Geosci. Model Dev. Discuss., 7, 6133–6171, 2014
www.geosci-model-dev-discuss.net/7/6133/2014/
doi:10.5194/gmdd-7-6133-2014

Geoscientific Model Development Discussions

Sensitivity analysis of PBL schemes by comparing WRF model and experimental data

A. Balzarini[1], F. Angelini[2], L. Ferrero[3], M. Moscatelli[3], M. G. Perrone[3], G. Pirovano[1], G. M. Riva[1], G. Sangiorgi[3], A. M. Toppetti[1], G. P. Gobbi[4], and E. Bolzacchini[3]

[1]Ricerca sul Sistema Energetico (RSE SpA), Via Rubattino 54, 20134 Milano, Italy
[2]ENEA/UTAPRAD-DIM, Via Enrico Fermi 45, 00044 Frascati, Italy
[3]POLARIS research center, Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milano, Italy
[4]CNR/ISAC, Via del Fosso del Cavaliere 100, 00133 Roma, Italy

Received: 28 April 2014 – Accepted: 14 May 2014 – Published: 17 September 2014

Correspondence to: A. Balzarini (alessandra.balzarini@rse-web.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This work discusses the sources of model biases in reconstructing the Planetary Boundary Layer (PBL) height among five commonly used PBL parameterizations. The Weather Research and Forecasting (WRF) Model was applied over the critical area of Northern Italy with 5 km of horizontal resolution, and compared against a wide set of experimental data for February 2008. Three non-local closure PBL schemes (Asymmetrical Convective Model version 2, ACM2; Medium Range Forecast, MRF; Yonsei University, YSU) and two local closure parameterizations (Mellor Yamada Janjic, MYJ; University of Washington Moist Turbulence, UW) were selected for the analysis. Vertical profiles of aerosol number concentrations and Lidar backscatter profiles were collected in the metropolitan area of Milan in order to derive the PBL hourly evolution. Moreover, radio-soundings of Milano Linate airport as well as surface temperature, mixing ratio and wind speed of several meteorological stations were considered too.

Results show that all five parameterizations produce similar performances in terms of temperature, mixing ratio and wind speed in the city of Milan, implying some systematic errors in all simulations. However, UW and ACM2 use the same local closure during nighttime conditions, allowing smaller mean biases (MB) of temperature (ACM2 MB = 0.606 K, UW MB = 0.209 K), and wind speed (ACM2 MB = 0.699 m s^{-1}, UW MB = 0.918 m s^{-1}).

All schemes have the same variations of the diurnal PBL height, since over predictions of temperature and wind speed are found to cause a general overestimation of mixing during its development in winter. In particular, temperature estimates seem to impact the early evolution of the PBL height, while entrainment fluxes parameterizations have major influence on the afternoon development. MRF, MYJ and ACM2 use the same approach in reconstructing the entrainment process, producing the largest overestimations of PBL height (MB ranges from 85.51–179.10 m). On the contrary, the best agreement between model and both Lidar and balloon observations was identified for YSU (MB = −27.54 m and 30.15 m, respectively).

1 Introduction

The knowledge and study of air quality are important because atmospheric pollutants can induce adverse effects on human health as well as natural ecosystems (e.g., Utell, 2006; Krupa et al., 2006).

Air quality is normally investigated at local or regional scale by the aid of different types of Chemistry and Transport Models (CTMs), which allow reproducing the fate of the main atmospheric pollutants, both primary and secondary, such as ozone, nitrogen oxides and particulate matter (PM). CTMs are generally driven by 3-D meteorological fields provided by a previous run of a mesoscale meteorological model. Consequently, the correct representation of air quality is strongly affected by the simulation of meteorological processes and parameters. The main controlling variables are wind, temperature, turbulent fluxes and, among others, the height of the Atmospheric Boundary Layer (ABL), also called Planetary Boundary Layer (PBL).

In recent years, there has been significant progress on the characterization of atmospheric turbulence, but the determination of the PBL remains one of the most uncertain parameters in model estimates, affecting the reconstruction of dispersion processes and, then, ground concentrations (Misenis and Zhang, 2010; Yerramilli et al., 2010).

The Weather Research and Forecasting (WRF; Skamarock, 2005) model is a state-of-art in meteorological applications that offers several schemes to reconstruct PBL heights, each adopting different assumptions when describing the turbulence or eddy activities in stable, neutral or convective conditions. Furthermore, new PBL schemes have been recently embedded into WRF version 3.3, such as the University of Washington Moist Turbulence (UW) scheme (Bretherton and Park, 2009).

Several studies have explored the sensitivity of PBL schemes in WRF model (Misenis et al., 2006; Zhong et al., 2007; Borge et al., 2008). All these works showed discrepancies between simulations with several PBL schemes and among these and experimental measures. However, all the aforementioned studies did not investigate the relationship between model performances and differences in PBL parameterizations.

More recently, this limit was overcome in the work of Ferrero et al. (2011a) where a sensitivity analysis of four PBL schemes over the North of Italy was conducted comparing results of the Fifth-Generation Penn State/NCAR Mesoscale Meteorological model (MM5; Grell et al., 1994) with particle concentration vertical profiles obtained by balloon soundings. However, due to the recently replacement of MM5 by WRF as mesoscale meteorological model for CTMs, the accuracy of WRF schemes in predicting the PBL has yet to be assessed.

Hu et al. (2010) attempted to evaluate the causes of model biases for three PBL schemes of WRF model in south-central United States. In contrast to previous sensitivity studies, the present application focuses on analyzing the differences in five model schemes allowing conclusions on the influence of PBL formulations on the modeled results. To this aim, WRF (version 3.3.1) has been applied to the critical area of the Po Valley (North of Italy) and compared to experimental data in order to evaluate five PBL parameterizations in winter: Medium Range Forecast (MRF) scheme (Hong and Pan, 1996), Yonsei University (YSU) scheme (Hong et al., 2006), and Mellor Yamada Janjic (MYJ) formulation (Janjic, 1994), as well as the Asymmetrical Convective Model version 2 (ACM2; Pleim, 2007) scheme and the new UW parameterization (Bretherton and Park, 2009).

The Po Valley is a European hot spot for atmospheric pollution due to the fact that the ventilation is generally poor and the atmospheric conditions are often stagnant, especially during winter when frequent thermal inversion at low altitude and prolonged foggy periods induce very low PBL depths (Rodriguez et al., 2007; Carbone et al., 2010; Ferrero et al., 2012). These conditions often limit model performances in estimating the PBL height (Ferrero et al., 2011a). This makes the Po Valley area, in winter, an interesting case study for modeling applications.

In this study many experimental techniques (ground level measurements, particle vertical profiles by balloon soundings, meteorological balloons and Lidar measurements) were combined together in order to obtain a correct and thorough representation of the PBL structure. Together with Ferrero et al. (2011a), this work represents

one of the few investigations of PBL heights over the Po Valley throughout a comparison between model results and observations.

The following section (Sect. 2) describes the WRF modeling set up, the main features of experimental techniques and the comparison method. In Sect. 3, a detailed analysis of results is presented. Finally, Sect. 4 summarizes the main findings and conclusions.

2 Model and observations

2.1 WRF description and modeling set up

In this study, WRF model version 3.3.1 has been applied. WRF is a non-hydrostatic meteorological model designed to simulate mesoscale and regional scale atmospheric circulation. It includes many physical and dynamical options for microphysics, radiation, cumulus processes and PBL (www.wrf-model.org).

The main physical parameterizations adopted here include the Rapid Radiative Transfer Model (RRTM) long wave radiation scheme (Mlawer et al., 1997) and the Goddard shortwave radiation scheme (Chou et al., 1998), the Noah land surface model (Chen and Dudhia, 2001), the Morrison double moment microphysics scheme (Morrison et al., 2009) and the Grell 3-D ensemble cumulus parameterization (Grell and Devenyi, 2002).

Five PBL schemes have been selected for the sensitivity test: MRF (Medium Range Forecast; Hong and Pan, 1996), YSU (Yonsei University; Hong et al., 2006), MYJ (Mellor Yamada Janjic; Janjic, 1994), UW (University of Washington Moist Turbulence; Bretherton and Park, 2009) and ACM2 (Asymmetrical Convective Model version 2; Pleim, 2007). The differences among the five PBL schemes are related to the turbulence or eddy diffusivity assumption, the parameterization of the PBL top and the treatment of the entrainment zone in stable, neutral and convective conditions.

The MRF scheme adopts a nonlocal-K approach proposed by Troen and Mahrt (1986) to simulate the mixed-layer diffusion with an implicitly parameterization of the

entrainment processes. The Boundary Layer height is enhanced by comparing the computed bulk Richardson number with a critical value (0.5; Hong et al., 2006). The entrainment effects are merely reproduced by additional mixing above the minimum flux level (Hong et al., 2006). The YSU scheme is a modification of the MRF approach. The major changes include the explicit treatment of the entrainment processes at the inversion layer by means of an asymptotic entrainment flux term added to the turbulence diffusion equation and a critical bulk Richardson number sets to zero (Hong et al., 2006). The YSU scheme is further modified in WRF version 3 by increasing the critical bulk Richardson number from zero to 0.25 over land during stable boundary conditions (Hong and Kim, 2008).

MYJ is an implementation of the Mellor Yamada level 2.5 model (Mellor and Yamada, 1982). It applies a local approach to determine eddy diffusion coefficients from prognostic turbulent kinetic energy (TKE). Since the TKE is largest within the PBL, MYJ defines its top as the height where the TKE becomes negative or drops to a prescribed lower bound (Hu et al., 2010; Janjic, 2001). A similar approach is used in the recently added UW parameterization, but turbulent kinetic energy is diagnosed rather than prognosed for different regimes (stable or convective) and an explicit entrainment closure is used at the edge of the convective layers (Bretherton and Park, 2009).

Finally, the ACM2 combines the nonlocal transport of the ACM1 model (Pleim and Chang, 1992) and the local eddy diffusivity. In this way, vertical fluxes are described as pure eddy diffusion in stable conditions and a combination of local gradient and nonlocal turbulent transport in unstable conditions (Pleim, 2007). In addition, ACM2 can simulate a convective upward transport from the lowest level to all other model layers and an asymmetrical layer-by-layer downward transport (Pleim, 2007). The height of the Boundary Layer is calculated starting from the bulk Richardson number. This scheme does not include and explicit treatment of the entrainment processes.

Because in WRF exists a particular surface layer scheme to which each PBL formulation is preferentially coupled, the surface layer schemes are also varied. The MRF, YSU and ACM2 are coupled to MM5 scheme, while MYJ and UW are associated to

Eta surface layer. Both surface schemes are based on similarity theory (Monin and Obukhov, 1954), but the second one includes a parameterization of a viscous sublayer following Janjic (1994).

The WRF model has been applied over three domains in a Lambert Conic Conformal projection for February 2008. The master domain covers the whole Europe with a horizontal resolution of 45 km. The first nested domain extends over the Italian Peninsula with a grid step of 15 km and the second one is centered over the Po Valley area with a spatial resolution of 5 km. The cumulus scheme is not applied in the highest resolution domain. All model domains have 27 vertical layers, and the model top is set to 50 hPa.

A grid nudging on wind speed, temperature and water vapor mixing ratio has been employed within the Boundary Layer in all model configurations. The nudging coefficients were set to $0.0003\,s^{-1}$.

Input terrestrial data were derived from the 24-land use category of US Geological Survey database (USGS; http://www.usgs.gov/). The ECMWF (European Centre for Medium-Range Weather Forecast; http://www.ecmwf.int/) analysis of 6 h and 0.5° grid resolution were used as initial and boundary conditions.

The model ran for February 2008 with a spin up time of five days.

2.2 Experimental vertical profiles by tethered balloon

Direct measurements of PBL height can be made through the use of tethered balloons (Laakso et al., 2007; Wiegner et al., 2006; McKendry et al., 2004; Maletto et al., 2003). To this aim, vertical aerosol profiles were carried out in the Milan metropolitan area (Torre Sarca; 45°31′19″ N, 9°12′46″ E), in the midst of the most industrialized area in the Po Valley, using a spherical helium-filled tethered balloon (PU balloon, = 4 m, volume 33.5 m^3, payload 15 kg). The balloon carried aloft a sampling platform consisting of an optical particle counter (OPC, 1.108 "Dustcheck" GRIMM, 15 particle class-sizes ranging from 0.3–20 µm) and a portable meteorological station (BABUC-ABC, LSI-Lastem: pressure, temperature and relative humidity); both instruments acquired

data at 6 s of time resolution. Ascent and descent rates were controlled by an electric winch; a fixed value of $30.0 \pm 0.1\,m\,min^{-1}$ was used for all profiles, giving 3.0 m of measuring vertical resolution. The maximum height reached during each launch depended on atmospheric conditions; for the majority of the profiles it was 600 m a.g.l.

Further details of the experimental approach can be found in Ferrero et al. (2010; 2012), while the site and the related aerosol properties are reported in Ferrero et al. (2014), in Perrone et al. (2013) and in Sangiorgi et al. (2011).

Basing on the observation that atmospheric particles act as tracers of atmospheric plumes, and integrate the effects of both thermal and mechanical forces, the Boundary Layer height was derived, in the present work, using a gradient method applied to the particle vertical profiles as also described in detail in Ferrero et al. (2010). At the same time the instantaneous thickness of the Entrainment Zone (EZ), which connects the layer of vertical mixing (mixing zone) with the layer above, was calculated too (Stull, 1998). Particularly, the EZ thickness is the layer around the mixing zone, measured using the gradient method, extending between the regions in which more than 5 %, but less than 95 %, of the air in the vertical profile possess above PBL characteristics.

A comparison of the aforementioned procedure with potential temperature, relative humidity, black carbon profiles and Lidar (laser radar) data was yet performed and it is reported in literature (Ferrero et al., 2010; 2011a, b; Angelini et al., 2009).

Vertical aerosol profile measurements were performed along three years from 2006–2008. The complete dataset of vertical particle profiles is described in Ferrero et al. (2010). In this work, only sixteen ascent and descent profiles were considered in the analysis. Measurements were generally performed in the morning, from sunrise to noon, during the field campaign of February 2008.

For a complete evaluation of modeled data, radio soundings of Milano Linate airport (about 9 km SE far from Torre Sarca; 45°26′ N, 9°17′ E) were considered too. Milano Linate station provides high resolution vertical profiles of pressure, temperature and dew point temperature, relative humidity, wind speed and wind direction every 12 h.

2.3 Lidar measurements

Information on PBL height can be also obtained through experimental techniques such as Lidar (Light Detection and Ranging) by observing from ground the vertical evolution of particle backscatter along the day (Amiridis et al., 2007; Kim et al., 2007; Cohn and Angevine, 2000). Planetary Boundary Layer height was thus estimated by means of an automated Lidar Ceilometer (Vaisala LD-40) installed at the Torre Sarca site on the same days as the balloon launching.

The Ceilometer acquired aerosols backscatter profiles every 15 s at the wavelength of 855 nm from January 2007 to February 2008. Measurements of February 2008 were only included in this study. Data were averaged every 15 min to increase the signal-to-noise ratio and corrected for the presence of clouds and haze.

Elastic backscatter Lidars can provide remote-sensing information on the aerosol distribution within the PBL, in terms of an aerosol cross section profile. Three methods have been proposed so far to infer the PBL height from Lidar data (Summa et al., 2013): the threshold method (Melfi, 1985), the gradient method (Endlich, 1967), and the variance method (Hooper and Eloranta, 1986; Hennemuth and Lammert, 2006). At the basis of all these methods there is the assumption that aerosols are produced at the ground and mixed in the PBL by the effect of turbulence-induced by convection. In this work, PBL heights were obtained using a gradient method (Angelini et al., 2009), which employs the inflection points in the corrected aerosol backscatter profile to determine the height of the PBL. For a detailed description of the PBL retrieval by elastic Lidar, see, among others: Angelini et al. (2009), Morille et al. (2007), Martucci et al. (2007) and Haeffelin et al. (2012). The latter work shows that while the algorithms for the determination of aerosol layer heights are nowadays rather efficient, the most difficult task for automated procedures is the attribution of the top of the boundary layer to one of the detected layers, since in many cases residual or advected aerosol layers may induce errors in the attribution. Low aerosol loads also represent a condition for uncertain determination of PBL height.

For this reason, in the present work a supervised analysis has been performed, and the boundary layer top has been attributed by visual inspection.

In this respect, particular attention must be paid to the afternoon transition between the convective boundary layer and the new stable boundary layer building up after sunset. At that time, in fact, the PBL height shows large ambiguities, depending on which criterion is adopted to measure it. Indeed, when the solar forcing is lowering, the convective atmosphere becomes first neutral and then stable, from a thermo-dynamical point of view. In such conditions, part of the aerosol lifted up by daytime turbulence will remain aloft and undergo slow sedimentation. In such conditions, Lidar-based estimates of the Boundary Layer height will keep giving the highest levels observed in the middle of the day, until new aerosol emitted at ground and confined now in a much shallower layer becomes “visible” to the Lidar, leading to the detection of the new stable layer.

Usually this process takes place along a rather sharp transition, and smooth drops of the PBL are rarely observed.

2.4 Methods

Simulation results were compared to experimental data through the Atmospheric Model Evaluation Tool (AMET 1.1, Appel et al., 2011). Evaluations focus on PBL heights as well as the main meteorological parameters, namely temperature, mixing ratio and wind speed.

Experimental PBL height at Torre Sarca was assessed using particle vertical profiles obtained by balloon soundings and Lidar data, while other meteorological fields were evaluated at 63 WMO (World Meteorological Organization) ground measurement stations located in the Po Valley. Meteorological radio-soundings at Milano Linate airport were considered too.

Particle vertical profiles and Lidar data enable to experimentally estimate the magnitude of PBL depth and then the accuracy of different PBL schemes in reconstructing its

structure and evolution. Furthermore, Lidar data allow obtaining a continuous temporal coverage of measurements.

Finally, meteorological stations and radio-soundings enable to elucidate model capability at reproducing the main meteorological parameters.

Comparisons have been done in terms of temporal variation, vertical profiles, and performance statistics. Several metrics could be included in the analysis (Lin et al., 2008; Zhang et al., 2006). In order to obtain a complete characterization, five statistical parameters were selected (Appendix A): Mean Observed, Mean Modeled, Mean Bias (MB), Root Mean Square Error (RMSE), and Pearson Correlation (r). MB and RMSE enable to consider the accuracy of model schemes in reconstructing the PBL magnitude, while r to characterize the hourly or seasonal evolution.

Modeled and observed data have been corrected removing outliers. Temperature data are compared when the difference between model and observation is lower than 20 K, while for mixing ratio a threshold of $10\,g\,kg^{-1}$ is set. Wind speeds lower than $0.5\,m\,s^{-1}$ and higher than $100\,m\,s^{-1}$ are rejected.

The aim of this study is to provide a sensitivity test of PBL scheme in the peculiar area of the Po Valley. For this reason, evaluations consider only the 5 km domain, interpolating model results to measurement sites and hours.

3 Results and discussion

3.1 Temperature, mixing ratio and wind speed

Table 1 displays the performance statistic of 2 m-temperature, 2 m-mixing ratio and 10 m wind speed at 63 WMO meteorological stations.

All simulations produce higher temperatures than the observed values in the lower atmosphere, implying some systematic errors in all model runs. The lowest bias is always related to ACM2 scheme (MB = 0.601 K), while the highest temperature biases are associated to MYJ and UW that use the Janjic Eta Monin–Obukhov surface layer

scheme. Indeed, one possible cause of the biases can be related to the heat fluxes delivered by this scheme with respect to Monin-Obukhov surface layer scheme used with YSU, ACM2 and MRF (Hu et al., 2010). In order to verify it, the mean upward sensible heat fluxes at the surface (HFX) are represented in Fig. 1.

The figure shows that the highest mean heat sensible fluxes are associated to MYJ and UW. In particular UW predicts the largest temperature bias (MB = 0.911 K) as well as the most pronounced differences of HFX over the Po Valley area. However when non-local closure schemes are compared together, it is possible to state that MRF predicts higher temperature than ACM2 and YSU, but lower heat fluxes. Downward flux at the ground surface of shortwave incoming solar radiation (SWDOWN) is represented in Fig. 2. All schemes have similar spatial distribution of the incoming solar radiation, even though some discrepancies are detectable on the North-East and South-West regions of the domain because of the differences in the cloud cover (not shown). UW is found to produce an incoming radiation that is on average larger than all other schemes. Indeed, this parameterization has the smallest cloud cover fraction. As discussed later, UW has also the weakest vertical mixing and, thus, low planetary boundary layer heights that tend to trap heat and moisture close to ground reducing the cloud formation at higher atmospheric levels. In fact, UW has also the highest latent heat flux at the surface (not shown).

ACM2 and YSU have a similar mean incoming shortwave radiation which is generally smaller than MRF and MYJ in the North-East of Italy, since the latest schemes have less cloud cover fraction. Non-local closure schemes have quite comparable latent heat fluxes as they use the same surface layer scheme, although in MRF it is slightly higher. Collectively, the comparison shows that discrepancies in temperature performances can be partially explained also with differences in SWDOWN field and latent heat fluxes.

Non-local closure schemes (MRF, YSU and ACM2) have also the best performance in terms of mixing ratio. The WRF model predicts a mixing ratio mean bias of −0.026, 0.041, 0.122, 0.265, and $0.278\,g\,kg^{-1}$ with the MRF, ACM2, YSU, UW and MYJ scheme, respectively. Moisture is generally overestimated by all configurations,

consistently with temperature over prediction. The only exception is the MRF scheme that slightly under predicts mixing ratio. The overall trend is well reproduced by models, showing a correlation of about 0.8.

As far as wind speed is concerned, performance analysis reveals a model difficulty at reconstructing this meteorological field in terms of both magnitude and time variation. Biases of MRF, YSU and ACM2 are lower than MYJ and UW, and they range from 0.354 m s^{-1} (ACM2) to 0.776 m s^{-1} (MYJ).

The goal of this work is to assess the ability of different parameterization schemes to predict the PBL behavior over the critical area of the Po Valley. However, as experimental data of PBL height were available over Milan, the accuracy of simulated temperature, wind speed and mixing ratio was also assessed for the urban city.

The performance indices at Milano Linate station are reported in Table 2. Runs follow the previous general pattern, over predicting temperature and wind speed, but an underestimation of mixing ratio is observed. However, the five schemes show different behaviors in the city of Milan.

In order to better analyze those differences, Fig. 3 shows the bias diurnal variation of 2 m temperature, 2 m mixing ratio and 10 m wind speed at Milano Linate station.

MRF and YSU produce the highest temperature overestimations (MRF MB = 0.924 K and YSU MB = 0.981 K), but all schemes show the same daily trend of the mean bias. In the morning, temperatures predicted with UW have lower bias than the other schemes, while during the day all schemes show positive biases. Discrepancies among the five schemes are mainly related to nighttime hours, especially when YSU parameterization is considered. As reported by Hu et al. (2010), the enhanced stable nighttime vertical mixing included in the YSU scheme (Hong and Kim, 2008) contributes to stronger downward fluxes that lead to higher temperature and lower moisture near the ground.

Moreover, under nighttime stable conditions, non-local transport of the ACM2 scheme is shut down and the vertical mixing is merely caused to eddy diffusivity as in MYJ (Hu et al., 2010). As a consequence, ACM2 and MYJ have similar magnitudes

in reconstructing temperature and mixing ratio during the night. On the contrary, during daytime ACM2 shows a temperature variation more similar to non-local closure schemes.

Concerning mixing ratio, all runs show good agreement with measurements in the early morning, while a negative bias is observed during the afternoon and the night. The maximum mixing ratio error varies from run to run, with very little positive bias around 6:00 UTC for YSU and MYJ. A more substantial negative bias is highlighted form 9:00–18:00 UTC. As discussed later in this work, this is consistent with the bias trend observed in PBL height and it can be partially explained by an enhancement in vertical mixing during these hours. Indeed, the temperature overestimation can increase the thermally induced convection inside the PBL, decreasing the mixing ratio at the ground. As a consequence, in Milan, all runs produce mixing ratio that are lower than observed. The higher biases are associated to MRF and ACM2, with values of -0.276, -0.208 g kg^{-1}, respectively. YSU has a similar behavior, but better overall performances, showing a bias of -0.131 g kg^{-1}.

During unstable daytime conditions, MRF, ACM2 and YSU use the same non-local closure approach to simulate mixing inside PBL, and this can lead to the similar behavior in reproducing the temperature and mixing ratio trend.

The diurnal variation of wind speed is analyzed in Fig. 3. All schemes confirm a difficulty in reconstructing the daily pattern of wind field. In particular, runs show a positive bias that ranges from 0.699–0.979 m s^{-1} (YSU), because of an overestimation of the observed values during the early morning and night. YSU scheme is found to generate the highest bias in the nighttime hours. Hong and Kim (2008) demonstrated that the increasing in the critical Richardson number during stable boundary conditions is responsible to the enhanced mixing when winds are generally weak. On the contrary, YSU has the lowest bias in correspondence of midday.

After the sunset ACM2 run produces lower overestimations than runs with MRF, MYJ or YSU. MYJ is also reported in Zhang and Zheng (2004) and Hu et al. (2010) to produce high wind speeds near the ground.

Finally, the modeled and observed vertical profiles are shown for comparison. Vertical profiles enable to further and better elucidate discrepancies observed in modeled results. In Fig. 4 is depicted the simulated and measured potential temperature, relative humidity and wind speed profiles at Milano Linate station for some days of the PBL experimental campaigns (6 February and 12 February at 11:00 UTC).

The early morning and night profiles show increasing potential temperature (inversion), decreasing humidity and wind shear with height. The local closure schemes have produced more realistic profiles than the non-local schemes.

In the convective conditions during daytime the above parameters are expected to have less pronounced variation with height in the well-mixed boundary layer. Wind speed shows important differences from run to run, but generally non-local closure schemes are closer to observations.

Local closure schemes have similar vertical wind profile on 6 February at 11:00 UTC with higher wind speed in the first 500 m than all other schemes. Differently, when the vertical profile of 12 February is analyzed, MYJ shows a larger surface wind speed than UW that results more similar to YSU. Both UW and YSU used an explicit term for simulating the entrainment zone, suggesting that either entrainment fluxes or the kind of closure play a key role in determining turbulence in the first meters of atmosphere during unstable conditions.

When analyzing the vertical profile of 6 Februay at 11:00 UTC it is quite clear that MYJ an UW have higher moisture and lower potential temperature than the other schemes below 500 m, while schemes are more similar above. Once again this can be partially related by the entrainment fluxes. As extensively discussed by Srinivas et al. (2007) and Hu et al. (2010), a possible explanation is a weak entrainment from the free troposphere in local schemes. The air above the PBL has higher potential temperature and less moisture than PBL air. A lack in the entrainment fluxes transport less warmer and drier air into the local PBL schemes with respect to YSU, ACM2 and MRF run, even though different entrainment approaches are used.

Relative humidity profiles can be also analyzed in order to obtain a first validation of PBL height. Since relative humidity is maximum inside the PBL (Seinfeld and Pandis, 1998; Ferrero et al., 2011a), it is possible to approximately estimate PBL depth looking at the height where relative humidity drops to a lower value. Vertical profiles indicate that relative humidity is characterized by an increasing from the ground level to a layer where a strong negative gradient is present. This layer is generally included in the first 500 m of the lower troposphere. The vertical profiles of 6 February shows that YSU, MYJ and UW predict the lowest PBL values around noon, while MRF and ACM2 overestimates the measured profile. On the other hand, the vertical profile of 12 February displays that all schemes overestimate PBL height at 11:00 UTC.

3.2 Planetary Boundary Layer height

Table 3 shows the performance statistics of the five parameterizations for both balloon and Lidar comparisons at Torre Sarca (Milan). It is worth noting that balloon soundings were available only in the morning, making them representative only of the early evolution of the PBL height (from 7:00–12:00 UTC).

In the morning, UW and YSU prove to have a quite coherent behavior in reproducing PBL height. Indeed, they have the best performances when compared to balloon data, even though UW is found to underestimate the observation (MB = −43.82 m) while YSU over predicts the overall morning values (MB = 30.15 m).

Also analyzing the diurnal variation of the mean bias, it is possible to state that YSU and UW give lower bias in reconstructing the hourly evolution of PBL until 12:00 UTC (Table 4). Both parameterizations under-predict balloon measurements from 7:00–9:00 UTC and overestimate them from 10:00–12:00 UTC. On the contrary, the largest overestimations are associated to ACM2 and MRF between 11:00–12:00 UTC (ACM2 MB = 248.05 m and MRF MB = 200.75 m). This is consistent with previous studies (Hong et al., 2006) in which it was demonstrated that YSU PBL increases the thermally induced mixing, while decreases the mechanically produced turbulence with respect to MRF, thus, partially resolving the problem of early development of PBL before noon.

Comparisons with Lidar measurements allow analyzing the performances over the whole day. Table 3 and Fig. 5 show that among the five schemes YSU and ACM2 are the most accurate overall (YSU MB = −27.54 m, ACM2 MB = 3 m). The YSU scheme is especially accurate during midday, while ACM2 gives the best performances during the evening and night thus improving its overall scores.

However, it is evident from Fig. 5 that all schemes have the same general daily trend of the mean bias. In the night and early morning, they show small biases which slightly increase around 9:00 UTC, reaching the highest values in the afternoon, and then decreasing again after 15:00 UTC. This general trend implies some common errors in all runs, such as the over-prediction of temperature and wind speed. Indeed the diurnal variations of the performance index are consistent with the daily trend of the temperature and mixing ratio bias discussed previously. All schemes were found to have the major temperature overestimation from 9:00–15:00 UTC when PBL mixing and biases increase. Moreover, the raising in PBL heights can help explaining the consequently decreasing in mixing ratio observed at those hours.

In particular, schemes underestimate morning values, while MRF, ACM2 and MYJ over predicted PBL depth around noon (ACM2 MB = 171.22 m, MRF MB = 85.51 m, MJY MB = 179.10 m). A slight overestimation is also highlighted for YSU from 10:00–11:00 UTC (MB ranges from 23.79–50.71 m), consistent with the balloon observations. UW generally under predicts Lidar data, even though a positive bias is visible around 11:00 UTC (MB = 8.61 m). As discussed before, UW has the smallest temperature overestimation, especially during the morning. This ranking seems to influence the early PBL development, thus, producing the lowest depths during the morning and over the whole day. On the contrary, in the afternoon the PBL evolution seems to be more related to entrainment fluxes parameterizations. Indeed, ACM2, MYJ and MRF adopt the same approach in reconstructing the entrainment processes.

A further comparison can be done in terms of hourly trend. Figures 6 and 7 depict the hourly modeled PBL height evolution compared to balloon and Lidar data, respectively.

As can be seen from these figures, in the early hours of the day there is closer agreement between models and among these and both Lidar and balloon observations. Deviation between simulated and observed grows with time. Indeed, the daytime development of PBL appears to be too rapid in all simulations. All schemes show a quick increase around 8:00–9:00 UTC, in line with the highest temperature bias, while observed values show a smoother variation. This pattern seems to confirm the previous hypothesis that relates the temperature estimations to the PBL early growth.

On the contrary, through afternoon there are substantial differences among the parameterizations. The largest overestimation exists for MRF, MYJ and ACM2 around noon. Under unstable conditions, ACM2 and MRF schemes compute PBL height using a similar method, and since the simulations with those schemes used the same surface layer, it makes sense that PBL heights are comparable. Moreover ACM2, MYJ and MRF adopt the implicit term to reproduce the entrainment zone, influencing the PBL development in a similar way.

The best agreement in reconstructing the PBL height can be shown in the YSU scheme in both Lidar and balloon comparisons, while UW generally under predicted its magnitude.

Finally all schemes show a too rapid decrease of PBL that collapses to the night-time value by 17:00 UTC, while Lidar data seem to report a smoother profile, especially on 11 and 12 February. The problem of the possible underestimation of the PBL height decay during the afternoon transition to nighttime condition in LIDAR data was then investigated. Lidar data were compared against hourly fine PM ($PM_{2.5}$) ground concentrations collected at Torre Sarca by means of OPC measurements (Optical Particle Counter, 1.107 “ENVIRO-check” GRIMM). Figure 8 shows a good agreement between the hourly variation of the PBL height and particle observations. $PM_{2.5}$ ground concentrations reveal a smooth increase during nighttime hours that correspond in time with the decrease in PBL Lidar estimations for both 11 and 12 February. These results confirm the robustness of the comparison between model and Lidar measurements even during the PBL afternoon decay. As a consequence, these results also confirm

the presence of common problems in all PBL schemes in urban areas, that could be partially related to the absence of an anthropogenic source of heating (Krpo et al., 2010), capable of limiting the too strong decrease of PBL height taking place at sunset.

4 Conclusions

The WRF model has been applied over the Po Valley area and compared against measurement stations, vertical profiles by balloon soundings and Lidar data in order to assess the skill of the meteorological model in reproducing PBL structure and evolution. Five PBL schemes were tested for a 5 km simulation in February 2008: three non-local closure schemes (ACM2, MRF and YSU) and two local closure parameterizations (MYJ and UW).

Vertical profiles of aerosol number concentrations and Lidar backscatter data were collected in the city of Milan (Torre Sarca site). The PBL height was derived in both cases using a gradient method. Additionally, meteorological radio-soundings of Milano Linate airport as well as surface temperature, mixing ratio and wind speed of different WMO meteorological sites were considered too.

At domain level, results show that all five parameterizations produce in the lower atmosphere higher temperatures, mixing ratio and wind speed values than the observed ones, implying some common biases in all model runs over the Po Valley in winter. The highest biases were associated to MYJ and UW parameterizations that use the Janjic Eta Monin–Obukhov surface layer scheme that was found to deliver a greater amount of sensible and latent heat fluxes with respect to Monin–Obukhov surface layer scheme.

Limiting the analysis to the observations collected in the city of Milan, the results follow the previous general pattern, over predicting temperature and wind speed, but underestimating mixing ratios. Moreover, the five schemes have different scores, with respect to the whole domain. UW and ACM2 show better performances than the other schemes for temperature, and wind speed. The best performance is related to

the ability of these schemes to predict more reliable results during the morning and evening. Indeed, ACM2 and UW use the same local closure during nighttime conditions that lead to an improvement in model performance.

Concerning PBL height, schemes have the same general trend in reproducing the daily variation. All runs show a closer agreement with observations during the night and early morning, a rapid increase around noon, and a fast collapse to the night values after 17:00 UTC. This general pattern suggests some systematic errors in all parameterizations. Over predictions of temperature and wind speed are found to cause a general overestimation of mixing during the PBL development. In particular, temperature estimations seem to mainly influence the early evolution of the PBL height. UW has the smallest temperature overestimation and the lowest PBL depths during the morning and over the whole day. Moreover, the diurnal variation of the performance index is consistent with the hourly trend of the temperature bias.

On the contrary during the afternoon and midday, the PBL evolution seems to be more related to the entrainment fluxes parameterizations. The largest afternoon overestimation exists for MRF, MYJ and ACM2 as they use the same implicit approach in reconstructing the entrainment process.

The best agreement in reconstructing the PBL height was highlighted for YSU. On the contrary, UW generally under predicted PBL magnitude at all daytime hours.

The obtained results enable to identify three key aspects that would require further analysis: (a) the modeling of the surface layer fluxes during the morning hours and their relationship with temperature and PBL growing; (b) the influence of the entrainment scheme in the development of the daytime development of PBL height; (c) the possible influence of anthropogenic sources of heating on the afternoon decay, by introducing a specific Urban Canopy Model in the computation of surface fluxes.

Appendix A: Statistical indicators

The statistical indicators selected to evaluate the model performances have been defined as follow:

Mean Bias (MB):

$$\mathrm{MB} = \frac{1}{N}\sum_{t=1}^{N} M(x,t) - O(x,t)$$

Root Mean Square Error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N} (M(x,t) - O(x,t))^2}$$

Pearson correlation index (r):

$$r = \frac{\sum_{t=1}^{N} \left(M(x,t) - \bar{M}(x)\right) \cdot \left(O(x,t) - \bar{O}(x)\right)}{\sqrt{\sum_{t=1}^{N} \left(M(x,t) - \bar{M}(x)\right)^2} \cdot \sqrt{\sum_{t=1}^{N} \left(O(x,t) - \bar{O}(x)\right)^2}}$$

$M(x,t)$ – Computed field
$O(x,t)$ – Observed field
N – Number of pairs

A cut-off threshold has been applied to observed and modeled fields to correct data from the presence of outliers. Data are rejected when differences between model and observation are:

Temperature $> 20\,^{\circ}\mathrm{K}$; Mixing ratio $> 10\,\mathrm{g\,kg^{-1}}$; Wind speed $< 0.5\,\mathrm{m\,s^{-1}}$ and Wind speed $> 100\,\mathrm{m\,s^{-1}}$.

Acknowledgements. RSE contribution to this work has been financed by the Research Fund for the Italian Electrical System under the Contract Agreement between RSE S.p.A. and the Ministry of Economic Development – General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency in compliance with the Decree of 8 March 2006.

References

Amiridis, V., Melas, D., Balis, D. S., Papayannis, A., Founda, D., Katragkou, E., Giannakaki, E., Mamouri, R. E., Gerasopoulos, E., and Zerefos, C.: Aerosol Lidar observations and model calculations of the Planetary Boundary Layer evolution over Greece, during the March 2006 Total Solar Eclipse, Atmos. Chem. Phys., 7, 6181–6189, doi:10.5194/acp-7-6181-2007, 2007.

Angelini, F., Barnaba, F., Landi, T. C., Caporaso, L., and Gobbi, G. P.: Study of atmospheric aerosols and mixing layer by LIDAR, Radiat. Prot. Dosim., 137, 275–279, 2009.

Appel, K. W., Gilliam, R. C., Davis, N., Zubrow A., and Howard, S. C.: Overview of the atmospheric model evaluation tool (AMET) v1.1 for evaluating meteorological and air quality models, Environ. Modell Softw., 26, 434–443, 2011.

Borge, R., Alexandrov, V., del Vas, J. J., Lumbreras, J., and Rodrıguez, E.: A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula, Atmos. Environ., 42, 8560–8574, 2008.

Bretherton, C. S. and Park, S.: A new moisture parameterization in the community atmosphere model, J. Climate, 22, 3422–3448, 2009.

Carbone, C., Decesari, S., Mircea, M., Giulianelli, L., Finessi, E., Rinaldi, M., Fuzzi, S., Marinoni, A., Duchi, R., Perrino, C., Sargolini, T., Vardè, M., Sprovieri, F., Gobbi, G. P., Angelini, F. and Facchini, M. C.: Size-resolved aerosol chemical composition over the Italian Peninsula during typical summer and winter conditions, Atmos. Environ., 44, 5269–5278, 2010.

Chen, F. and Dudhia, J.: Coupling an advanced land surface-hydrology model with the penn state–NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Mon. Weather. Rev., 129, 569–585, 2001.

Chou, M.-D., Suarez, M. J., Ho, C.-H., Yan, M. M.-H., and Lee, K.-T.: Parameterizations for cloud overlapping and shortwave single-scattering properties for use in general circulation and cloud ensemble models, J. Climate, 11, 202–214, 1998.

Cohn, S. A. and Angevine, W. M.: Boundary layer height and entrainment zone thickness measured by lidars and wind-profiling radars, J. Appl. Meteorol., 39, 1233–1247, 2000.
Endlich, R. M., Ludwig, F. L., and Uthe, E. E.: An automatic method for determining the mixing depth from lidar observations, Atmos. Environ., 13, 1051–1056, 1967.
Ferrero, L., Perrone, M. G., Petraccone, S., Sangiorgi, G., Ferrini, B. S., Lo Porto, C., Lazzati, Z., Cocchi, D., Bruno, F., Greco, F., Riccio, A., and Bolzacchini, E.: Vertically-resolved particle size distribution within and above the mixing layer over the Milan metropolitan area, Atmos. Chem. Phys., 10, 3915–3932, doi:10.5194/acp-10-3915-2010, 2010.
Ferrero, L., Riccio, A., Perrone, M. G., Sangiorgi, G., Ferrini, B. S., and Bolzacchini, E.: Mixing height determination by tethered balloon-based particle soundings and modeling simulations, Atmos. Res., 102, 145–156, 2011a.
Ferrero, L., Mocnik, G., Ferrini, B. S., Perrone, M. G., Sangiorgi, G., and Bolzacchini, E.: Vertical profiles of aerosol absorption coefficient from micro-aethalometer data and Mie calculation over Milan, Sci. Total Environ., 409, 2824–2837, 2011b.
Ferrero, L., Cappelletti, D., Moroni, B., Sangiorgi, G., Perrone, M. G., Crocchianti, S., and Bolzacchini, E.: Wintertime aerosol dynamics and chemical composition across the mixing layer over basin valleys, Atmos. Environ., 56, 143–153, 2012.
Ferrero, L., Castelli, M., Ferrini, B. S., Moscatelli, M., Perrone, M. G., Sangiorgi, G., Rovelli, G., D'Angelo, L., Moroni, B., Scardazza, F., Mocnik, G., Bolzacchini, E., Petitta, M., and Cappelletti, D.: Impact of black carbon aerosol over Italian basin valleys: high resolution measurements along vertical profiles, radiative forcing and heating rate, Atmos. Chem. Phys. Discuss., 14, 541–591, doi:10.5194/acpd-14-541-2014, 2014.
Grell, G. A. and Dévényi, D.: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, Geophys. Res. Lett., 29, 38-1–38-4, doi:10.1029/2002GL015311, 2002.
Grell, G., Dudhia, J., and Stauffer, D.: A description of the fifth-generation penn state/NCAR Mesoscale Model (MM5), National Center for Atmospheric Research, Boulder, 117 pp., 1994.
Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G. P., Lolli, S., O'Dowd, C. D., Sauvage, L., Xueref-Rémy, I., Wastine, B., Feist, D. G.: Evaluation of mixing height retrievals from automatic profiling lidars and ceilometers in view of future integrated networks in Europe, Bound.-Lay. Meteorol., 143, 49–75, doi:10.1007/s10546-011-9643-z, 2012.

Hennemuth, B. and Lammert, A.: Determination of the atmospheric boundary layer height from radiosonde and lidar backscatter, Bound.-Lay. Meteorol., 120, 181–200, 2006.
Hong, S.-Y. and Kim, S. K.: Stable boundary layer mixing in a vertical diffusion scheme, in: Proceeding of the Ninth Annual WRF User's Workshop, Boulder, CO, 3.3., 2008.
Hong, S.-Y. and Pan, H.-L.: Nonlocal boundary layer vertical diffusion in a medium-range forecast model, Mon. Weather Rev., 124, 2322–2339, 1996.
Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Mont. Weather Rev., 134, 2318–2341, 2006.
Hooper, W. P. and Eloranta, E. W.: Lidar measurements of wind in the planetary boundary layer: the method, accuracy and results from joint measurements with radiosonde and kytoon, J. Clim. Appl. Meteorol., 25, 990–1001, 1986.
Hu, X.-M. and Nielsen-Gammon, J. W.: Evaluation of three planetary boundary layer schemes in the WRF model, J. Appl. Meteorol. Clim., 49, 1831–1844, 2010.
Janjic, Z. I.: The step-mountain eta coordinate model: further developments of the convection, viscous sublayer, and turbulence closure schemes, Mont. Weather Rev., 122, 927–945, 1994.
Janjic, Z. I.: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP meso model, NOAA/NWS/NCEP, 61 pp., 2001.
Kim, S. W., Yoon, S. C., Won, J. G., and Choi, S. C.: Ground-based remote sensing measurements of aerosol and ozone in an urban area: a case study of mixing height evolution and its effect on ground-level ozone concentrations, Atmos. Environ., 41, 7069–7081, 2007.
Krpo, A., Salamanca, F., Martilli, A., and Clappier, A.: On the impact of anthropogenic heat fluxes on the urban boundary layer: a two-dimensional numerical study, Bound.-Lay. Meteorol., 136, 105–127, 2010.
Krupa, S. V., Grunhage, L., Jager, H. J., Nosal, M., Manning, W. J., Legge, A. H., and Hanewald, K.: Ambient ozone (O_3) and adverse crop response: a unified view of cause and effect, Environ. Pollut., 87, 119–126, 2006.
Laakso, L., Grönholm, T., Kulmala, L., Haapanala, S., Hirsikko, A., Lovejoy, E. R., Kazil, J., Kurtén, T., Boy, M., Nilsson, E. D., Sogachev, A., Riipinen, I., Stratmann, F., and Kulmala, M.: Hot-air balloon as a platform for boundary layer profile measurements during particle formation, Boreal. Environ. Res., 12, 279–294, 2007.

Lin, J.-T., Youn, D., Liang, X.-Z., and Wuebbles, D. J.: Global model simulation of summertime US ozone diurnal cycle and its sensitivity to PBL mixing, spatial resolution, and emissions, Atmos. Environ., 42, 8470–8483, 2008.

Maletto, A., McKendry, I. G., Strawbridge, K. B.: Profiles of particulate matter size distributions using a balloon-borne lightweight aerosol spectrometer in the planetary boundary layer, Atmos. Environ., 37, 661–670, 2003.

Martucci, G., Matthey, R., Mitev, V., and Richner, H.: Comparison between Backscatter Lidar and Radiosonde Measurements of the Diurnal and Nocturnal Stratification in the Lower Troposphere, J. Atmos. Ocean. Tech., 24, 1231–1244, doi:10.1175/JTECH2036.1, 2007.

McKendry, I. G., Sturman, A. P., and Vergeiner, J.: Vertical profiles of particulate matter size distributions during winter domestic burning in Christchurch, New Zealand, Atmos. Environ., 38, 4805–4813, 2004.

Melfi, S. H., Spinhirne J. D., Chou, S.-H., and Palm, S. P.: Lidar observations of vertically organized convection in the planetary boundary layer over the ocean, J. Clim. Appl. Meteorol., 24, 806–821, 1985.

Mellor, G. L. and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Rev. Geophys. Space. Ge., 20, 851–875, 1982.

Misenis, C. and Zhang, Y.: An examination of sensitivity of WRF/Chem predictions to physical parameterizations, horizontal grid spacing, and nesting options, Atmos. Res., 97, 315–334, 2010.

Misenis, C., Hu, X.-M., Krishnan, S., Zhang, Y., and Fast, J.: Sensitivity of WRF/Chem predictions to meteorological schemes, in: Proceeding of the 86th Annual Conference/14th Joint Conf. on the Applications of Air Pollution Meteorology with the A&WMA, Atlanta, GA, 1.8, 2006.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, J. Geophys. Res., 102, 16663–16682, 1997.

Monin, A. S. and Obukhov, A. M.: Basic laws of turbulent mixing in the ground layer of the atmosphere, Trans. Geophys. Inst. Akad. Nauk. USSR 151, 163–187, 1954.

Morille, Y., Haeffelin, M., Drobinski, P., and Pelon, J.: STRAT: An automated algorithm to retrieve the vertical structure of the atmosphere from single-channel lidar data, J. Atmos. Ocean. Tech., 24, 761–775, 2007.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud micrpohysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one- and two-moment schemes, Mon. Weather. Rev., 137, 991–1007, 2009.

Perrone, M. G., Gualtieri, M., Consonni, V., Ferrero, L., Sangiorgi, G., Longhin, E., Ballabio, D., Bolzacchini, E. and Camatini, M.: Particle size, chemical composition, seasons of the year and urban, rural or remote site origins as determinants of biological effects of particulate matter on pulmonary cells, Environ. Pollut., 176, 215–27, 2013.

Pleim, J. E.: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing, J. Appl. Meteorol. Clim., 46, 1383–1395, 2007.

Pleim, J. E. and Chang, J. S.: A non-local closure model for vertical mixing in the convective boundary layer, Atmos. Environ., 26, 965–981, 1992.

Rodríguez, S., Van Dingenen, R., Putaud, J.-P., Dell'Acqua, A., Pey, J., Querol, X., Alastuey, A., Chenery, S., Ho, K.-F., Harrison, R., Tardivo, R., Scarnato, B., and Gemelli, V.: A study on the relationship between mass concentrations, chemistry and number size distribution of urban fine aerosols in Milan, Barcelona and London, Atmos. Chem. Phys., 7, 2217–2232, doi:10.5194/acp-7-2217-2007, 2007.

Sangiorgi, G., Ferrero, L., Perrone, M. G., Bolzacchini, E., Duane, M., and Larsen, B. R.: Vertical distribution of hydrocarbons in the low troposphere below and above the mixing height: tethered balloon measurements in Milan, Italy, Environ. Pollut., 159, 3545–3552, 2011.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics, From Air Pollution to Climate Change, John Wiley and Sons, Inc., NY, 1998.

Skamarock, W. C., Klemp, J. B., Dudhia J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, National Centre of Atmospheric Research, Boulder, Colorado, 2008.

Srinivas, C. V., Venkatesan R., and Bagavath Singh, A.: Sensitivity of mesoscale simulations of land-sea breeze to boundary layer turbulence parameterization, Atmos. Environ., 41, 2534–2548, 2007.

Stull, R. B.: An introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, the Netherlands, 1989.

Summa, D., Di Girolamo, P., Stelitano, D., and Cacciani, M.: Characterization of the planetary boundary layer height and structure by Raman lidar: comparison of different approaches, Atmos. Meas. Tech., 6, 3515–3525, doi:10.5194/amt-6-3515-2013, 2013.

Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer: sensitivity to surface evaporation, Bound.-Lay. Meteorol., 37, 129–148, 1986.

Utell, M. J.: Inhalation of ultrafine particles alters blood leukocyte expression of adhesion molecules in humans, Environ. Health. Persp., 114, 51–58, 2006.

Wiegner, M., Emeis, S., Freudenthaler, V., Heese, B., Junkermann, W., Münkel, C., Schäfer, K., Seefeldner, M., and Vogt, S.: Mixing layer height over Munich, Germany: variability and comparisons of different methodologies, J. Geophys. Res., 111, D13201, doi:10.1029/2005JD006593, 2006.

Yerramilli A., Challa V. S., Dodla V. B. R., Dasari H. P., Young J. H., Patrick C., Baham J. M., Hughes R. L., Hardy M. G., and Swanier S. J.: Simulation of surface ozone pollution in the Central Gulf Coast Region using WRF/ChemModel: sensitivity to PBL and land surface physics, Adv. Meteorol., 2010, 1–24, doi:10.1155/2010/319138, 2010.

Zhang, Y. and Zheng, W.-Z.: Diurnal cycles of surface winds and temperatures as simulated by five boundary layer parameterizations, J. Appl. Meteorol., 43, 157–169, 2004.

Zhang, Y., Liu, P., Pun, B., and Seigneur, C.: A comprehensive performance evaluation of MM5-CMAQ for the summer 1999 Southern Oxidants Study Episode – Part I. Evaluation protocols, databases and meteorological predictions, Atmos. Environ., 40, 4825–4838, 2006.

Zhong, S., In, H., and Clements, C.: Impact of turbulence, land surface, and radiation parameterizations on simulated boundary layer properties in a coastal environment, J. Geophys. Res., 112, D13110, doi:10.1029/2006JD008274, 2007.

Table 1. Comparison of PBL schemes against 63 WMO ground-based meteorological stations in the Po Valley for 2 m-temperature, 2 m-mixing ratio and 10 m-wind speed. The best performances are bold.

	ACM2	MRF	MYJ	UW	YSU
Temperature (K)					
Mean Obs	277.487	277.487	277.487	277.487	277.487
Mean Mod	278.088	278.346	278.363	278.398	278.255
MB	**0.601**	0.859	0.876	0.911	0.767
RMSE	3.501	3.670	**3.285**	3.360	3.465
r	0.833	0.827	**0.840**	0.839	0.837
Mixing Ratio ($g\,kg^{-1}$)					
Mean Obs	4.577	4.577	4.577	4.577	4.577
Mean Mod	4.618	4.551	4.855	4.842	4.699
MB	0.041	**−0.026**	0.278	0.265	0.122
RMSE	**0.811**	0.848	0.966	0.933	0.832
r	**0.852**	0.842	0.822	0.822	0.850
Wind Speed ($m\,s^{-1}$)					
Mean Obs	3.237	3.237	3.237	3.237	3.237
Mean Mod	3.591	3.617	4.013	4.005	3.694
MB	**0.354**	0.381	0.776	0.768	0.457
RMSE	2.916	2.921	**2.870**	2.905	2.935
r	0.469	0.465	**0.521**	0.509	0.471

Table 2. Performance statistics of the five configurations at Milano Linate station for 2 m-temperature, 2 m-mixing ratio and 10 m-wind speed. The best performances are highlighted bold.

	ACM2	MRF	MYJ	UW	YSU
Temperature (K)					
Mean Obs	278.980	278.980	278.980	278.980	278.980
Mean Mod	279.586	279.904	279.369	279.189	279.961
MB	0.606	0.924	0.389	**0.209**	0.981
RMSE	**1.972**	2.170	1.975	2.110	2.093
r	0.896	0.886	0.885	0.878	**0.897**
Mixing Ratio ($g\,kg^{-1}$)					
Mean Obs	4.735	4.735	4.735	4.735	4.735
Mean Mod	4.527	4.459	4.669	4.620	4.605
MB	−0.208	−0.276	**−0.066**	−0.115	−0.131
RMSE	0.693	0.729	0.637	**0.636**	0.688
r	0.874	0.871	0.886	**0.888**	0.869
Wind Speed ($m\,s^{-1}$)					
Mean Obs	1.721	1.721	1.721	1.721	1.721
Mean Mod	2.419	2.536	2.593	2.639	2.699
MB	**0.699**	0.815	0.873	0.918	0.979
RMSE	**1.563**	1.702	1.738	1.845	2.092
r	0.331	0.270	0.218	**0.407**	0.300

Table 3. Performance indexes in the five configurations for PBL height; statistics derived from a comparison with both balloon and Lidar data. The best performances are highlighted in bold.

	ACM2	MRF	MYJ	UW	YSU
Balloon					
Meab Obs (m)	260.88	260.88	260.88	260.88	260.88
Mean Mod (m)	376.22	369.81	360.65	217.06	291.03
MB (m)	115.34	108.93	99.77	−43.82	**30.15**
RMSE (m)	211.76	186.89	185.17	137.39	**104.64**
r	0.60	0.71	0.59	0.51	**0.76**
Lidar					
Meab Obs (m)	271.57	271.57	271.57	271.57	271.57
Mean Mod (m)	274.57	222.04	321.60	171.76	244.03
MB (m)	**3.00**	−49.53	50.04	−99.81	−27.54
RMSE (m)	170.75	189.96	168.82	181.32	**156.94**
r	**0.84**	**0.84**	0.66	0.83	0.68

Table 4. Hourly variation of the mean bias index in the five configurations for PBL height balloon campaign. The best performances are highlighted bold.

MB (m)	ACM2	MRF	MYJ	UW	YSU
7:00–8:00	−121.13	−122.97	**−78.33**	−144.85	−132.27
8:00–9:00	**−2.66**	38.73	19.21	−81.91	−5.25
9:00–10:00	175.34	159.63	160.85	−82.96	**52.10**
10:00–11:00	220.27	191.81	169.96	**49.82**	103.43
11:00–12:00	248.05	200.75	164.44	**59.85**	78.42

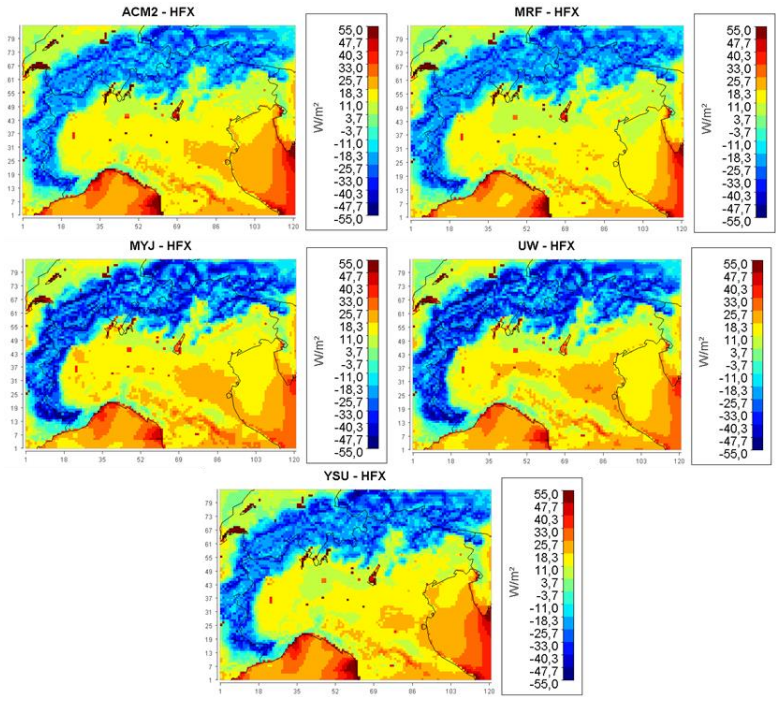

Figure 1. Monthly mean (February 2008) of upward sensible heat fluxes at the surface (HFX) in the five configurations.

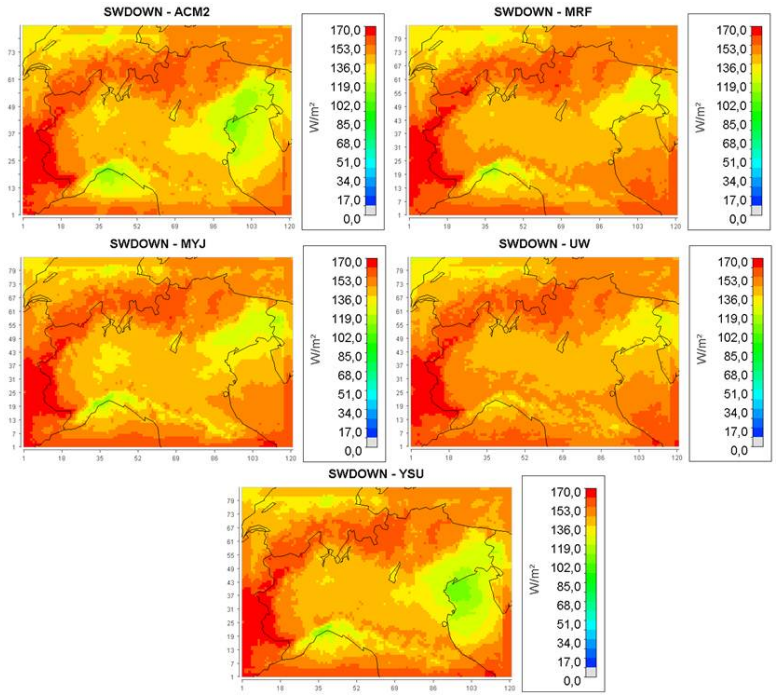

Figure 2. Monthly mean (February 2008) of downward shortwave flux at the ground surface (SWDOWN) in the five configurations.

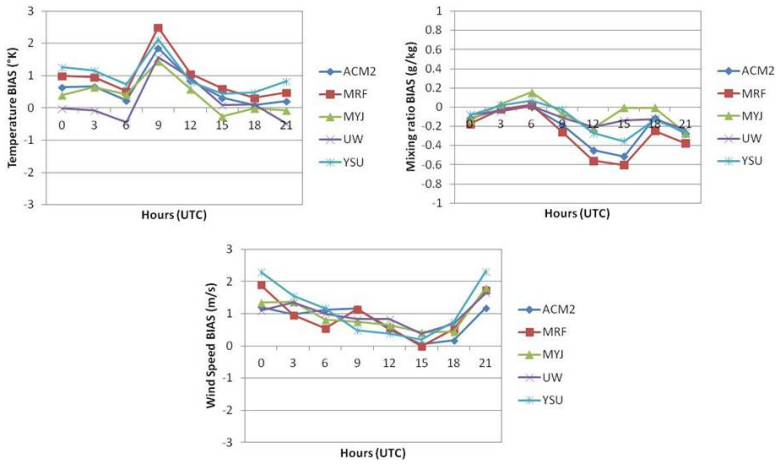

Figure 3. Diurnal variation of mean bias index at Milano Linate station.

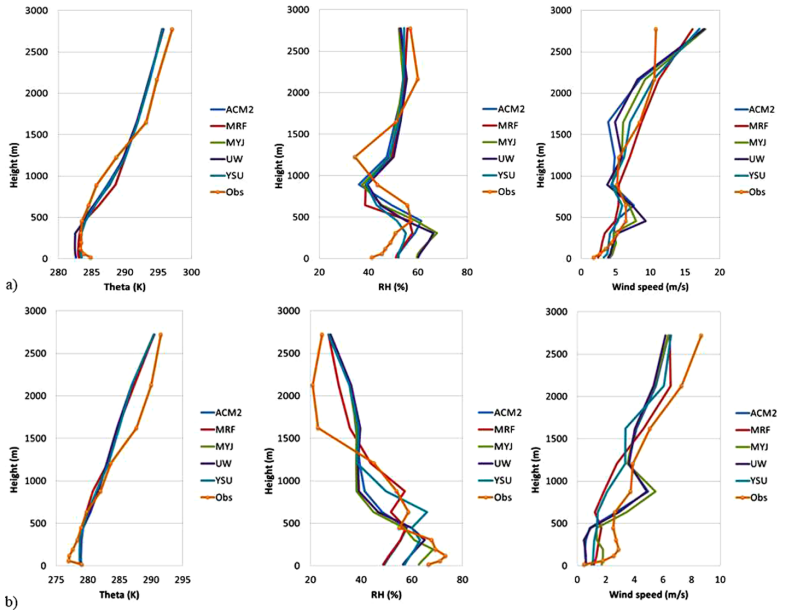

Figure 4. Vertical profile of potential temperature (K), relative humidity (%) and wind speed ($m\,s^{-1}$) on 6 February 2008 at 11:00 UTC **(a)** and on 12 February 2008 at 11:00 UTC **(b)**.

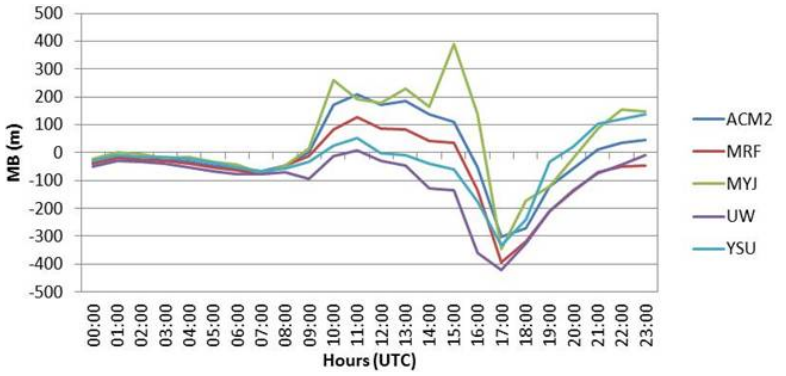

Figure 5. Diurnal trend of mean bias (m) index in the five configurations for PBL height Lidar data.

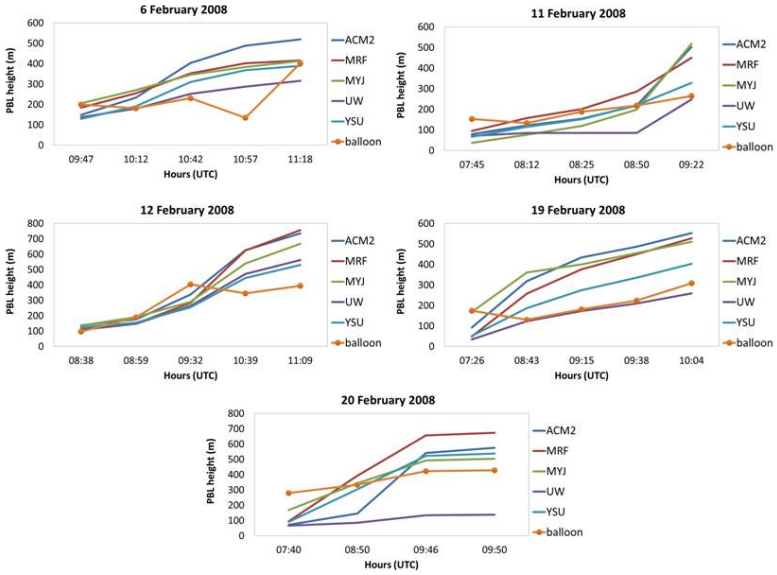

Figure 6. Comparison between five PBL schemes and balloon observations at Torre Sarca at different days, February 2008.

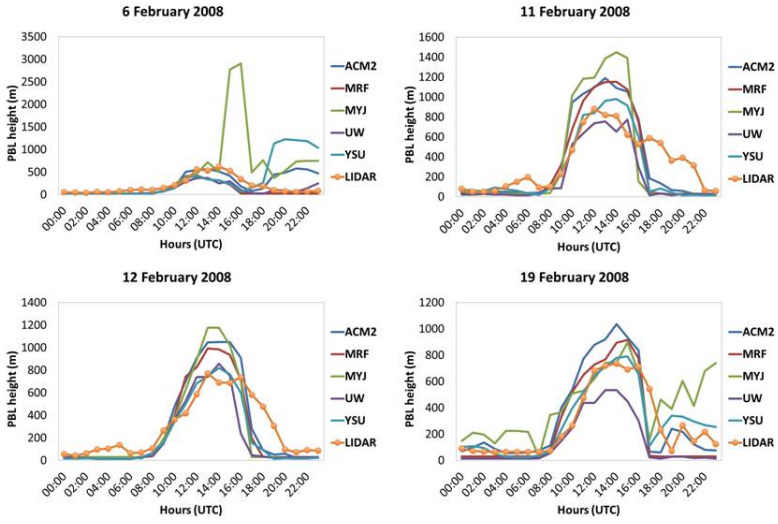

Figure 7. Comparison between five PBL schemes and Lidar measurements at Torre Sarca at different days, February 2008.

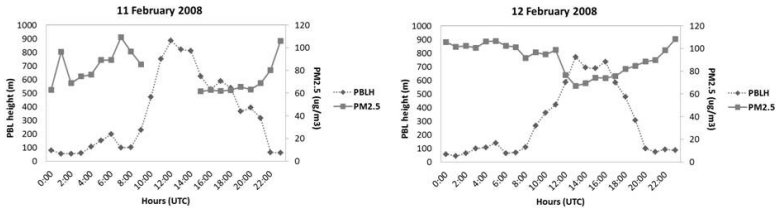

Figure 8. Comparison between PBL height Lidar estimations and hourly fine PM ($PM_{2.5}$) ground concentrations at Torre Sarca at two different days of February 2008.