73	2.18	0.41	0.74	2.39
	QNSE	−0.69	0.81	2.47	−1.01	0.72	2.54	−0.14	0.71	2.69
	MYNN2	−1.32	0.82	2.44	−1.23	0.75	2.47	−0.66	0.73	2.52
	ACM2	−0.41	0.82	2.08	−0.54	0.74	2.16	0.2	0.75	2.31
	BOULAC	−0.54	0.82	2.08	−0.41	0.76	1.99	0.24	0.77	2.25
	UW	−0.62	0.82	2.1	−0.75	0.74	2.22	0.1	0.77	2.2
	TEMF	−1.16	0.74	2.76	−0.97	0.72	2.46	0.15	0.65	2.61
RH2 (%)		(%)		(%)	(%)		(%)	(%)		(%)
	YSU	3.52	0.29	12.88	−0.26	0.4	11.74	−3.62	0.43	12.58
	MYJ	4.93	0.32	13.87	1.64	0.49	11.54	−2.97	0.39	13.47
	QNSE	2.72	0.32	13.31	0.55	0.47	11.24	−3.42	0.42	13.23
	MYNN2	1.59	0.34	14.02	−0.56	0.45	12.6	−4.83	0.45	14.27
	ACM2	−0.07	0.32	11.95	−2.49	0.46	11.43	−6.21	0.48	13.05
	BOULAC	2.71	0.28	12.65	−1.33	0.42	11.05	−4.75	0.39	13.02
	UW	2.67	0.3	12.38	−0.17	0.43	11.51	−4.52	0.4	13.09
	TEMF	9.33	0.23	16.95	1.85	0.4	11.81	−2.05	0.42	12.72
WS10 ($m\ s^{-1}$)		($m\ s^{-1}$)		($m\ s^{-1}$)	($m\ s^{-1}$)		($m\ s^{-1}$)	($m\ s^{-1}$)		($m\ s^{-1}$)
	YSU	2.17	0.38	3.05	2.48	0.45	3.09	2.39	0.37	3.13
	MYJ	2.93	0.41	3.79	3.38	0.46	3.97	2.97	0.4	3.73
	QNSE	2.99	0.38	3.91	3.43	0.46	4.07	3.1	0.4	3.88
	MYNN2	2.53	0.43	3.34	2.91	0.48	3.51	2.6	0.42	3.28
	ACM2	2.4	0.39	3.23	2.86	0.46	3.44	2.46	0.4	3.31
	BOULAC	2.19	0.35	3.09	2.96	0.37	3.54	2.36	0.33	3.28
	UW	2.12	0.37	3.06	2.57	0.45	3.17	2.37	0.36	3.21
	TEMF	2.19	0.26	3.17	2.65	0.36	3.14	2.48	0.42	3.32

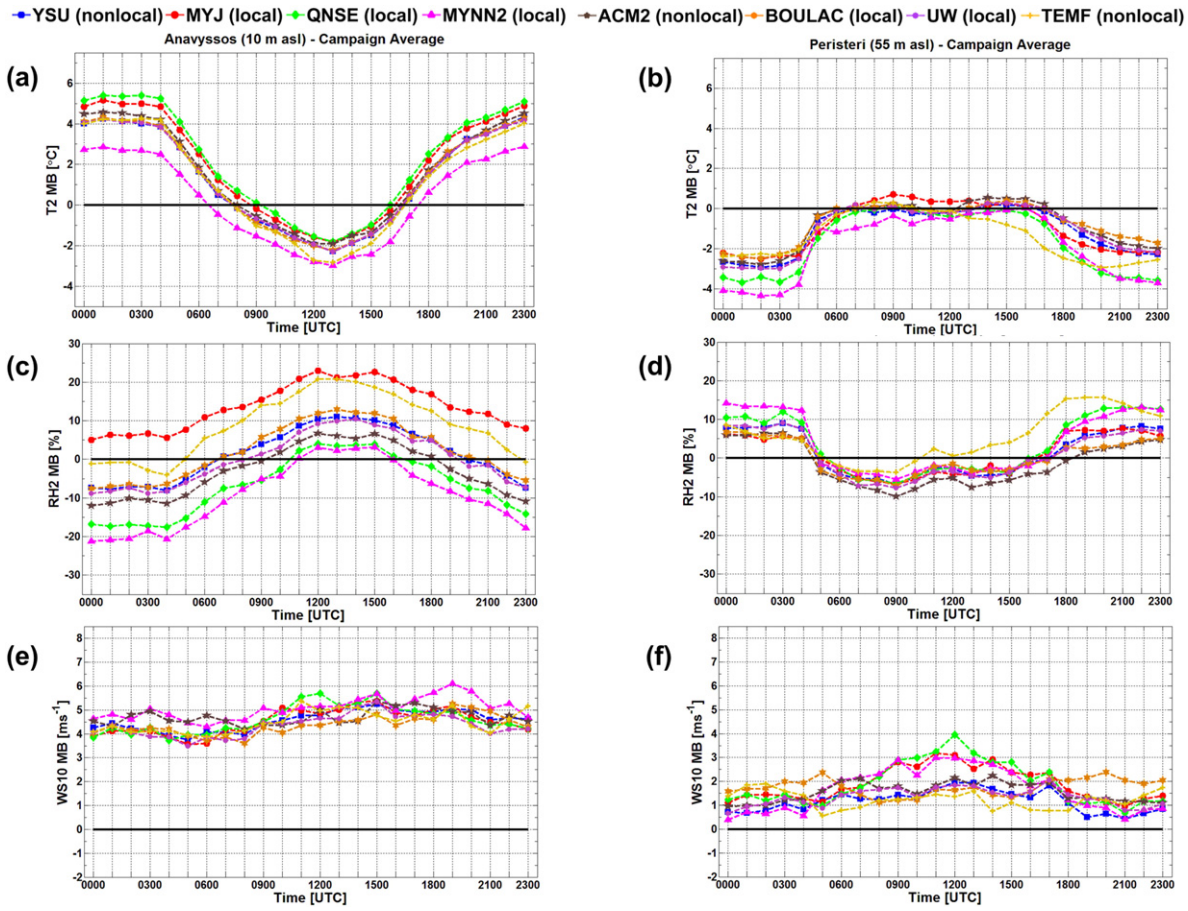

Fig. 5. Campaign-averaged mean bias (MB, WRF—observation) for (a, b) 2-m air temperature (T2, °C), (c, d) 2-m relative humidity (RH2, %), and (e, f) 10-m wind speed (WS10, $m\ s^{-1}$) at two stations (Anavyssos and Peristeri) from the NOA network, representative of coastal and inland influences, respectively.

In order to investigate the performance of WRF model PBL schemes in greater detail, Table 2 shows the performance indicators of the surface meteorological variables between each WRF PBL scheme and the observations. The results in Table 2 have been computed to represent all 14 stations for the campaign average ($N = 12{,}096$) over the full diurnal period. Overall, the performance of WRF PBL schemes is quite ambiguous depending on the meteorological parameter being analyzed. The BouLac scheme shows the best performance with respect to T2 ($R^2 = 0.8$ and RMSE 2.09 °C), while the MYJ scheme has the lowest MB (−0.23 °C). The BouLac scheme also performs well in simulating RH2 (MB = −0.41%), but the ACM2 scheme shows a closer correspondence ($R^2 = 0.4$) and RMSE (12.1%) to the observed values. All PBL schemes consistently over-estimate WS10, with the non-local YSU scheme having the lowest error (MB = 2.33 m s^{-1}).

The differences between model simulations and observations have influences from the underlying daily changes in atmospheric dynamics. The three primary synoptic situations identified with the FLEXPART-WRF backtrajectories (Fig. 2) can be used to further analyse the impacts of the meteorological situation. These results are shown in Figs. 7–9, again using Anavyssos and Peristeri as proxy locations for coastal and inland impacts, respectively.

First are the results when averaging over the days influenced by Continental synoptic flow (Fig. 6). The overall impression of MB for both locations looks similar to the campaign average; however, subtle differences are noted for T2 (Fig. 6a, b) and RH2 (Fig. 6c, d). At Peristeri, the TEMF scheme shows a cooler, moister bias after 1200 UTC (T2: up to −3 °C and RH2: up to 25%) progressing into the evening. All PBL schemes show a slight improvement with WS10 (≈1 m s^{-1}) when compared to the campaign average (Fig. 6e, f), especially in the morning hours.

Table 3 shows the performance indicators of the surface meteorological variables between each WRF PBL scheme and the observations for each synoptic flow type. Statistics are calculated to represent all 14 stations averaged for each type. STDEV calculated for each synoptic flow was similar to the campaign average, and is not shown in the table.

For Continental synoptic flow ($N = 5040$), the ACM2 and BouLac schemes show the closest representation (MB = −0.41 °C and RMSE = 2.08 °C, respectively) to observed T2 values. Also, the ACM2 scheme performs well with respect to RH2 (MB = −0.07%). The lowest errors in simulated WS10 are shown for the YSU (2.17 m s^{-1}) and UW (2.12 m s^{-1}) schemes.

The next most frequent atmospheric flow type is the Etesians. The MB of surface meteorological variables are averaged for days in this synoptic flow group for Anavyssos and Peristeri stations (Fig. 7). It is shown that daytime simulated T2 at Anavyssos (Fig. 7a) is around 0.5 °C colder than the campaign average. The spread of T2 between the PBL schemes is similar, with MYNN2 being the coldest. With respect to simulated RH2 (Fig. 7c, d), the differences between the PBL schemes are largest at Anavyssos. During daytime, the WRF model tends to reproduce an RH2, which is too moist (5–25%) at Anavyssos and too dry (5–10%) at Peristeri. Simulated values of WS10 (Fig. 7e, f) continue to show over-estimates with all schemes, up to 6 m s^{-1} with MYNN2.

Table 3 shows the performance indicators of the surface meteorological variables between each WRF PBL scheme and the observations, now computed to represent all 14 stations averaged for the synoptic flow group, Etesians ($N = 4368$). The BouLac scheme reproduces the closest T2 (MB = −0.41 °C) when compared with other schemes. The closest schemes with respect to the MB of RH2 are UW (−0.17%), followed by QNSE (0.55%). Again, WS10 is the most well reproduced by the YSU scheme (MB = 2.48 m s^{-1}).

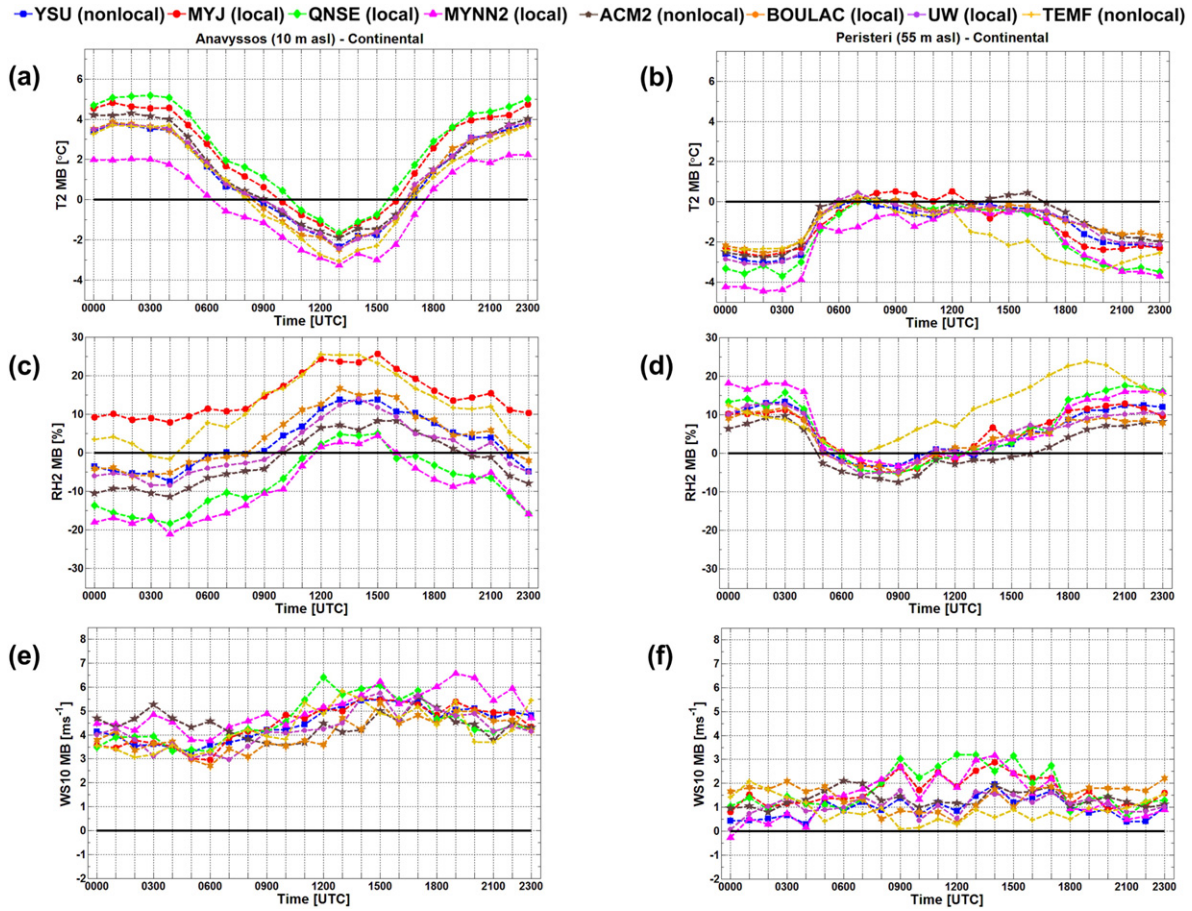

Fig. 6. Same as in Fig. 5, but time-averaged for Continental synoptic flow days.

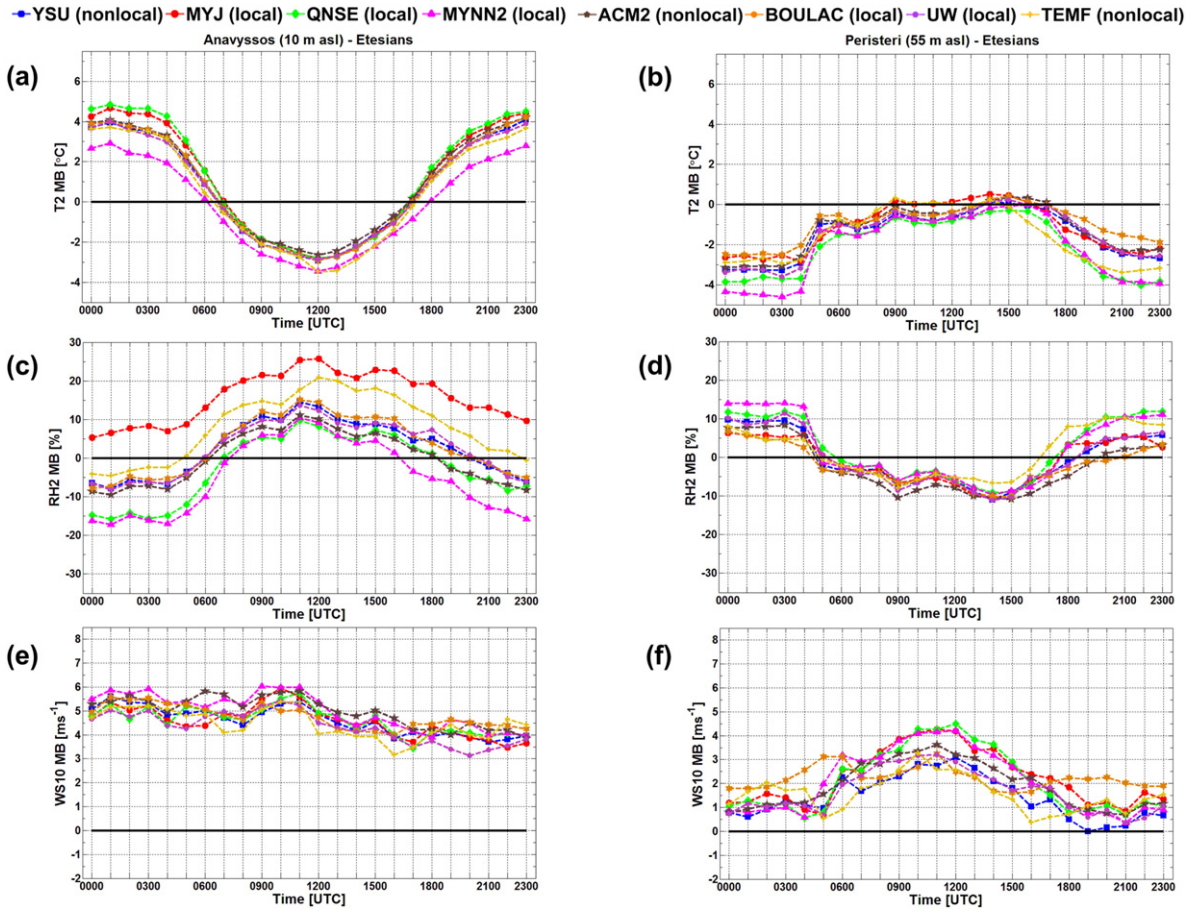

Fig. 7. Same as in Fig. 5, but time-averaged for Etesians synoptic flow days.

In the final synoptic flow type, Saharan, we show the MB of surface meteorological variables for Anavyssos and Peristeri stations (Fig. 8). In this synoptic situation, simulated T2 at both locations (Fig. 8a, b) increases by approximately 1 °C, resulting in an improvement at Anavyssos, but a degradation at Peristeri. The difference is most notable during daytime. Wind flow from the southwest also promotes a drier environment than the campaign average, which is shown in Fig. 8c, d. Simulated RH10 at Peristeri is very well reproduced with a ± 10% error. Finally, simulated WS10 at Anavyssos (Fig. 8e) shows a large deviation from the observations (>4 m s^{-1}) after 1200 UTC, which is approximately 2 m s^{-1} higher than the campaign average.

The performance indicators for surface meteorological variables between individual WRF PBL schemes and the observations, calculated to represent all 14 stations averaged for the Saharan synoptic flow group are shown in Table 3. The UW and BouLac schemes correspond closest (R^2 = 0.77 for both) to the observed T2 values, with the UW and YSU schemes having the lowest MB (0.10 and 0.06 °C, respectively). The TEMF scheme simulates RH2 closest (MB = −2.05% and RMSE = 12.72%) to the observed values, even though it is a poor performer with the other two synoptic flow types and the campaign average. With respect to WS10, the BouLac scheme slightly outperforms (MB = 2.36 m s^{-1}) other schemes; however, YSU is a close second (RMSE = 3.13 m s^{-1}).

In summary, the campaign-averaged near-surface variables showed that the WRF model tended to have a systematic cold, moist bias during daytime, most prominent at the coastal location. The BouLac scheme reproduced T2 and RH2 well with the campaign average, and with Etesians synoptic flow. ACM2 showed the closest T2 and RH2 during Continental flows. With Saharan synoptic flows, the UW and BouLac schemes well-represented T2, while TEMF best-reproduced RH2. WRF with the YSU scheme showed the closest WS10 to the observed values with the campaign average, and during Continental and Etesian synoptic flows. The BouLac scheme only slightly outperformed YSU during Saharan events.

3.2. PBL height (PBLH) intercomparison

In this section, we show the results for the evaluation of the PBLH. First will be the comparison between PBLH estimated from lidar and radiosoundings. Next, we use the lidar-EKF estimates of PBLH to perform an evaluation of WRF model-simulated PBLH using the different PBL schemes. Finally, we show representative cases of lidar and WRF model PBLH for each synoptic flow to discuss the strengths and weaknesses of the lidar-EKF technique.

3.2.1. PBLH from lidar and radiosoundings

PBLH is estimated from daytime lidar measurements using the EKF technique. Lidar-estimated PBLH was compared with 13 radiosounding-derived estimates using a bulk Richardson number method (Holtslag et al., 1990). The MB (lidar–radiosonde) shows that the lidar-EKF overestimated PBLH around 0.40 km at the EMY site. Possibly, this can be explained by the approximate 15 km distance between NTUA and EMY. The mean difference between WRF model-simulated PBLH at EMY and NTUA was 0.56 km for the campaign. The stable nocturnal boundary layer is not evaluated in this work due to a high complete overlap region (700 m) of the lidar instrument.

Over-estimates by the lidar were largest when the daytime (1100 UTC) PBLH was above 2 km. The performance of PBLH between lidar and radiosoundings would most likely be better with more launches from EKPA (average radiosonde PBLH = 2.29 km, N = 2), which has a

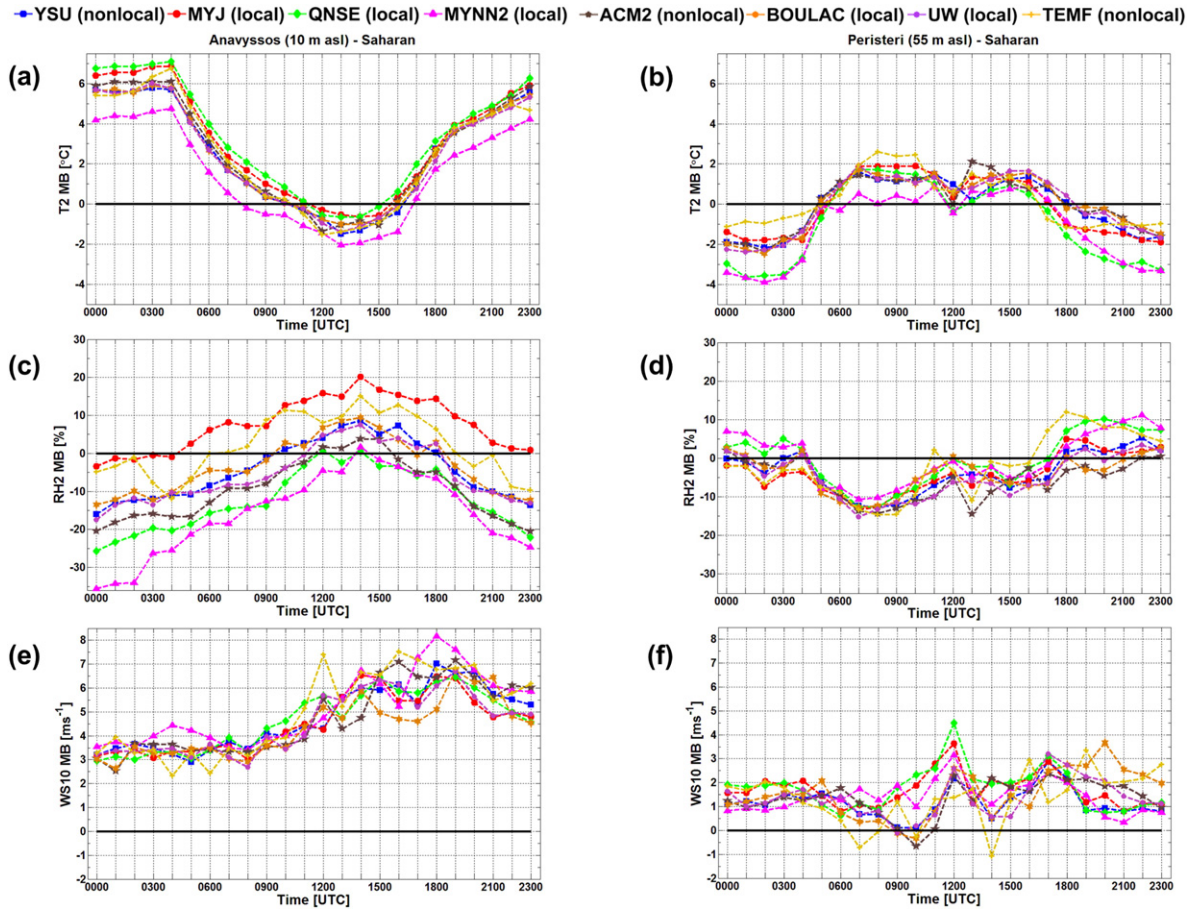

Fig. 8. Same as in Fig. 5, but time-averaged for Saharan synoptic flow days.

similar atmospheric background to NTUA. In the following, PBLH retrieved from the lidar are used to evaluate model-simulated PBLH from WRF. Strengths and limitations of the lidar-EKF method are discussed at the end of the section.

3.2.2. Comparison of WRF model-simulated PBLH against lidar

Fig. 9 shows the full diurnal cycle of WRF model-simulated PBLH from the different PBL schemes against the hourly averaged lidar-estimated PBLH. The hourly standard deviation of lidar-estimated

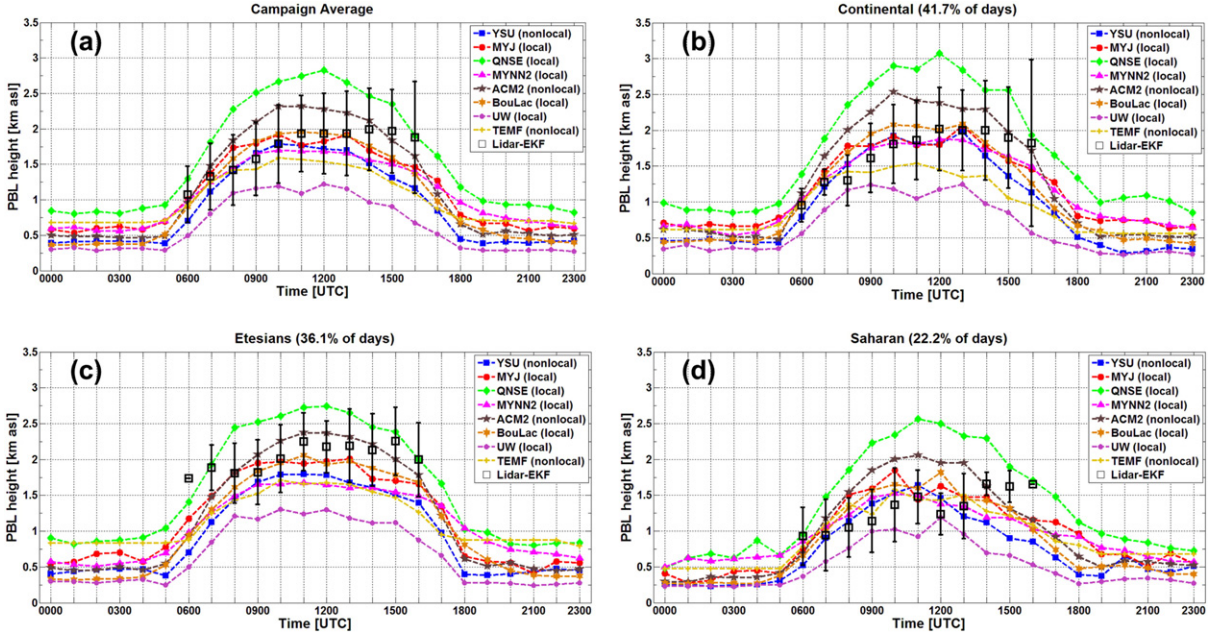

Fig. 9. Hourly mean PBL height (PBLH) comparisons between the NTUA lidar (open black squares) and WRF PBL schemes (colored lines) for (a) campaign average, and (b) Continental flow, (c) Etesians flow, and (d) Saharan flow types. PBLH estimated by lidar with extended Kalman filter (EKF) technique.

PBLH is represented with error bars computed by the total number of lidar estimates at each synoptic hour. Concerning the campaign average (Fig. 9a), the spread among the PBL schemes is quite large, with differences approaching 1.5 km between the lowest (UW ≈ 1.25 km) and highest (QNSE ≈ 2.75 km) simulated daytime maximum PBLH. The spread between the schemes is reduced by around 50% during the nighttime and early morning hours. In addition, PBLHs reproduced with the local UW and QNSE schemes are the lowest and highest, respectively, for all synoptic flow conditions.

The local BouLac scheme follows close to the lidar estimates during the campaign average (Fig. 9a) and Continental flows (Fig. 9b), while the non-local ACM2 scheme captures the PBLH well during the Etesians (Fig. 9c). Lidar-estimated PBLH during Etesians shows a faster growing boundary layer than other synoptic flow types, with an average PBLH already around 1.75 km at 0600 UTC. The lowest PBLHs observed by the lidar occur during Saharan synoptic flows (Fig. 9d), where the daytime maximum struggles to reach around 1.5 km. PBLH calculated from radiosoundings was as low as 0.37 km. Low PBLHs during Saharan events could be caused by negative radiative feedback from the dust aerosols (Pérez et al., 2006b). Between 0600 UTC and 1300 UTC, the non-local TEMF scheme reproduces the closest heights during Saharan flows.

Table 4 shows the performance indicators (MB, STDEV, RMSE, R^2) of PBLH between individual WRF PBL schemes and the estimates from lidar using the EKF method. The performance indicators have been calculated to represent the campaign average ($N = 286$) and the three synoptic flow groupings: Continental ($N = 125$), Etesians ($N = 101$), and Saharan ($N = 60$). Again, considerable differences are evident between model-simulated PBLH from the PBL schemes.

The determination coefficient (R^2) between model-simulated and lidar-estimated PBLH is practically insignificant for all PBL schemes, with the MYNN2 and ACM2 schemes showing the highest correspondence ($R^2 = 0.15$ for both) to the observed values with the campaign average. The MYNN2 scheme shows slightly better correlations ($R^2 = 0.22$) with the Continental and Saharan synoptic flows. The BouLac scheme follows with the third best correlation ($R^2 = 0.13$) based on campaign totals. The UW and TEMF schemes have the lowest correlations ($R^2 = 0.05$ and 0.09, respectively) against the lidar-estimated PBLH, consistent in all synoptic flow types and the campaign average.

The MB and RMSE show better results than the goodness of fit approximations for evaluating the schemes. Simulations with the BouLac scheme only slightly under-estimate PBLH (MB = −0.12 km) and RMSE (0.65 km) with the campaign average, in addition to the Continental synoptic flow (MB = −0.02 km, RMSE = 0.68 km). The MYJ scheme shows the lowest MB with the campaign average (−0.11 km) and the second lowest during Continental (−0.10 km) and Etesians (−0.30 km) flows. The ACM2 scheme performs the best during the Etesians with a slight under-estimate (MB = −0.11 km), while the TEMF scheme reproduces the best PBLH during Saharan synoptic flows (MB ≈ 0 km, RMSE = 0.58 km).

3.2.3. Representative cases of synoptic flows observed

A few examples of the lidar-EKF technique are shown in Fig. 10 to demonstrate the strengths and limitations of the method for PBLH detection. The corresponding WRF model-simulated PBLH are also shown, for additional analysis of the model results presented earlier in the section. Both the lidar and the model results are shown for representative cases of the three synoptic flow types.

First is a Continental synoptic flow on 25 May 2014 (Fig. 10a), a complex case characterized by winds from the northwest. There are a few sharp gradients below 2.5 km, and these are probably due to stratification of the mixed aerosol layers. The additional lofted layer between 1.5 and 2 km around 12 UTC is coupled with the PBL top, most likely due to hygroscopic growth. From 0701 to 0900 UTC the residual layer from the previous day is still evident (light green shading) between 1 km and 1.5 km. An additional aerosol layer (green and yellow shading) around 3 km altitude at 0701 UTC slowly descends towards the boundary-layer top during the day, merging into the boundary layer around 1100 UTC. PBLH estimated with the EKF method is around 0.90 km at 0800 UTC, growing to 1.83 km by 1200 UTC. Lidar-EKF PBLH estimates are considered reasonable as they fit the erf-like curve in the individual lidar profiles. However, a narrow transition amplitude was selected to avoid complications from the coupled aerosol layer later in the day. The daily mean standard deviation (1σ) of the hourly lidar retrievals is 19.87 m, indicating reliable hourly estimates.

The lidar-estimated PBLH compares well with model-simulated PBLH from WRF with Continental flow (Fig. 10d). All PBL schemes show an MB less than 0.50 km, except for the QNSE and UW schemes.

Table 4
Statistics for PBL height showing performance between eight WRF PBL schemes and lidar-EKF estimates. Performance indicators include coefficient of determination (R^2), mean bias (MB), standard deviation (STDEV), and root mean squared error (RMSE). MB and RMSE are calculated as WRF model—lidar. Statistical sets are grouped according to campaign average (ALL) and three primary synoptic flows. Number (N) of observations are shown in parentheses.

	YSU	MYJ	QNSE	MYNN2	ACM2	BOULAC	UW	TEMF
R^2								
ALL ($N = 286$)	0.09	0.1	0.09	0.15	0.15	0.13	0.05	0.09
Continental ($N = 125$)	0.08	0.07	0.06	0.22	0.16	0.11	0.02	0.12
Etesians ($N = 101$)	0.07	0.02	0.05	0.04	0.1	0.1	0.05	0.07
Sahara ($N = 60$)	0.05	0.13	0.08	0.22	0.13	0.13	0	0
MB								
ALL ($N = 286$)	−0.30	−0.11	0.58	−0.26	0.18	−0.12	−0.76	−0.42
Continental ($N = 125$)	−0.20	−0.10	0.72	−0.15	0.30	−0.02	−0.73	−0.40
Etesians ($N = 101$)	−0.59	−0.30	0.33	−0.59	−0.11	−0.39	−0.97	−0.60
Sahara ($N = 60$)	−0.02	0.19	0.73	0.07	0.39	0.13	−0.46	0.00
STDEV								
ALL ($N = 286$)	0.67	0.68	0.75	0.58	0.68	0.64	0.65	0.60
Continental ($N = 125$)	0.72	0.73	0.81	0.57	0.70	0.68	0.71	0.63
Etesians ($N = 101$)	0.51	0.56	0.56	0.49	0.55	0.51	0.51	0.47
Sahara ($N = 60$)	0.66	0.66	0.77	0.46	0.67	0.59	0.60	0.59
RMSE								
ALL ($N = 286$)	0.74	0.69	0.95	0.63	0.70	0.65	1.00	0.73
Continental ($N = 125$)	0.74	0.74	1.08	0.59	0.76	0.68	1.02	0.75
Etesians ($N = 101$)	0.78	0.63	0.65	0.77	0.56	0.64	1.10	0.77
Sahara ($N = 60$)	0.65	0.68	1.06	0.46	0.77	0.60	0.75	0.58

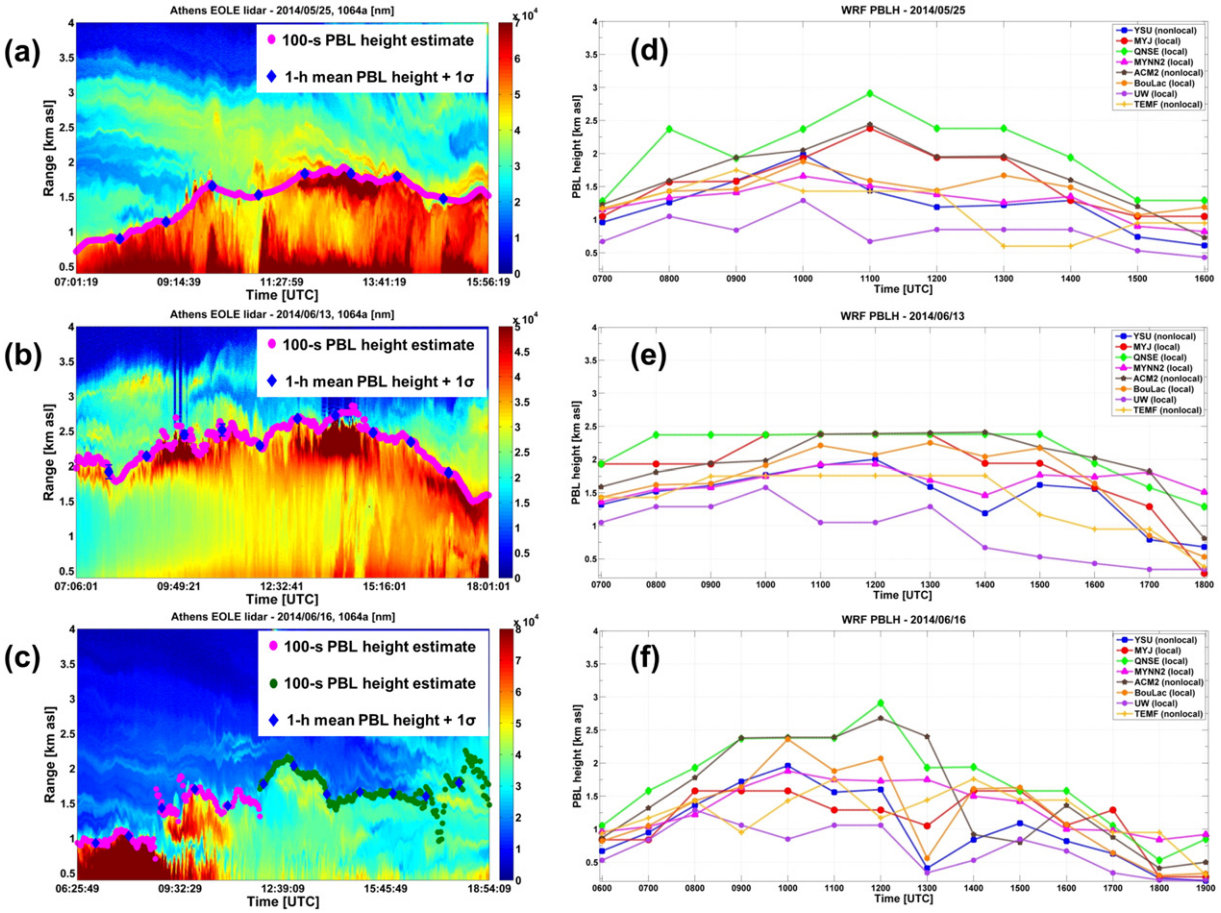

Fig. 10. Lidar range-corrected power time-range color plots (a–c) at 1064 nm wavelength and simulated PBL height from the WRF model (d–f) for (top row) Continental, (middle row) Etesians, and (bottom row) Saharan synoptic flow types. Lidar time-range plots are overlaid with PBL height estimates (circles, 100-s resolution) using the extended Kalman filter technique, in addition to the 1-h mean PBL height (blue diamonds) calculated with nine 100-s estimates, along with 1σ standard deviation.

The BouLac and MYJ schemes perform the best with this case and are similar to the average MB for this synoptic flow.

Second is the Etesians synoptic flow type (Fig. 10b), represented by lidar recorded on 13 June 2014. The lidar-EKF method estimates the highest PBLH of the three synoptic flows, with a daytime maximum near 2.5 km. However, here we show a case in which the EKF technique fails to estimate the true mixing layer. Instead the lidar estimates are closer to the residual layer. Most likely the mixing layer starts under 0.50 km in the morning, then grows to around 1.5 km after 1200 UTC. Over the GAA, the Etesians act as a powerful ventilator, advecting PBL aerosols away from the land and towards the sea. The result is a decreased aerosol load in the PBL. Here, the EKF technique has limitations with the state vector initialization early in the day, mainly due to the overlap characteristics of the instrument, with a full overlap around 0.7 km.

It is shown that WRF model-simulated PBLH (Fig. 10e) also have some complications with Etesians synoptic flow. All PBL schemes simulate the residual layer in the morning, with the UW scheme following the lowest PBL around 1 km. In contrast with the previous section, the ACM2 and MYJ schemes are not any closer to the true PBL top as the other schemes.

The third synoptic flow type (Fig. 10c) is Saharan (16 June 2014), and we observe the lowest PBLH of all the days during the campaign. Daytime maximum PBLH estimated by the lidar-EKF is around 1.75 km. In this case, we had to initialize the EKF a second time (green dots) around 1230 UTC as the aerosol signature was too low to follow the initial trajectory. The concentration of aerosols may be low, even though mixing is occurring up to 2 km. We follow the mixing layer correctly until around 1100 UTC, then the EKF method fails by jumping to the higher aerosol gradient.

Most of the PBL schemes from the WRF model appear to follow the PBLH closely with this case of Saharan synoptic flow. The WRF model results confirm the incorrect lidar estimates beginning in the afternoon hours. The TEMF scheme simulates the best PBLH values, similar to the synoptic flow average presented earlier.

In summary, the lidar-EKF technique is a useful tool for PBLH detection from lidar. However, the method has limitations in certain situations, including Etesians and Saharan synoptic flow types. WRF model-simulated PBLH during the three synoptic flows shows similar results to the averages, except in the case of Etesians flows. During this flow, both the lidar-estimated and WRF model-simulated PBLH are closer to the residual layer.

3.3. Impact of PBL schemes on vertical profiles

Daytime boundary-layer vertical profiles at EMY are compared with WRF model-simulated potential temperature (θ in K), water vapor mixing ratio (q_v in g kg^{-1}), and wind speed (WS in m s^{-1}) in Fig. 11. Vertical profiles from the WRF model are selected to the closest hour of each radiosounding launch time. Representative cases are shown for Continental (Fig. 11a–c), Etesians (Fig. 11d–f), and Saharan (Fig. 11g–i) synoptic flow types.

First, vertical profiles of Continental synoptic flow are represented by 2 June 2014 launched at 1116 UTC (Fig. 11a-c). On this day, the radiosonde-estimated PBLH is 1.82 km from the Richardson bulk method. Below 1 km, all PBL schemes over-predict θ with the TEMF scheme

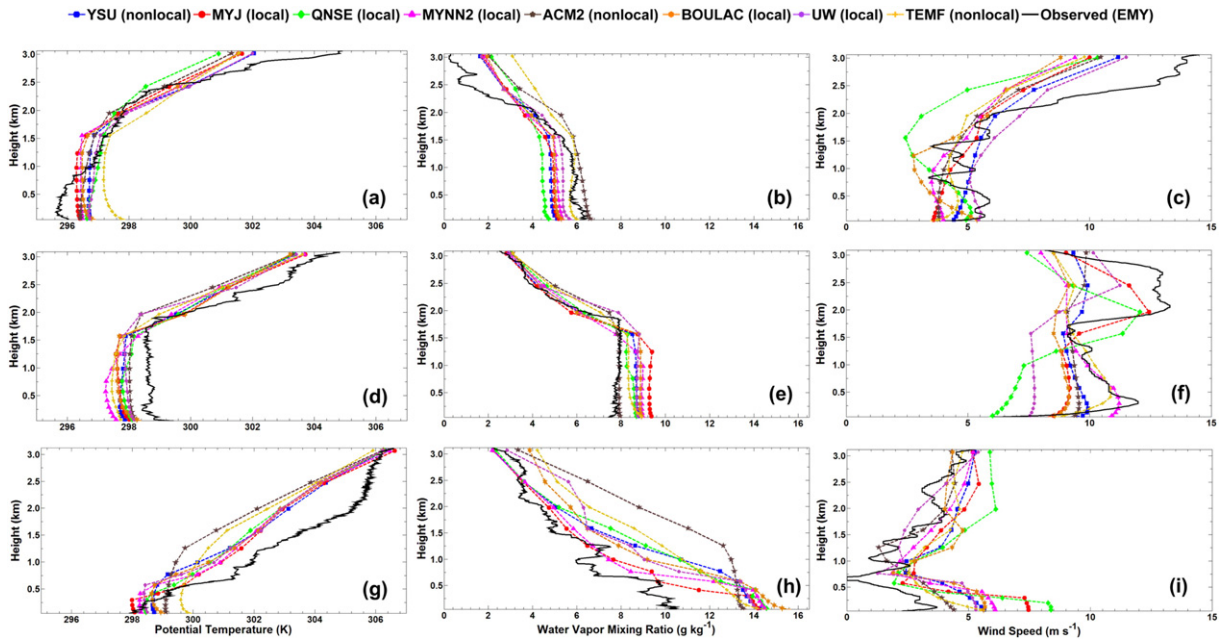

Fig. 11. Representative vertical profiles comparison of WRF model-simulated potential temperature (K), water vapor mixing ratio (g kg^{-1}), and wind speed (m s^{-1}) versus radiosonde launches at EMY (37.88 N, 23.73 E, 10 m asl). One set of profiles for (a–c) Continental (2 June 2014), (d–f) Etesians (22 May 2014), and (g–i) Saharan (16 June 2014) synoptic flow types.

showing the largest deviation (1–2 K) from the observed through the whole boundary layer. On the other hand, TEMF reproduces q_v closest to the observed values. ACM2 also performs well with only a small moist bias, less than 1 g kg^{-1} through the PBL. Other PBL schemes simulate a drier PBL ($\approx$1–2 g kg^{-1}). The model spread is rather large with the vertical profiles of WS, with no scheme able to capture the detailed structure shown with the radiosonde profile. However, the YSU and UW schemes follow the general pattern of the WS profile.

Next are comparisons of vertical profiles for the Etesians synoptic flow, most representative on 22 May 2014 with a launch time at 1120 UTC (Fig. 11d-f). For this day, we estimate that the PBLH is 2.12 km from the radiosounding. WRF model-simulated values of θ show a cold bias in the PBL ($\approx$1 K). Above the PBL, the MYJ and BouLac schemes perform best compared with the radiosounding. MYJ also reproduces q_v well, but only above the PBL. In the boundary layer ACM2 simulates q_v closest to the observed values. With respect to WS, the MYNN2 and TEMF schemes are best reproduced in the PBL, while UW and MYJ simulate accurately the WS above the PBL.

Finally, vertical soundings representing Saharan synoptic flow types are shown by a 1107 UTC launch on 16 June 2014 (Fig. 11g–i). Radiosonde-estimated PBLH is 1.19 km on this day. The largest spread among the PBL schemes is in the lowest 0.5 km. Again, the MYJ scheme simulates θ well, but still with a slight cold bias (1–2 K) compared to the sounding. Also, MYJ and MYNN2 reproduce q_v values closest to the observations above the PBL. Below 1 km, all PBL schemes simulate too moist (4–6 g kg^{-1}). WRF over-predicts the WS in the boundary layer, then the UW and ACM2 schemes follow the structure well above the PBL.

In summary, we have shown a cold, moist bias with WRF model-simulated vertical profiles of potential temperature and water vapor mixing ratio during Etesians and Saharan flows, with the closest simulated values by the MYJ scheme. However, during Continental flow, we found a slightly warm and dry bias with the WRF model, with the TEMF and ACM2 schemes showing the best results. With respect to wind speed profiles, it is more difficult to reproduce the detailed structure of the radiosoundings, with large spread among the various schemes. The closest simulated wind speed was found with the YSU scheme during Continental and Etesians flows, while the UW and ACM2 schemes work best in Saharan flows.

3.4. Sensible heat flux comparison

Comparisons of model PBL schemes are not meaningful unless the model-simulated surface heat fluxes are examined. Surface meteorological variables are very sensitive to the model surface-layer schemes, which provide surface fluxes of heat and moisture to the PBL schemes. Without a proper investigation of the surface heat fluxes, it is difficult to ascertain whether model performance is due to the impact of the PBL scheme or the surface-layer physics.

We show the diurnal cycle of surface sensible heat flux (W m^{-2}) simulated by the WRF model at the coastal station, Anavyssos, and the inland station, Peristeri (Fig. 12). We use the same representative simulation days as in Section 3.3 to evaluate the surface heat fluxes for Continental, Etesians, and Saharan synoptic flow types. Noticeably unique patterns of the diurnal cycle are observed between the coastal and inland locations. Peristeri exhibits a classic diurnal cycle in all synoptic flow types with large daytime surface heat flux values (>250 W m^{-2}), while Anavyssos shows a nearly constant cycle below 100 W m^{-2}, most likely due to its moderating location near the water. Lower daytime maximum surface heat fluxes are simulated at Peristeri during Saharan synoptic flows, which is probably a result of increased aerosols in the PBL. The increased aerosol load acts as a limiter to the amount of solar radiation, which reaches the boundary layer. A drastic drop-off in surface heat fluxes at Peristeri is noted starting around 1300 UTC in Saharan flow.

All PBL schemes simulate similar surface heat fluxes at both locations during nighttime and the early morning hours. The largest spread between the PBL schemes is seen at Peristeri during daytime in Continental synoptic flow. The PBL schemes tied to the MM5 surface-layer scheme are grouped similarly, with the lowest model-simulated values of all schemes. The three PBL schemes (QNSE, MYNN2, and TEMF) that use unique surface-layer schemes simulate the highest surface heat fluxes. It is unknown what causes the large deviation in model-simulated surface heat fluxes using the TEMF scheme at Anavyssos. In summary, it is found that WRF model-simulated surface heat fluxes are more dependent on the land surface and surface-layer schemes than PBL schemes, and should be subject to further investigation against observations.

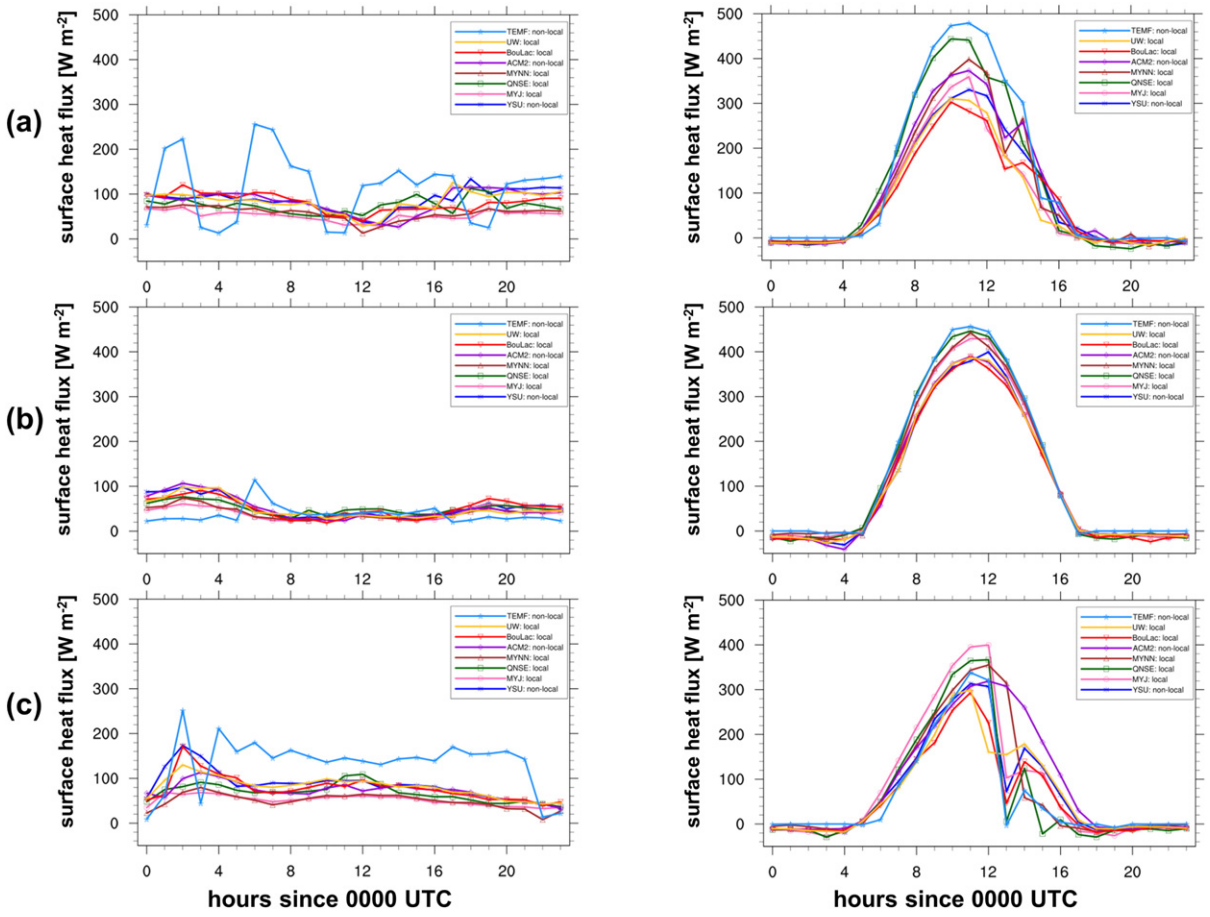

Fig. 12. Model surface sensible heat flux (W m-2) simulated with eight WRF PBL schemes at Anavyssos (column 1) and Peristeri (column 2) stations for representative synoptic flows (rows): (a) Continental (2 June 2014), (b) Etesians (22 May 2014), and (c) Saharan (16 June 2014) synoptic flow types.

3.5. Comparison to previous studies

The WRF results presented here show some similarities to previous works evaluating model PBL schemes in complex urban areas. Past evaluations have been performed with the WRF model (Kleczek et al., 2014; Banks et al., 2015) and the legacy MM5 model (Tombrou et al., 2007; Bossioli et al., 2009).

An earlier study by Bossioli et al. (2009) analyzed four PBL schemes implemented in the MM5 model over the GAA during typical summer and winter conditions. The schemes were two non-local (Blackadar and Pleim-Xiu), one semi-non-local (Medium Range Forecast; MRF), and one local scheme (Gayno-Seaman). They reported that the non-local Pleim-Xiu (PX) scheme reproduces the mean observed surface values at all stations analyzed. The ACM2 scheme, which is an improved version of PX, also is favored in our study for campaign average and Continental flow types. The enhanced turbulence of the non-local scheme is attributed to the better performance of its peers.

Tombrou et al. (2007) also used the MM5 model to evaluate simulated PBLH over the GAA. They used the same four PBL schemes as in Bossioli et al. (2009) for two different simulation days in September 1994 and 2002, in concert with the ICAROS-NET and MEDCAPHOT-TRACE field campaigns, respectively. It was found that the non-local schemes generally provide higher values of the PBLH during the daytime, which is similar to our WRF model-simulated PBLH with the ACM2 scheme.

In addition, Banks et al. (2015) investigated the performance of eight PBL schemes from WRF version 3.4.1 over the complex urban area of Barcelona, Spain. The GAA and Barcelona can be characterised with similar atmospheric conditions (complex topography, sea breeze influences, etc). WRF model-simulated PBLH were validated against estimates from a lidar as in the current study. It was determined that the ACM2 scheme most well reproduced the PBLH at 1200 UTC, with a slight under-estimate of 0.01 km. Similar to the findings of this contribution, the BouLac scheme also performed well in Barcelona with a mean error around −0.35 km.

Simulated vertical profiles of meteorological parameters show some agreements to those found in Kleczek et al. (2014) during the GABLS3 (GEWEX Atmospheric Boundary-Layer Study) campaign in Cabauw, The Netherlands. They analyzed the performance of six PBL schemes in WRF version 3.4.1 against observations from a meteorological tower and radiosondes. They found that all model simulations show a similar structure for θ and q_v, with a consistent cold ($\approx$2 K) and moist (up to 4 g kg^{-1}) bias in the upper PBL. This is similar to our results where the cold, moist bias is evident with Etesians and Saharan synoptic flows, and in turn, we find a slight warm, dry bias in the PBL during Continental flows.

However, we must be careful when comparing studies between the GAA and Cabauw. One must also take into consideration the contrasting atmospheric situations prevalent between a coastal Mediterranean site like the GAA, and a continental European site such as Cabauw, which can lead to significant differences in model performance. Model physics is sensitive to topographic differences, land-sea exchange, and latitudinal changes in the atmospheric radiation, just to name a few.

WRF model-simulated surface heat fluxes shown by Madala et al. (2015) over eastern India show similar findings to those presented in Section 3.4 with regards to model-simulated surface heat fluxes. They found that all PBL schemes tested could capture nighttime-minimum values well, but significant differences were found in the daytime maximum surface heat fluxes. We found similar differences, especially with the model-simulated surface heat fluxes at Peristeri during Continental

synoptic flows, grouped closely to the corresponding surface-layer schemes. Madala et al. (2015) determined the ACM2 and MYNN2 schemes performed better than other schemes when compared with observations from a fast response sonic anemometer paired to an eddy correlation technique.

4. Summary and conclusions

In this study, we evaluated the sensitivity of planetary boundary-layer (PBL) variables to various PBL parameterization schemes available in the Weather Research and Forecasting (WRF) numerical weather prediction model. The study used data we collected during the HygrA-CD (Hygroscopic Aerosols to Cloud Droplets) experimental campaign, which took place from mid-May to mid-June 2014 over the complex, urban terrain of the Greater Athens Area (GAA). Proper representation of the PBL from meteorological models is a necessary component in air quality forecast systems. We have shown the WRF model can be a valuable source for this information, however, is dependent on several factors.

The PBL schemes were evaluated under diverse synoptic flow types identified with 2-day backtrajectories from the FLEXPART-WRF dispersion model. Three typical atmospheric flow types were observed during the 39-day campaign: Continental, Etesians, and Saharan, which represented 41.7%, 36.1%, and 22.2% of the days, respectively.

Eight PBL schemes (5 local, 3 non-local) from WRF-ARW version 3.4.1 were tested using daily simulations on a 1 km × 1 km grid over the GAA with hourly output resolution. Near-surface observations of 2-m air temperature (T2) and relative humidity (RH2) and 10-m wind speed (WS10) were collected from surface meteorological instruments at multiple locations. Estimates of the PBL height (PBLH) are retrieved using elastic-channel (1064-nm) backscatter measurements from a multiwavelength Raman lidar using an adaptive extended Kalman filter technique. In addition, vertical profiles of atmospheric variables are obtained from radiosonde launches. The PBLH is estimated from the radiosoundings using a bulk Richardson number approach. It is found that daytime maximum PBL heights ranged from 2.57 km during Etesian flows, to as low as 0.37 km attributed with Saharan dust episodes. Lidar-estimated PBLH compared relatively well to the radiosoundings.

WRF model simulations yield drastically different solutions depending on the PBL scheme used, the meteorological parameter analyzed, and the general synoptic conditions. The largest differences between model and observations are associated with simulated values of the PBLH (>400 m on average) during Saharan dust events.

The largest spread between the lowest and highest WRF model-simulated PBLH was shown to be as high as 1.5 km. It is shown that there are influences from the underlying synoptic conditions. The local BouLac scheme reproduced PBLH well with the campaign average and Continental synoptic flows, with slight under-estimates. ACM2, a non-local scheme, is a top performer during the Etesians synoptic flow, while the TEMF scheme is best during Saharan synoptic flow.

Campaign-averaged near-surface variables showed that the WRF model tended to have a systematic cold, moist bias during daytime, most prominent at the coastal locations. The BouLac scheme reproduced T2 and RH2 well with the campaign average, and with Etesians synoptic flow. ACM2 showed the closest T2 and RH2 during Continental flows. With Saharan synoptic flows, the UW and BouLac schemes well-represented T2, while TEMF best-reproduced RH2. WRF with the YSU scheme showed the closest WS10 to the observed values with the campaign average and during Continental and Etesian synoptic flows. The BouLac scheme only slightly outperformed YSU during Saharan events.

WRF model-simulated vertical profiles of θ mostly show an across-the-board cold, moist bias, except a slightly warm and dry bias in Continental flow. The MYJ scheme simulated the closest θ and q_v during Etesians and Saharan synoptic flows. Vertical soundings of simulated WS have a difficult time reproducing the detailed structure of the radiosoundings, with large spread among PBL schemes. The YSU scheme reproduces the closest WS with Continental and Etesians flows, while the UW and ACM2 schemes work best in Saharan flows.

Future work should further address the physical explanations of the numerous differences between the WRF PBL schemes in greater detail. Use of the urban parameterization option in WRF should be explored. In addition, the study areas should be expanded to include more experimental sites and complex locations, but also areas with a more stable atmospheric regime. Additional measurements for a future study include flux measurements from a meteorological tower, more frequent upper-air soundings, and continuous nighttime backscatter measurements from a ceilometer to analyze the nocturnal boundary layer.

Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013): People, ITN Marie Curie Actions Programme (2012–2016) in the frame of ITaRS under grant agreement no. 289923. Simulations were executed on the MareNostrum supercomputer at the Barcelona Supercomputing Center, under grants SEV-2011-00067 of Severo Ochoa Program and CGL2013-46736-R awarded by the Spanish Government. Also, credit (UE FP7-INFRA-2010-1.1.16) to the ACTRIS (Aerosols, Clouds, and Trace gases Research InfraStructure) network, and (2014 SGR 522). Stavros Solomos was supported by the European Union Seventh Framework Programme (FP7-REGPOT-2012-2013-1), in the framework of the project BEYOND, under grant agreement no. 316210.

References

Angevine, W.M., Jiang, H., Mauritsen, T., 2010. Performance of an eddy diffusivity–mass flux scheme for shallow cumulus boundary layers. Mon. Weather Rev. 138, 2895–2912. http://dx.doi.org/10.1175/2010MWR3142.1.

Banks, R.F., Tiana-Alsina, J., María Baldasano, J., Rocadenbosch, F., 2014. Retrieval of boundary layer height from lidar using extended Kalman filter approach, classic methods, and backtrajectory cluster analysis. Proc. SPIE 9242, Remote Sensing of Clouds and the Atmosphere XIX; and Optics in Atmospheric Propagation and Adaptive Systems XVII, 92420F, 17 October 2014 http://dx.doi.org/10.1117/12.2072049.

Banks, R.F., Tiana-Alsina, J., Rocadenbosch, F., Baldasano, J.M., 2015. Performance evaluation of the boundary-layer height from lidar and the Weather Research and Forecasting model at an Urban Coastal Site in the North-East Iberian Peninsula. Bound.-Layer Meteorol. 157 (2), 265–292. http://dx.doi.org/10.1007/s10546-015-0056-2.

Blackadar, A.K., 1978. Modeling pollutant transfer during daytime convection. Preprints Fourth Symposium on Atmospheric Turbulence, Diffusion, and Air Quality. American Meteorological Society, Reno, pp. 443–447.

Bösenberg, J., CoAuthors, 2001. EARLINET: a European aerosol research lidar network. Laser Remote Sensing of the Atmosphere, Selected Papers of the 2001 International Laser Radar Conference, pp. 155–158.

Bossioli, E., Tombrou, M., Dandou, A., Athanasopoulou, E., Varotsos, K.V., 2009. The role of planetary boundary-layer parameterizations in the air quality of an urban area with complex topography. Bound.-Layer Meteorol. 131, 53–72. http://dx.doi.org/10.1007/s10546-009-9349-7.

Bougeault, P., Lacarrere, P., 1989. Parameterization of orography-induced turbulence in a mesobeta–scale model. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2.

Bretherton, C.S., Park, S., 2009. A new moist turbulence parameterization in the community atmosphere model. J. Clim. 22, 3422–3448. http://dx.doi.org/10.1175/2008JCLI2556.1.

Brioude, J., CoAuthors, 2013. The Lagrangian particle dispersion model FLEXPART-WRF version 3.1. Geosci. Model Dev. 6, 1889–1904. http://dx.doi.org/10.5194/gmd-6-1889-2013.

Brioude, J., Angevine, W.M., McKeen, S.A., Hsie, E.-Y., 2012. Numerical uncertainty at mesoscale in a Lagrangian model in complex terrain. Geosci. Model Dev. Discuss. http://dx.doi.org/10.5194/gmdd-5-967-2012.

Brown, R.G., Hwang, P.Y.C., 1982. Introduction to Random Signals and Applied Kalman Filtering. first ed. Wiley, New York.

Chen, S.-H., Dudhia, J., 2000. Annual Report: WRF Physics. Air Force Weather Agency (38 pp).

Coniglio, M.C., Correia, J., Marsh, P.T., Kong, F., 2013. Verification of convection-allowing WRF model forecasts of the planetary boundary layer using sounding observations. Weather Forecast. 28, 842–862. http://dx.doi.org/10.1175/WAF-D-12-00103.1.

Draxl, C., Hahmann, A.N., Peña, A., Giebel, G., 2014. Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. Wind Energy 17, 39–55. http://dx.doi.org/10.1002/we.1555.

Dudhia, J., 1989. Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. J. Atmos. Sci. http://dx.doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Dudhia, J., 1993. A non-hydrostatic version of the Penn State-NCAR mesoscale model: validation tests and simulation of an Atlantic cyclone and cold front. Mon. Weather Rev. 121, 1493–1513.

Gent, P.R., Danabasoglu, G., Donner, L.J., Holland, M.M., Hunke, E.C., Jayne, S.R., Lawrence, D.M., Neale, R.B., Rasch, P.J., Vertenstein, M., Worley, P.H., Yang, Z.L., Zhang, M., 2011. The community climate system model version 4. J. Clim. 24, 4973–4991. http://dx.doi.org/10.1175/2011JCLI4083.1.

Holtslag, A.A.M., De Bruijn, E.I.F., Pan, H.-L., 1990. A high resolution air mass transformation model for short-range weather forecasting. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1990)118<1561:AHRAMT>2.0.CO;2.

Hong, S.-Y., 2010. A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. Q. J. R. Meteorol. Soc. 136, 1481–1496. http://dx.doi.org/10.1002/qj.665.

Hong, S.-Y., Pan, H.-L., 1996. Nonlocal boundary layer vertical diffusion in a medium-range forecast model. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1996)124<2322:NBLVDI>2.0.CO;2.

Hong, S.-Y., Dudhia, J., Chen, S.-H., 2004. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

Hong, S.-Y., Noh, Y., Dudhia, J., 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Weather Rev. http://dx.doi.org/10.1175/MWR3199.1.

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. http://dx.doi.org/10.1175/2010JAMC2432.1.

Janjić, Z.I., 1990. The step-mountain coordinate: physical package. Mon. Weather Rev. http://dx.doi.org/10.1175/1520-0493(1990)118<1429:TSMCPP>2.0.CO;2.

Janjić, Z.I., 2002. Nonsingular implementation of the Mellor–Yamada Level 2.5 scheme in the NCEP meso model. NCEP Office Note 437 (61 pp).

Jiménez, P.a., Dudhia, J., González-Rouco, J.F., Montávez, J.P., García-Bustamante, E., Navarro, J., Vilà-Guerau De Arellano, J., Muñoz-Roldán, a., 2013. An evaluation of WRF's ability to reproduce the surface wind over complex terrain based on typical circulation patterns. J. Geophys. Res. Atmos. 118, 7651–7669. http://dx.doi.org/10.1002/jgrd.50585.

Kain, J.S., 2004. The Kain–Fritsch convective parameterization: an update. J. Appl. Meteorol. http://dx.doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kim, Y., Sartelet, K., Raut, J.C., Chazette, P., 2013. Evaluation of the weather research and forecast/urban model over greater Paris. Bound.-Layer Meteorol. 149, 105–132. http://dx.doi.org/10.1007/s10546-013-9838-6.

Kleczek, M.A., Steeneveld, G.J., Holtslag, A.A.M., 2014. Evaluation of the Weather Research and Forecasting mesoscale model for GABLS3: impact of boundary-layer schemes, Boundary Conditions and Spin-Up. Bound.-Layer Meteorol. 152, 213–243. http://dx.doi.org/10.1007/s10546-014-9925-3.

Kokkalis, P., Papayannis, A., Mamouri, R.E., Tsaknakis, G., Amiridis, V., 2012. The EOLE lidar system of the National Technical University of Athens. 26th International Laser Radar Conference (26th ILRC). Porto Heli, pp. 629–632.

Lange, D., Tiana-Alsina, J., Saeed, U., Tomás, S., Rocadenbosch, F., 2014. Atmospheric boundary layer height monitoring using a kalman filter and backscatter lidar returns. IEEE Trans. Geosci. Remote Sens. 52, 4717–4728. http://dx.doi.org/10.1109/TGRS.2013.2284110.

Lange, D., Rocadenbosch, F., Tiana-alsina, J., Frasier, S., 2015. Atmospheric-boundary-layer height estimation using a Kalman filter and a frequency-modulated continuous-wave radar. IEEE Trans. Geosci. Remote Sens. 53, 3338–3349. http://dx.doi.org/10.1109/TGRS.2014.2374233.

LeMone, M.A., Tewari, M., Chen, F., Dudhia, J., 2012. Objectively-determined fair-weather CBL depths in the ARW-WRF model and their comparison to CASES-97 observations. Mon. Weather Rev. 141, 30–54. http://dx.doi.org/10.1175/MWR-D-12-00106.1.

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. http://dx.doi.org/10.1016/j.atmosenv.2015.02.059.

Martilli, A., Clappier, A., Rotach, M.W., 2002. An urban surface exchange parameterisation for mesoscale models. Bound.-Layer Meteorol. 104, 261–304. http://dx.doi.org/10.1023/A:1016099921195.

Melas, D., Ziomas, I.C., Zerefos, C.S., 1995. Boundary layer dynamics in an urban coastal environment under sea breeze conditions. Atmos. Environ. 29, 3605–3617.

Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res. http://dx.doi.org/10.1029/97JD00237.

Nakanishi, M., Niino, H., 2006. An improved Mellor–Yamada level-3 model: its numerical stability and application to a regional prediction of advection fog. Bound.-Layer Meteorol. 119, 397–407. http://dx.doi.org/10.1007/s10546-005-9030-8.

Papayannis, A., Mamouri, R.E., Amiridis, V., Kazadzis, S., Pérez, C., Tsaknakis, G., Kokkalis, P., Baldasano, J.M., 2009. Systematic lidar observations of Saharan dust layers over Athens, Greece in the frame of EARLINET project (2004–2006). Ann. Geophys. 27, 3611–3620. http://dx.doi.org/10.5194/angeo-27-3611-2009.

Papayannis, A., Argyrouli, A., Bougiatioti, A., Remoundaki, E., Eleftheriadis, E., Nenes, A., Van de Hey, J., Komppula, M., Solomos, S., Kazadzis, S., Banks, R.F., Labzovskii, L., Kalogiros, I., Tzanis, C.G., 2015. From hygroscopic aerosols to cloud droplets: the HYGRA-CD campaign in the Athens Basin—an outline. Sci. Total Environ. (Under prep.).

Pérez, C., Jiménez, P., Jorba, O., Sicard, M., Baldasano, J.M., 2006a. Influence of the PBL scheme on high-resolution photochemical simulations in an urban coastal area over the Western Mediterranean. Atmos. Environ. 40, 5274–5297. http://dx.doi.org/10.1016/j.atmosenv.2006.04.039.

Pérez, C., Nickovic, S., Pejanovic, G., Baldasano, J.M., Özsoy, E., 2006b. Interactive dust-radiation modeling: a step to improve weather forecasts. J. Geophys. Res. Atmos. 111. http://dx.doi.org/10.1029/2005JD006717.

Pleim, J.E., 2007. 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. http://dx.doi.org/10.1175/JAM2539.1.

Richardson, H., Basu, S., Holtslag, A.A.M., 2013. Improving stable boundary-layer height estimation using a stability-dependent critical bulk Richardson Number. Bound.-Layer Meteorol. 148, 93–109. http://dx.doi.org/10.1007/s10546-013-9812-3.

Rocadenbosch, F., Vázquez, G., Comerón, A., 1998. Adaptive filter solution for processing lidar returns: optical parameter estimation. Appl. Opt. 37, 7019–7034. http://dx.doi.org/10.1364/AO.37.007019.

Rocadenbosch, F., Soriano, C., Comerón, A., Baldasano, J.M., 1999. Lidar inversion of atmospheric backscatter and extinction-to-backscatter ratios by use of a Kalman filter. Appl. Opt. 38, 3175–3189.

Skamarock, W.C., Klemp, J.B., 2008. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. J. Comput. Phys. 227, 3465–3485. http://dx.doi.org/10.1016/j.jcp.2007.01.037.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., Powers, J.G., 2005. A description of the Advanced Research WRF Version 2. NCAR Tech. Note NCAR/TN-468 + STR (doi:10.1.1.127.5949).

Sukoriansky, S., Galperin, B., Perov, V., 2005. Application of a new spectral theory of stably stratified turbulence to the atmospheric boundary layer over sea ice. Bound.-Layer Meteorol. 117, 231–257. http://dx.doi.org/10.1007/s10546-004-6848-4.

Tewari, M., CoAuthors, 2004. Implementation and verification of the unified NOAH land surface model in the WRF model. 20th Conference on Weather Analysis and Forecasting/16th Conference on Numerical Weather Prediction, pp. 11–15.

Tombrou, M., Dandou, A., Helmis, C., Akylas, E., Angelopoulos, G., Flocas, H., Assimakopoulos, V., Soulakellis, N., 2007. Model evaluation of the atmospheric boundary layer and mixed-layer evolution. Bound.-Layer Meteorol. 124, 61–79. http://dx.doi.org/10.1007/s10546-006-9146-5.

Troen, I.B., Mahrt, L., 1986. A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. Bound.-Layer Meteorol. 37, 129–148. http://dx.doi.org/10.1007/BF00122760.

Tsaknakis, G., Papayannis, A., Kokkalis, P., Amiridis, V., Kambezidis, H.D., Mamouri, R.E., Georgoussis, G., Avdikos, G., 2011. Inter-comparison of lidar and ceilometer retrievals for aerosol and planetary boundary layer profiling over Athens, Greece. Atmos. Meas. Tech. 4, 1261–1273. http://dx.doi.org/10.5194/amt-4-1261-2011.

Wilks, D.S., Department of E. and A.S.C.U., 2006a. Statistical Methods in the Atmospheric Sciences, Meteorological Applications.

Wilks, D.S., Department of E. and A.S.C.U., 2006b. Statistical Methods in the Atmospheric Sciences, Meteorological Applications.