A Planetary Boundary Layer Height Climatology Derived from ECMWF Reanalysis Data

AXEL VON ENGELN
EUMETSAT, Darmstadt, Germany

JOÃO TEIXEIRA
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 17 June 2012, in final form 5 February 2013)

ABSTRACT

A planetary boundary layer (PBL) height climatology from ECMWF reanalysis data is generated and analyzed. Different methods are first compared to derive PBL heights from atmospheric temperature, pressure, and relative humidity (RH), which mostly make use of profile gradients, for example, in RH, refractivity, and virtual or potential temperature. Three methods based on the vertical gradient of RH, virtual temperature, and potential temperature were selected for the climatology generation. The RH-based method appears to capture the inversion that caps the convective boundary layer very well as a result of its temperature and humidity dependence, while the temperature-based methods appear to capture the PBL better at high latitudes. A validation of the reanalysis fields with collocated radiosonde data shows generally good agreement in terms of mean PBL height and standard deviation for the RH-based method. The generated ECMWF-based PBL height climatology shows many of the expected climatological features, such as a fairly low PBL height near the west coast of continents where stratus clouds are found and PBL growth as the air is advected over warmer waters toward the tropics along the trade winds. Large seasonal and diurnal variations are primarily found over land. The PBL height can exceed 3 km, mostly over desert areas during the day, although large values can also be found in areas such as the ITCZ. The robustness of the statistics was analyzed by using information on the percentage of outliers. Here in particular, the sea-based PBL was found to be very stable.

1. Introduction

The atmospheric boundary layer has a profound influence on the physics of the atmosphere and of the climate system as a whole. Cloud processes, land and ocean surface fluxes, and the atmospheric hydrological cycle in general, are strongly influenced by the boundary layer. Boundary layer processes play a role in the modulation of the atmosphere at weather and climate time scales and as such have an important role in extreme weather and climatic events.

In spite of some recent progress, weather and climate prediction models still do not fully, realistically represent boundary layer processes, since most of the processes associated with the boundary layer occur at subgrid scales when compared to typical climate and weather prediction horizontal grid scales. As a consequence, boundary layer parameterizations that represent the statistical behavior of boundary layer processes at the model grid scale have to be developed. The development of realistic boundary layer parameterizations, representing the complex variety in the earth’s atmosphere, is still an unresolved challenge in climate and weather prediction (Teixeira et al. 2008).

One major issue in this context is the lack of global observational datasets. The atmospheric boundary layer has been notoriously difficult to observe from space, preventing a detailed understanding of its properties at global scales. One of the most relevant boundary layer properties to investigate is the planetary boundary layer (PBL) height (or inversion height) (Medeiros et al. 2005). The PBL height characterizes the PBL in a fairly integrated manner and can be closely related to fundamental variables such as cloud cover and water, surface
(and boundary layer) heat fluxes, contaminant dispersion, and ducting.

A 10-yr PBL climatology based on a pure general circulation model run is presented in Medeiros et al. (2005), along with a detailed analysis of the PBL drivers. In their model setup, the PBL is synonymous with the lowest model layer. The horizontal resolution is rather coarse at 4° by 5°. The authors identify several PBL regimes that tend to exhibit large-scale geographic organization where locally generated buoyancy fluxes and static stability control PBL height nearly everywhere. However, the model uses no real observations and is based on early climate model developments. More recently, work by Seidel et al. (2010) used radiosonde data to lay some groundwork for the generation of a PBL climatology, albeit thus generally restricted to locations on land. A first climatology based on this earlier work is available in Seidel et al. (2012), focusing on the U.S. and European regions.

There have also been some studies of boundary layer properties from space-based remote sensing instruments, for example by von Engeln et al. (2005), Sokolovskiy et al. (2006), and Basha and Ratnam (2009), who used radiosonde occultation observations to infer the PBL height. This kind of observation has a high vertical resolution, down to below 100 m, albeit at a low horizontal resolution around a few hundred kilometers. Hence the derived PBL height is an average over larger horizontal areas. Other satellite-based methods used, for example, infrared observations from the Atmospheric Infrared Sounder (AIRS) instrument (Fetzer et al. 2004; Martins et al. 2010) or Moderate Resolution Imaging Spectroradiometer (MODIS) and Tropical Rainfall Measuring Mission (TRMM) data to derive the PBL height in selected regions (Wood and Bretherton 2004). Other authors (e.g., Karlsson et al. 2010) have used highly accurate estimates of cloud-top height from the Multiangle Imaging SpectroRadiometer (MISR) to investigate the subtropical PBL height evolution.

A global climatology for PBL height can be derived from reanalysis data, which combines information from observations and models; it can provide important insight into the physics of the boundary layer from a global perspective and can be used to evaluate weather and climate prediction models. Long-term reanalysis runs [such as the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) or the ECMWF Interim Re-Analysis (ERA-Interim) (Simmons and Gibson 2000; Simmons et al. 2007; Dee et al. 2011)] are in several respects still the best source of global climatological information for many variables. Reanalysis systems assimilate a variety of measurements ranging from in situ to space-based instruments into weather prediction models.

In this work we generate a PBL height climatology from 20 years of reanalysis data. The ERA-Interim reanalysis focuses on the data-rich period from 1989 onward; it is provided continuously in parallel to the operational analysis to support climate monitoring. It is thus well suited for such a task. The PBL height is derived from several atmospheric parameters, as detailed below. The primary focus is on relative humidity, however, virtual and potential temperatures are also discussed in more detail.

The paper is structured as follows: we first introduce the used datasets for this study and the processing setup (section 2) and then discuss several possibilities to determine the PBL height (section 3), where we preselect different methods to determine the PBL height global climatology. A validation of the ECMWF fields with collocated radiosonde data is performed in section 4. In section 5 the general characteristics of the PBL climatology are presented, and section 6 presents conclusions.

2. Input data and general processing setup

In this study, the ERA-Interim reanalysis and forecast fields (Simmons et al. 2007) are used. Climatological records are derived from 20 years of data (1990 to 2009). This includes the most recent and “best” quality observation period that starts from about 1989 onward and already assimilates several standard satellite observations from geostationary orbit (Meteosat, Geostationary Operational Environmental Satellite (GOES), and Geostationary Meteorological Satellite (GMS)), from low-earth-orbit (National Oceanic and Atmospheric Administration (NOAA) and Defense Meteorological Satellite Program (DMSP)), and stretches into the modern era with hyperspectral sounders such as the Infrared Atmospheric Sounding Interferometer (IASI) and AIRS, and radio occultation instruments such as the Challenging Minisatellite Payload (CHAMP), Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC), and Global Navigation Satellite System Receiver for Atmospheric Sounding (GRAS).

For the most part of the study, temperature, surface pressure, geopotential, and specific humidity fields of ERA-Interim are used, and variables for PBL height determination are derived from these core fields. In addition, selected ERA-Interim fields such as the planetary boundary layer height variable are directly used for the year 2000 in order to perform some further validations and comparisons; these are available at 6-h and 12-h forecast times. The horizontal resolution of the ERA-Interim fields is 1.0° and in total 60 vertical hybrid levels are available [about 21 levels between 0 km
and 5 km, starting with a resolution of about 25 m near the ground, decreasing to about 200 m around 1 km (Teixeira 1999a). Four analysis times are used at 0000, 0600, 1200, and 1800 UTC (coordinated universal time). Additionally, ERA-40 data for the years up to 2001 (at a lower resolution of 1.5°) are also used for comparison.

The methods to estimate the PBL height in this paper are generally making use of vertical gradients. All vertical profiles were thus interpolated to a 20 m resolution up to 5 km (a 20-m interpolation assures that the underlying ECMWF or sonde resolution is kept, but it has no influence on the PBL height calculation methods). To avoid selecting a very low PBL height caused by temperature inversions near the surface, we excluded all data below 50 m in our analysis, again both in ECMWF and in sonde data. Similar screening is also used by other authors to remove noisy data near the ground (e.g., Liu and Liang 2010; Seidel et al. 2010).

Monthly averages for each analysis time and month were calculated initially. For seasonal averages, all analysis times of the corresponding months were averaged as follows: December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON). For UTC time-based averaging, all years and months were averaged at that specific time step.

For a validation of the ERA-Interim gradients (primarily in RH) we used two sets of radiosondes, covering ocean and land conditions. For ocean observations, we used data from the Alfred Wegener Institute (AWI). For land observations, we also obtained data from selected sites through the University of Wyoming website, spanning 10 years from 2000 to 2009.

AWI sondes are launched from the research vessel Polarstern (König-Langlo and Marx 1997). Starting late 1982, VAISALA RS80 radiosondes were launched during research cruises of the ship. Cruises were mostly in the polar regions and the Atlantic (see, e.g., Beyerle et al. 2006) for a typical location plot. In total, there are more than 7300 sondes available from the 1990 to 2009 time span, where about 900 are found at low latitudes, 1700 at midlatitudes, and the majority at high latitudes. The vertical resolution is up to 30 m. The land-based sonde observations cover more than 30 000 profiles; however, we did limit the validation to specific times since most stations do not cover all investigated UTC times.

The preprocessing is essentially identical to the one of the ECMWF fields mentioned above, where the lowest 50 m are excluded and all data are interpolated to a 20-m vertical grid. In addition, a further refinement was introduced for PBL heights based on the RH gradient, where the algorithm only searches first for altitudes with gradients $< -100.0 \% \text{ km}^{-1}$. The lowest found layer of this search is denoted as PBL height. Note that this additional step was introduced in order to remove the large variations in such a high-resolution dataset.

The Wyoming data provide good annual coverage; these data were thus statistically evaluated per month and used to validate the averaged month-by-month variations. The AWI data were used as is to validate the ECMWF fields at the time and location of the sonde observations.

All mean and standard deviation calculations performed here are based on a robust estimator (Tukey’s biweight; Hoaglin et al. 1983). The robust estimator is an effective tool to deweight outliers from noisy distributions, returning standard deviation and percentage of data points falling into the $\pm 2 \sigma$ interval. This would be 95% for an ideal Gauss curve. This tool thus also allows us to determine how Gaussian or robust such a climatology is.

Note that all PBL heights reported here are with respect to the local elevation.

3. Methods to determine PBL height

Several possibilities exist to determine the PBL height from atmospheric fields, mostly based on vertical gradients. Within this study we investigate the following gradient-based methods on profiles derived from ERA-Interim reanalysis fields: 1) minimum gradient of relative humidity, PBL$_{RH}$, 2) maximum gradient of potential temperature, PBL$_{Tp}$, 3) maximum gradient of virtual temperature, PBL$_{Tv}$, 4) minimum gradient of specific humidity $q$, PBL$_q$, and 5) minimum gradient of refractivity $N$, PBL$_N$.

In addition, we also investigated several other methods that are reported in the literature, or are used by researchers to estimate PBL growth and height: 6) scaled $q$ that identifies the altitude where $q$ decreases by 50% from its surface value, PBL$_{sq}$, and 7) breakpoint method, which searches for the break point in the N profile, PBL$_{BP}$.

These methods rely on different thermodynamic properties of the vertical profile. The PBL$_{Tp}$ uses information on the vertical gradient of potential temperature, while PBL$_{Tv}$ also includes humidity information and, in practice, finds the strongest gradient in the profile of atmospheric buoyancy. The PBL$_q$ method identifies the sharp gradient in humidity at the top of the PBL while PBL$_{sq}$ uses the fact that within a well-mixed PBL the specific humidity varies little with height. PBL$_N$ is often used for deriving the PBL height from radio occultation data (see, e.g., Ao et al. 2012), just as PBL$_{BP}$, following a method developed by Sokolovskiy et al. (2006) and applied, for example, as well in Guo et al. (2011).
It is fair to say that there is no perfect way to determine the height of the PBL and that the appropriateness of each method depends on the topic being investigated. In our case, the main interest is in characterizing in the best and most complete manner possible, boundary layers associated with clouds over the tropical, subtropical, and midlatitude oceans. Here, a method based on the gradient of RH provides a fairly robust way of determining the PBL inversion. This is due to the fact that RH combines both temperature and specific humidity; hence gradients in both temperature and specific humidity will be “amplified” when looking at RH. In addition, RH is intimately associated with clouds and the strong decrease in RH at the PBL top is a strong indication of the cloud top. Since we are particularly interested in capturing the properties of the cloudy boundary layer, this is an important property.

Given the fact that the boundary layer is often associated with stronger mixing (as compared to the free troposphere) due to increased levels of turbulence, it would be natural to investigate properties associated with turbulence. Other authors have done this by, for example, associating the top of the PBL with a certain value of the Richardson number (which characterizes the degree of turbulence). There are, however, several issues with this approach. For one, it is often more difficult to accurately reproduce the value of the Richardson number (which is based on ratios of both dynamic and thermodynamic vertical gradients) than just obtaining a value for the gradient of temperature and/or humidity. On the other hand, Richardson-based methods often use dry variables (as opposed to moist conserved variables), leading to results that may differ significantly from the real cloud or PBL top in cloudy situations. Finally, an essential problem is that the Richardson number is just a measure of local turbulence and is often unable to properly characterize the turbulent properties of convective boundary layers such as stratocumulus, shallow cumulus, and dry convective PBLs. To properly address this final issue, a variable such as turbulent kinetic energy (TKE) needs to be used, but TKE as a prognostic variable is rarely used in global models and as such is not available.

An estimate of the PBL height based on the Richardson number (referred to as boundary layer height) is, however, provided in the ERA-Interim data (PBLRI). It is only available at 6-h and 12-h forecasts, but it is likely very similar to the analysis period since forecasting skills over these short periods are well developed. Essentially it is based on the level where the bulk Richardson number reaches a critical value of 0.25, using the difference between this level and the lowest level as an estimator for the vertical stability. But, as stated, since the stability estimation is based on dry thermodynamic variables (and not on moist conserved variables) there is a tendency (see below) for this method to provide an estimate of the PBL height that is often closer to the cloud-base height in marine cloudy boundary layers, rather than the PBL height itself (Janssen and Bidlot 2003).

Figure 1 shows these different PBL heights along a transect from the coast of California to the equator, using only data from JJA 2000. This transect was introduced by the Global Energy and Water Cycle Experiment/Working Group on Numerical Experimentation (GEWEX/WGNE) Pacific Cross-Section Intercomparison (GPCI) to study important physical regimes and transitions in models and observations, for example, stratocumulus, shallow cumulus, and deep convection (Teixeira et al. 2011). The transect runs from 18S, 173W to 358N, 125W. In general, ERA-Interim data are used, although PBLRI also uses the ECMWF dataset ERA-40 for illustration of improvements.

Most PBL height definitions show fairly similar results for the transition from stratocumulus to cumulus in the area from 30°N to 20°N, although the PBLRI and PBLN do not grow up to a mean value of about 1.6 km near 20°N as in other estimates. More symptomatic of a potential problem with these two methods, in addition to the PBLBP one, is the significant decrease in PBL height south of 20°N, where the mean values of PBL height become about 600 m around 7°N. The refractivity Na entering PBLN is dominated by the water vapor pressure at lower altitudes and decreases rapidly with height (von Engeln et al. 2007). Thus a method based on a gradient search in a refractivity profile, or the breakpoint method, can be dominated by water vapor pressure variations at lower altitudes when a strong
gradient is found close to the warm ocean surface [this could also affect the PBL height determination using radio occultation data; however, it is likely too close to the surface for this measurement technique to be affected; see also Ao et al. (2012)]. This problem also affects the method based on $q$, not excluding the lowest 50 m in the gradient search leads to much lower values for the methods based on $N$ and $q$ (the breakpoint method automatically removes this lowest range).

The method based on a scaled specific humidity ($\text{PBL}_{sq}$) shows significantly larger PBL heights closer to the intertropical convergence zone (ITCZ) where more water vapor is present at the surface and, because of deep convection, larger values of water vapor can be found deeper in the troposphere above the PBL. Although the boundary layer is not as well defined in its vertical structure in the deep convection regions as it is in the subtropics, a more detailed investigation also shows that the RH-based method works better than the scaled humidity around the ITCZ (see discussion below). More generally, the $\text{PBL}_{sq}$ method also results in unrealistically large PBL heights at polar latitudes where little water vapor is present.

These generally high PBL heights are also found for the respective hemisphere wintertime over polar regions for $\text{PBL}_{N}$; it does, though, produce fairly realistic high values over desert regions. Methods based on virtual or potential temperature do not take into account the full cloud moist vertical structure of the PBL and also show higher PBL heights than, for example, the RH-based one over the transect.

The two reanalysis datasets ERA-40 and ERA-Interim show similar results when using $\text{PBL}_{\text{RH}}$, with a boundary layer growing from values around 700 m in the strato-cumulus regions close to the California coast to relatively steady values of around 1.5 km for latitudes below 20°N.

The $\text{PBL}_{\text{RI}}$ is always found below the $\text{PBL}_{\text{RH}}$ method. Similar results were also already reported by other researchers (e.g., Palm et al. 2005). The ECMWF method to estimate $\text{PBL}_{\text{RI}}$ can be considered more akin to an estimate of PBL cloud-base height in cloudy boundary layers, since it uses a temperature variable that is not conserved in phase transitions. The differences seen in Fig. 1 are consistent with this fact (climatological values of cloud base around 500 m are closely related to RH near the surface and in the east Pacific subtropics) and provide an approximate measure of cloud depth (since $\text{PBL}_{\text{RH}}$ captures the cloud-top height). In this context, it is interesting to note the realistic growth of this difference (cloud depth) from the shallow stratocumulus-topped PBLs toward the deeper trade cumulus and deep convection PBLs.

All other seasons actually do not show such a good agreement between 30°N and 20°N, instead showing more pronounced differences between the investigated methods with the DJF season having the widest spread, which is related to the prevalence of midlatitude systems. MAM and SON show fairly different behavior even among the two RH-based methods, in particular south of 10°N (MAM) or between 5° and 17°N (SON), while DJF shows very good agreement. Generally, the ECMWF-provided PBL height $\text{PBL}_{\text{RI}}$ is always found at the lower end, while the $\text{PBL}_N$ and $\text{PBL}_{sq}$ are always found very low toward the equator within the ITCZ.

It is difficult to fully evaluate these results since there are no reliable independent observations of the PBL height from a large-scale perspective. In addition, climate and weather prediction models still produce fairly divergent results in terms of cloud-top and PBL height. Teixeira et al. (2011) analyzed over 20 weather and climate prediction models and showed, for example, that cloud-top and PBL height model results at 20°N in the stratocumulus-to-cumulus transition region along the GPCI transect can basically be anywhere between 1 and 3 km. Nevertheless, there have been some observational studies that hint at results similar to the RH-based method. Von Engeln et al. (2005) used GPS radio occultation (RO) data along the GPCI transect to show a PBL height growing from around 1 km in the strato-cumulus regions to values between 1.5 and 2 km south of 20°N. Karlsson et al. (2010) analyzed MISR observations, which provide a fairly accurate estimate of boundary layer cloud-top height (which for strato-cumulus and transition to cumulus is a good measure of the PBL height). Along the GPCI transect, they show a PBL growing from values between 500 m and 1 km close to the coast to values around 1.5 km at 15°N.

To investigate the possibility of other indirect (but still physically based) predictors for PBL height, the correlations between the ERA-Interim $\text{PBL}_{\text{RH}}$ for JJA with additional parameters such as cloud cover and total water vapor are shown in Fig. 2 for points along the GPCI transect. Within the ECMWF model, low, medium, and high cloud covers are defined at altitudes where the current to surface pressure ratio ($p_R$) is $1.0 > p_R > 0.8, 0.8 \geq p_R > 0.45$, and 0.45 $\geq p_R$.

The most consistent feature in this figure is the negative correlation between low cloud cover and PBL height in the transition region between about 20° and 30°N. The total cloud cover correlation is very similar to the low cloud one here since low clouds are the dominating type. The negative correlation is consistent with the notion that as the boundary layer deepens when moving south, cloud cover diminishes as a result of the transition from stratocumulus regimes to cumulus
regimes. Note also the fairly drastic change in low cloud cover correlation from values around 0.3 close to the coast to values of about −0.5 at 30°N (and farther south), illustrating the radically different behavior of boundary layer growth in the “pure” stratocumulus and the transition region. These features can also be found in correlations of, for example, PBLTv and PBLTp.

In the transition, deeper boundary layers are often due to increasing shallow convection activity (versus well-mixed stratocumulus clouds) and as such are associated with less low cloud cover. In the stratocumulus regime, deeper boundary layers are associated with a deeper well-mixed cloud layer, larger values of integrated liquid water, and, as illustrated in Fig. 2, larger values of low cloud cover. There is also a fairly large and consistent positive correlation with medium cloud cover in the transition and cumulus region between 15° and 25°N, which is associated with shallow and midlevel convection activity; larger values of shallow and midlevel convection associated with deeper boundary layers lead to larger values of clouds in the midtroposphere. Interestingly, the total water vapor (TWV), which is a measure of the vertically integrated water vapor content of the atmosphere and is often directly associated with the water vapor content in the PBL, does not show a consistent pattern of correlation, oscillating between positive and negative values across the GPCI transect.

Correlations of clouds with, for example, PBLTv and PBLTp are generally similar, in particular at latitudes >15°–20°N; farther south they tend to show a lower positive correlation with medium clouds and a slightly negative correlation with low clouds.

We mentioned above that the determination of a PBL height around/within the ITCZ region with strong convection can be difficult. There are very few investigations into the PBL development across the ITCZ. Here, we briefly compare our results against an investigation that was performed with a focus on the ITCZ region around 95°W, following the analysis of the fall seasons for the years 1999, 2000, and 2001 in Pyatt et al. (2005, hereafter Pyatt05), using data from research cruises. Despite the fact that research cruises provide very localized pieces of information and do not fully cover the investigated times here (we use the SON season), the PBLRH structure found is very similar to, for example, the sounding data shown in Fig. 14 of Pyatt05.

The PBLRH is slightly decreasing when moving north for latitudes around 8° to 4°S with values around 1.2 km (for comparison, Pyatt05 finds values around 1.1 km). A more rapid decrease is found for latitudes 4°S to 0°N with a PBL height of around 0.75 km at the minimum (Pyatt05 finds the minimum slightly more to the north). Farther north, a rapid increase up to about 2.5 km is found in our data, the maximum being around 6°N (more closely following the satellite data in Fig. 14 of Pyatt05 in this region), followed by a rapid decrease to a minimum of about 0.6 km around 12° to 15°N (results in Pyatt05 are found even lower). In contrast, the PBLsq data are only similar at the very southern end of the investigated region, moving to altitudes around 2.7 km at 2°N and showing only a minor altitude reduction farther north. The two temperature-based methods actually show a similar behavior as the RH one with a fairly flat PBL height south of the equator (albeit about 150 m higher), an increase north of the equator to about 2.5 km (PBLtp) and about 2.0 km (PBLtv), and a less rapid decrease around 7° to 8°N. Note that, PBLsq actually also shows a similar behavior across the 95°W longitude but with generally lower values than the temperature or RH-based methods.

To summarize the main findings of this section: some presented methods for determining PBL height have difficulties in fully capturing the PBL height, for example, they 1) tend to be dominated by lower altitude water vapor gradients (PBLsq, PBLtv, PBLtp), 2) are too simplistic and do not work at all latitudes (PBLsq), or 3) are close to cloud-base height (PBLri). Given this analysis, the following results for a global PBL height climatology are based on PBLRH, PBLtv, PBLtp, and use of ERA-Interim data since this new reanalysis appears to have some advantages over the older ERA-40 data (Simmons et al. 2007). In particular, ERA-Interim has shown significant improvements of the simulations of marine cloudy boundary layers due to the implementation of the eddy-diffusivity mass flux (EDMF) approach (e.g., Soares et al. 2004; Köhler 2005; Siebesma et al. 2007). Other possible PBL height estimators, such

---

**Fig. 2. ERA-Interim PBLRH JJA correlation with several other model parameters over all analysis times.**
as cloud cover, were shown to be less effective in finding the PBL height.

4. Data validation

The quality of the ERA-Interim reanalysis fields was already shown in Simmons et al. (2010) in an analysis of low-frequency variations in surface atmospheric humidity, temperature, and precipitation. Within the presented work, the quality of the ECMWF fields to reproduce gradients in the vertical profile is first tested by comparing them to high-resolution AWI radiosonde data. The latitudinal coverage is sufficient to allow a meridional analysis of the atmosphere. These sondes are available for assimilation into the ECMWF weather prediction system and thus do not provide a fully independent validation. However, such radiosondes represent only a fairly localized piece of information in a data assimilation system and the forecasting skills of the ECMWF model are actually good enough to reproduce the observed gradients even without the sonde information available.

One can actually test the radiosonde impact on the model-based PBL height estimates by comparing the here used ERA-Interim reanalysis data with the forecast data provided by this reanalysis, where the sonde data have not been used. These are available from 0000 and 1200 UTC in 6-h and 12-h steps. We performed this comparison for 1 year only; results confirm that the analysis and the forecast-based PBL heights are similar, thus PBL heights differ only by a few meters. This limited impact of the sonde data can be understood by considering that the height estimates used here do not rely on accurate temperature or humidity information (which is provided by the sonde data as, e.g., shown in Benjamin et al. 2010), but only on the gradient of these profiles, which is less variable over the investigated forecast range. In addition, the ocean-based PBL is strongly constrained by the sea surface temperature while over land the PBL development is much more complex and thus the sondes impact will be higher.

The AWI dataset was also used previously in a similar study where the focus was on the production of a ducting climatology from ERA-40 fields (von Engeln and Teixeira 2004). Results obtained in that study confirmed the high reliability of this earlier ECMWF reanalysis for this type of application and have also led to the generation of a ducting climatology for use with ground-based weather radars (Lopez 2009). Here we perform a validation of the calculated ERA-Interim PBL height against the AWI dataset. We derive three different PBL heights from the ERA-Interim fields: PBLRH, PBLTP, and PBLTV. Collocated ECMWF observations are generated at the nearest analysis hour (0000, 0600, 1200, and 1800 UTC), interpolated to the corresponding latitude, longitude position. As mentioned above, the lowest 50 m of the AWI and the ECMWF profile are not used in order to avoid surface inversions.

The initial validation of PBL heights from the AWI dataset against ECMWF-based ones shows that all three investigated methods show heights in the AWI dataset that are a few hundred meters higher caused by gradients/strong variability in the sonde data at higher altitudes. A further restriction on the AWI sondes to generate a PBLRH dataset is not to use the altitude of the minimal gradient but further screen first for gradients in RH that are $\leq -100$ % km$^{-1}$, which removes about 6% of the sondes from this validation mostly at higher latitudes, and use the altitude of the lowest of these gradients instead. The mean gradient in RH for ECMWF data is also about $-100$ % km$^{-1}$ at the PBLRH altitude with slightly lower values over sea than over land. This PBLRH reduces to meridional mean results as shown in Fig. 3. The monthly-mean-based altitude is also included for comparison. Results have only been binned by latitude band since the AWI sampling locations do not really allow zonal gradients to be assessed.

For the collocated and monthly-mean results, agreement is generally good for all latitude bands in the PBLRH and PBLRH heights, although AWI data show generally slightly lower values except near the equator. Correlations between AWI and the collocated ECMWF data are above 0.8. Deviations are similar for the monthly-mean dataset, even though AWI sondes are generally launched at specific times. The difference between this ECMWF data and AWI is more evident in the correlation, which is here only 0.6. Note that a linear interpolation in time between the ERA-Interim 6-h time step yields similar results. The standard deviations in PBL height also agree well with the ones of the collocated ERA data. The same is not true, as would be expected, for the monthly-mean data with values between 0.2 and 0.4 km as opposed to values of up to 1.1 km in the observations.

It is also noted that agreement in PBL and PBLTV with the relevant ECMWF values is not as good as for PBLRH, where AWI data are generally a few hundred meters higher. Thus further refinements of the validation are also required for these parameters, similar to the gradient criteria used in the sondes relative humidity validation.

Performing the same validation exercise with ERA-40 illustrates several improvements of ERA-Interim over the older reanalysis, as is also shown in Fig. 3. Agreement between the PBLRH and PBLRH altitudes is
improved in particular at midlatitudes; however, around the equator ERA-40 captures the strong gradient from about 5° to 15°N slightly better. Variability (as seen here in standard deviation) is also improved when using collocated ERA-Interim data; note that the ERA-40 data cover only the years 1990 to 2001, hence a larger standard deviation is expected as a result of the smaller sample size. The overall better performance of ERA-Interim is, however, also found when restricting all datasets to the ERA-40 coverage.

The AWI dataset covers ocean observations. A validation that also includes the temporal development of the PBL height from ERA-Interim data over land is presented in Fig. 4, showing a selection of radiosonde monthly-mean data. We use monthly-mean data at specific UTC times over the years 2000 to 2009, primarily focusing on local noon when the PBL is well developed, again applying the lowest gradient selection PBLRH LG when identifying the PBL height in sonde data.

The figure clearly shows strength and limitations of these validations over land areas, where the PBL is much more variable. In particular, at high latitudes (McMurdo station) ECMWF PBL heights are significantly lower during the winter season than for sonde-based data. These low values are actually expected for these latitudes where the PBL height can even be below 50 m (Handorf et al. 1999) occasionally. The very low humidity will likely lead to problems for this PBL method using sondes; it is, however, unclear why sondes are also affected for temperature-based methods. The comparison for the other sonde stations shows generally good agreement between the ECMWF and sonde-based relative humidity height; this good agreement is nevertheless not found for the sonde-based PBLTP and PBLTv methods, which can either be higher (thus requiring a further selection akin to the cutoff used in relative humidity) or lower (thus requiring the exclusion of higher surface-based inversions). It should also be noted that the validation is generally closest to midday with a well-developed PBL; at night the residual layer can lead to problems in the detection of the PBL height, both for the sonde and the ECMWF data. This might also explain the deviations in PBL height found, for example, for Denver very early or late in the year when days are short and residual layers can impact the sonde data.

The ECMWF data only provide a 6-h resolution of the diurnal development, which can lead to over- or underestimation of average PBL heights, in particular over land where the PBL is much more variable. This sampling error depends on the actual local time of the observation. We estimate it based on PBL height data from Liu and Liang (2010), collected at the Atmospheric Radiation Measurement Program (ARM) Southern Great Plains site. The higher temporal sampling of this dataset allows a time resolution of 3 h and we find that the 6-h ECMWF data can lead to over- or underestimation in PBL height of up to 6%.

5. PBL height climatology results

Two different PBL height analyses are performed below. First, the data are averaged over all years to determine the PBL climatology. This is further separated into seasonal and diurnal variations. The second part looks at the robustness of the climatology by analyzing the percentage of outliers. The primary focus is on the PBLRH method, and differences with PBLTP and PBLTv are discussed in the text.

a. PBL height climatology

Figure 5 shows the mean PBLRH height for the four different seasons. From an annual mean climatology
perspective, a few features stand out. The PBL height clearly increases from the subtropical west coast of continents (e.g., California) with typical values between 500 m and 1 km toward the equator along the trade winds, reaching typical values close to 2 km in the shallow cumulus and deep tropical regions. These results naturally confirm previous studies of the evolution of PBL and cloud-top height in this transition region (e.g., von Engeln et al. 2005; Karlsson et al. 2010; Teixeira et al. 2011). The highest values of PBL height are reached in summer over dry subtropical land regions with values larger than 3 km. The lowest mean values of PBL height are reached in the polar regions during local winter when strong inversions close to the surface are fairly common.

Over the oceans the seasonal cycle is not as pronounced as over land, with the subtropical regions
(particularly off the west coast of continents) exhibiting the clear transition discussed above from a shallow PBL over the stratocumulus regions to a deeper cumulus PBL downwind the trades. In the ITCZ, the regions with larger values of PBL height are in JJA over the Atlantic and generally over the eastern Pacific. Over the oceans there are, however, some regions with a significant seasonal cycle, in particular off the northeast coast of continents, with larger values during DJF, presumably associated with baroclinic systems. During Northern Hemisphere (NH) summer these regions off the east coast of Asia and North America are dominated by fog (Klein and Hartmann 1993; Teixeira 1999b) and as such have fairly low values of PBL height. There are also regions that show more seasonal variability, for example, over the Mediterranean Sea, where the summer season shows lower PBL values possibly due to less contrast between air and sea temperatures [this pattern is also found in Seidel et al. (2012)].

The seasonal cycle is fairly obvious over land for regions such as North Africa, the Middle East, North America, and Australia, exhibiting a very marked difference between summer and winter. Mountainous regions over North America or Asia also show a strong seasonal cycle in PBL height with lower values in winter and higher ones in summer. There are also some intriguing results such as the low values over the tropical forests, which will be further discussed below.

The PBL$_{TP}$ and PBL$_{TV}$ methods are less variable over polar areas, where, generally, lower heights are found. But they also do not produce seasonally as high values over desert areas as the PBL$_{RH}$ method. Most land areas show, in fact, higher values for the RH-based method; tropical rain forest areas are, however, very similar for all three methods. Over ocean regions, the temperature-based methods are typically higher for low to midlatitudes over the different seasons, generally also around the South Pacific convergence zone (SPCZ) and the ITCZ (as is also visible in Fig. 1). The PBL$_{TV}$ is closer to the PBL$_{RH}$ method over oceans for low to mid-latitudes since both methods are affected by humidity variations; this is also visible in Fig. 1. In the ITCZ and SPCZ regions, it often shows even smaller values than the RH method.

Figure 6 shows the PBL$_{RH}$ height for the four different analysis times. It is quite significant how different the land and ocean are in terms of the diurnal cycle of PBL height, with almost no discernible diurnal cycle over the ocean. Over land, as expected, the diurnal cycle is intense, with large values of PBL$_{RH}$ height over Asia and Australia at 0600 UTC, Africa and the Middle East at 1200 UTC, and the Americas at 0000 and 1800 UTC. Hence, these large values are found at very similar local times of around 12 to 17 h. Some interesting features include the fact that for dry convective PBL regions like North Africa, the PBL can be fairly deep even several hours after the maximum of solar insulation, around 1800 local time. There is also a significant diurnal cycle over the tropical forests with values above 1 km at local noon (which corresponds to 1200 UTC over the African forest and around 1800 UTC for the Amazon forest) and clearly below 1 km for the rest of the day. The diurnal
mean of these values leads to the relatively low values of PBL\textsubscript{RH} height over the tropical forests as mentioned above when discussing Fig. 5.

The differences among PBL\textsubscript{RH}, PBL\textsubscript{TP}, and PBL\textsubscript{TV} are similar for all times over the oceans, as expected. Over land, PBL\textsubscript{TP} shows primarily higher values during local noon (e.g., around 1200 UTC over Africa and at 1800 UTC over South America) and lower ones during the rest of the day, except for high latitudes where it is generally lower. Both temperature-based methods also show a much more rapid decline of the PBL over land in the early evening, hence the RH-based method shows for Africa again higher values at 1800 UTC.

Figure 7 shows the PBL\textsubscript{RH} standard deviations for the four different seasons as a measure of the variability. It shows on the one hand the profound differences between the subtropical regions over the ocean (stratocumulus and transition to cumulus) with very low

![Fig. 6. Derived PBL\textsubscript{RH} height for different analysis times. Averaged over all years.](image)

![Fig. 7. Standard deviation of PBL\textsubscript{RH} height for different seasons. Averaged over all years.](image)
values and, on the other hand, the tropical deep convection and midlatitude regions over the ocean and much of the land with large values. Over the dry land subtropical areas, in particular during local spring and summer, values can easily reach about 2 km. The large values in the ITCZ region of the east Pacific (as compared to the warm pool) are presumably associated with a much larger variability in terms of cloud regimes in these regions.

The larger variability in the midlatitude regions, particularly over the southern oceans, is presumably due to the heterogeneous nature of baroclinic frontal systems. In the subtropical land regions, the variability is associated with the large values that the PBL height can attain there in spring and summer (due to the very large values of sensible heat flux) and the large diurnal variability.

Figure 8 shows the standard deviation as a measure of the variability of the PBLRH height for the four different analysis times. The diurnal cycle of PBL height standard deviation is clear. The lowest values are again in the stratocumulus and cumulus regions in the subtropical oceans and over most of Antarctica throughout the day. The highest values are over the subtropical land regions. For local nights, this is likely associated with the fact that the method mostly detects the strongest PBL inversion close to the surface, while other times the method may detect the height of the residual layer, which has remained from the previous day’s well-mixed convective layer.

The PBLTP and PBLTV standard deviations over the ocean are generally similar to the PBLRH ones, since the ocean variability is low. There are, however, areas that show higher standard deviation values for the temperature-based methods than for the RH one, in particular over the western Pacific warm pool, the northern Indian Ocean, and around the ITCZ region. Generally, the RH method shows higher standard deviations over land, but this cannot be fully explained by the residual layer since it is found for most UTC times, albeit the smallest differences are found around local noon. Over the tropical forests, the RH-based method actually shows less variability during local noon. High latitudes are found to be less variable in the temperature-based PBL height estimates. The main difference between the standard deviations of the two temperature-based methods is found at the ITCZ where PBLTP shows higher values.

The strength of the PBL inversion for the PBLRH height is shown in Fig. 9 for the different seasons. It is calculated for the year 2000 only. It illustrates well the key differences between the deep tropics and the subtropical regions over the oceans in terms of inversion strength. Over the subtropics a transition from strong inversions over stratocumulus-topped boundary layers (with values below \(-200\ \text{% km}^{-1}\)) to weaker inversions over the shallow convection regions is clear. In regions where deep convection occurs, the PBL inversion is fairly weak reaching values close to 0, which means that a clear PBL inversion in the deep tropics is not always present. Over the subtropical deserts where dry convection is prevalent, particularly during spring and summer, the inversion is fairly weak as well because of
a low (e.g., compared with the stratocumulus regions for example) contrast between the moister and colder (in terms of potential temperature) PBL and the drier and warmer free troposphere.

b. PBL height robustness

As mentioned above, we use the Tukey's biweight method to calculate mean and standard deviations at a specific grid point per time. This method, in addition, gives information on the percentage of data falling into the $\pm 2\sigma$ interval, which for an ideal Gauss curve would be about 95%. This percentage or weight of the distribution can thus be used to investigate the robustness of the derived climatology. Actually, the application of a robust estimator should always include an investigation into the weights since two very different distributions can lead to the same mean and standard deviation with such a method.

Figures 10 and 11 show the percentage of data points, averaged over all years, entering the statistical calculation for the PBL RH data. Both figures show large variability (thus lower weights) for Greenland and Antarctica, in particular at/near the ice–sea border, areas with higher precipitation. This high variability over Antarctica is generally caused by single peaks in the PBLRH time series, but there are also persistent features such as frequently found higher PBL heights in the summer months over Antarctica, during December and January. Another persistent feature is the lower weight near the west coasts of Africa and South America on the Southern Hemisphere. A more detailed investigation of the variability around 10$^\circ$S and 90$^\circ$W shows the influence of the ITCZ early in the year with higher PBL values of around 1.8 to 2 km, followed by higher variability/lower weights for the later seasons for all UTC times. A bit farther south, there is more variability throughout the year and no clear seasonal pattern.

These figures also show that over land, the PBLRH is more stable during local noon and early evening, and outliers are more frequently detected for local midnight, likely caused by the collapse of the PBL and the false detection of the residual layer. In addition, particular mountainous regions, for example, the Rockies in the United States, show larger variability during the night and early morning (0600 and 1200 UTC), developing into a stable PBL over the day (1800 and 0000 UTC). The patterns over mountainous regions can be observed for all seasons, hence the reduced weight visible in Fig. 10. The ITCZ region shows lower weights in particular during the MAM season.

The results look generally very similar for the PBL$_{TV}$ and PBL$_{TP}$ methods, although a more detailed look shows that the PBL$_{RH}$ is more robust over land during local noon, while the temperature-based methods are found to be more robust during local night and early morning (likely better in detecting the actual PBL height than the residual layer). At high latitudes over sea, the temperature-based methods have a higher weight and also show lower PBL height values. Seasonally, the RH-based method often shows higher weights over land, however, some coastal areas over Antarctica and Greenland show higher weights for the temperature-based methods. The difference between the two temperature-based methods is very small over land;
over the ocean the PBL$_{TP}$ method has generally higher weights, in particular over the western Pacific warm pool and the ITCZ.

6. Conclusions

We generate a PBL height climatology from ECMWF ERA-Interim reanalysis data over the years 1990 to 2009. This reanalysis series covers already standard satellite observations from space for the earlier years, as well as all the generally available ground-based observations. Several methodologies to estimate PBL height from this reanalysis are discussed and tested. An analysis of results from the different methods along a transect in the Pacific Ocean from California to the ITCZ that encompasses stratocumulus, cumulus, and deep convection regimes and the transitions between them, highlights some potential problems with methods associated

---

**Fig. 10.** Weight of PBL$_{RH}$ height for different seasons.

**Fig. 11.** Weight of PBL$_{RH}$ height for different analysis times.
with the vertical gradients of specific humidity and re-
fractivity and with a provided PBL height from the
ERA-Interim dataset (which is more akin to cloud-base
height). Overall the methods associated with the iden-
tification of vertical gradients of either relative hu-
midity PBLRH, potential temperature PBLTp, or virtual
temperature PBLTv, appear to be the most robust. Within
our analysis, we focus on PBLRH heights and outline
differences found when using the temperature-based
methods.

A detailed evaluation of these three methods to de-
rive PBL heights from ERA-Interim data is performed
against sea- and land-based radiosondes. These results
support the conclusion that in general these heights are
fairly close to the radiosonde observations, not only in
the mean sense but also in terms of variability. There is,
however, a need to further filter the sonde-based PBL
heights to minimize the impact of this high-resolution
dataset. Within this analysis, we use a simple gradient-
based method for PBLRH only. The temperature-based
methods in addition generally need a screening of low
altitudes, in order to avoid picking up surface inver-
sions. Depending on how often these particular radio-
sondes are assimilated into the ECMWF system, this
may actually be an interesting result about the capa-
bility of a model/analysis system to achieve a realistic
balance between large-scale and boundary turbulence
motions, which control the height of the PBL.

A detailed analysis of the global-mean PBL height
climatology (based on the relative humidity gradient
method) and its seasonal and diurnal variations illus-
trates significant differences between land and sea. Over
the global oceans there appears to be virtually no di-
urnal cycle in PBL height and a fairly modest seasonal
cycle when compared to land. The most striking fea-
tures over the oceans are the fairly persistent patterns
of transition along the trade winds from low values of
PBL height off the west coast of continents to much
larger values deeper in the tropics. An interesting sea-
sonal pattern is clear off the east coast of continents (Asia
and North America) with relatively large values of PBL
height during NH winter associated with cold air out-
breaks and frontal systems and low values during NH
summer associated with fog. Over land there are clear
seasonal and diurnal cycles of PBL height with large
values during local summer and local daytime.

The PBLTv and PBLTp climatologies are similar; how-
ever, they are less noisy and generally lower over polar
regions since relative humidity is very low in these areas.
However, they do not produce such high values over
desert areas as the PBLRH method.

An analysis of the global PBLRH height standard de-
viation climatology illustrates the essential differences
between the subtropical oceans and the deep tropical
regions. Over the subtropical regions, dominated by the
cloudy boundary layer where stratocumulus and cumu-
lus clouds are prevalent, there are low values of PBL
height standard deviation. These low values are asso-
ciated with the persistent nature of these clouds (and
associated vertical thermodynamic structure) and the
transition from shallow stratocumulus regimes to deeper
cumulus regions. Large values over the subtropical dry
land are associated with strong diurnal cycles, oscilla-
tions between large values of PBL height during the
day and low values at night, and the remnants in the
free troposphere of dry convective boundary layers. An
analysis of the PBL inversion strength in terms of rel-
ative humidity vertical gradients and the PBL height
highlights again the fundamental differences between
tropical and subtropical oceans, with fairly large values
over the stratocumulus regions and low values over the
deep convection regions. The results for PBLTv and
PBLTp are similar; however, they again show less vari-
bility for polar regions.

An analysis of the robustness of the statistics—
measuring the percentage of outliers—shows that the
sea-based PBL is generally very stable over all in-
vested seasons and UTC times. Over land, the PBLRH
was found to be more robust during local noon, while in
particular during the night there is more scatter and thus
more outliers found. The PBLTv and PBLTp methods
tend to be more robust over land during the night, likely
less frequently picking up the residual layer.

The PBL height dataset derived in this study is
available for researchers upon request from the authors.

Acknowledgments. The authors thank C. Accadia,
J. Ackermann, P. Albert, T. Hewison (EUMETSAT,
Germany), and M. Schröder (DWD, Germany) for
helpful discussions, the AWI and the University of
Wyoming for access to the radiosonde data, and ECMWF
for providing the different data sets used in this study.
An easy to use storage and computing environment
for this project that started in 2002 was provided by
the SAT group. Special thanks go to O. Lemke and
S. Bühler (Lulea Technical University, Kiruna, Sweden).
JT acknowledges the support provided by the Office
of Naval Research, Marine Meteorology Program
under Awards N0001411IP20087 and N0001411IP20069,
the NASA MAP Program, and the NOAA/CPO MAPP
Program. Part of this research was carried out at the
Jet Propulsion Laboratory, California Institute of Tech-
ology, under a contract with the National Aeronautics
and Space Administration. Two reviewers (D. Seidel
and one anonymous) are also acknowledged for valu-
able inputs.