# Urban daytime boundary layer height determination and its relationship with high pollutant concentrations in Barcelona

Author: Miguel García Dalmau
Supervisor: Mireia Udina Sistach, mudina@ub.edu
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.**

**Abstract:** The planetary boundary layer height (PBLH) is a key parameter to diagnose the amount of pollutant concentration since it takes part in vertical dispersion, limiting the volume where the pollutants can be dispersed. Nevertheless, there is not a reference technique to compute it. In this work, 4 different algorithms -Parcel Method (RSD0), Gradient Detector (WW), Function Adjustment (ERES) and BL-View Enhanced Gradient using a Candidate Selector (CEIL)- are applied to Barcelona non-rain midday radiosounding and ceilometer data from August 2015 to December 2018 in order to estimate PBLH, focusing on pollutant episodes. The results suggest that RSD0 algorithm use is more suitable during summer, when boundary layer is highly developed. Instead, CEIL yield better results during colder seasons, when boundary layer is not highly developed and, thus, noise do not dominate over backscatter signal. WW is useful to detect strong temperature gradients and ERES is effective at detecting cloud presence. Among the four algorithms, WW and CEIL are selected to study high pollution concentration days. The PBLH median for $NO_2$ and $PM_{10}$ episodes is lower than the 25th percentile for all days studied, which suggest the role of low PBLH in pollutant accumulation. WW and CEIL point higher PBLH for tropospheric $O_3$ episodes influenced by high temperature and clear sky conditions.

## I. INTRODUCTION

Boundary layer (BL) is the lower atmospheric layer, from the surface up to few kilometres above it. BL main characteristic is being affected by the surface, which forces the atmosphere to react to derived effects (forcings) caused by its presence, such as sensible and latent heat transfer, terrain induced flow modification, frictional drag and pollutant emission, among others, in a short timescale of one hour or less (*Stull* 1988). Under the influence of high pressure large scale forcings and clear skies, over land, the BL shows a marked diurnal cycle: it starts developing vertically after sunrise, when convection begins to dominate over dissipative stresses, and reach its maximum height -usually between 10 m and 3000 m- before sunset. As night approaches, surface cools and develops a radiative surface inversion which avoids the turbulence processes and creates a stable nocturnal boundary layer. Above it, remains the residual layer: part of the daytime originated BL which maintains its characteristics.

Daytime BL is dominated by turbulence, usually driven by convection. The solar heating creates eddies (thermal turbulence) which produce mixing, so that meteorological variables tend to uniformity. In a ideal profile, BL shows a neutral stratification with constant potential temperature, constant water vapour content and constant pollutant concentration -except directly over surface, where strong temperature and humidity gradients are expected due to stronger turbulence. The entrainment zone is located at the top of BL: it is a transition zone between the BL and the free atmosphere characterized by a sharp potential temperature increase (entrainment inversion) and a water vapour decrease. It usually defines the top of the BL, except when clouds are present. In this case, if cloud height is lower than the entrainment zone, BL top is usually defined at the cloud base.

BL depth or BL height (hereafter PBLH) strongly influences the pollutant vertical dispersion. High pollutant concentrations are often related to high pressure large scale forcings leading to stability and low dispersion. In that situations a PBLH can be well defined. If PBLHs are low, the volume of air where the pollutants can be mixed is limited and its concentration tends to increase due to poor vertical dispersion. High pollution levels are linked to health problems causing globally more than 2 million premature deaths every year (*WHO* 2005). The main pollutants that affect large cities, and the standard air quality thresholds are usually exceeded, are nitrogen dioxide, particulate matter and troposferic ozone.

$NO_2$, a toxic gas at high concentrations originated from fuel combustion processes, have been shown to reduce lung function (*WHO* 2005). $PM_{10}$ are defined as any particulate matter -e.g. dust, which in Mediterranean areas may be transported from nearby North African deserts; soot, ashes, metal particles and pollen- with a diameter lower than 10 micrometers. Health effects are broad, mostly related to cardiovascular and respiratory systems (*WHO* 2005). Tropospheric $O_3$ is a secondary pollutant formed in complex chemical reactions with nitrogen oxides and volatile organic compounds in presence of high-intensity ultraviolet radiation. Its affectations range from agricultural losses (*Arya* 1999) to chronic effects in human health, specially those related to the respiratory system (*WHO* 2005). $O_3$ is not a pollutant that usually affects main cities, but may become a regional issue if precursors emitted in urban areas are transported.

* Electronic address: mgarcida8@alumnes.ub.edu

Because of its relation to vertical dispersion, PBLH is a key parameter to pollutant prediction and diagnose, especially for models. Nevertheless, there is not a unique approach to compute PBLH. Many authors have explored various calculation methods based on different atmospheric characteristics: *Holzworth* (1964), *Oke* (1988), *Seidel et al.* (2010), *Wang and Wang* (2014), among others, developed algorithms using radiosounding data to detect the entrainment zone or a elevated inversion. Ceilometers, originally used to detect cloud height, also serve to estimate PBLH since an ideal backscattering profile shows a decrease near the entrainment zone that it is able to be detected by this devices (*Vaisala* 2014). This is used, for instance, in *Eresmaa et al.* (2006), *Endlich et al.* (1979) and *de Haij et al.* (2007).

In this study, we explore the daytime PBLH in the urban area of Barcelona, and its relation with high pollution episodes. Two objectives can be defined: i) evaluate four different PBLH determination algorithms, two of them based on radiosoundings measurements: Parcel Method (*Holzworth* 1964), hereafter RSD0, and Gradient Detector (*Wang and Wang* 2014), hereafter WW; and the other two based on ceilometer data: Function Adjustment (*Eresmaa et al.* 2006), hereafter ERES, and Vaisala BL-View Enhanced Gradient (*Vaisala* 2010) applying Candidate Selector (*Montolio et al.* 2018), hereafter CEIL. The second objective is ii) study the relation between PBLH and high pollutant concentration episodes in the urban area of Barcelona.

## II. METHODOLOGY

### A. Study area

This study is centred in the Low Emission Zone *Rondes de Barcelona* (ZBE), which was set up in 2017 affecting partially or entirely the municipalities of Barcelona, Cornellà de Llobregat, Esplugues de Llobregat, l'Hospitalet de Llobregat and Sant Adrià de Besòs, in Catalonia region -NE Iberian Peninsula, see figure 1-. It affects a densely populated area -the five councils register a population over 2 million people in an area of roughly 129 $km^2$ (*IDESCAT* 2020)-, forming an urban continuum. The main purpose of the Low Emission Zone is to improve air quality in one of most vulnerable zones of the Catalan Special Protection Zone (ZPE), 40 municipalities around Barcelona city where legal pollutant thresholds are frequently exceeded. Some measures included the application of traffic restrictions in case of an activation of the environmental episode procedure due to high pollutant concentration. This restrictions became permanent to the most pollutant vehicles since January 2020 (*BCN-CC* 2020).

*Reial Decret* 102/2011 (*RD* 2011) fixes maximum pollutant concentration in Spain following 2004/107/EC and 2008/50/EC European Directives (*ED* 2019). Table I presents the legal thresholds set for $NO_2$, $PM_{10}$ and $O_3$.

XVPCA (*Xarxa de Vigilància i Protecció de la Contaminació Atmosfèrica*) is the public network responsible of pollution measurements in Catalonia. It operates under the supervision of the Catalan Territory and Sustainability Department. XVPCA station network counts, nowadays, with 79 automatic stations and 107 manual stations, distributed in the 15 Air Quality Zones (ZQA) in which Catalonia is divided, also shown in figure 1. Nowadays, all municipalities included in ZBE are part of ZQA 1 - Barcelona Area. The stations are classified according to urbanization degree -urban, suburban or rural- and to main pollution source -background, industrial or traffic- (*XVPCA a* 2014).

Barcelona Area usually exceeds legal pollutant thresholds (*Vidal et al.* 2019). From 2015 to 2018, $NO_2$ VLa was exceeded each year in the traffic measurement stations of Eixample and Gràcia, while $NO_2$ VLd and LLA were never reached, as well as $PM_{10}$ VLd. $PM_{10}$ annual mean was gradually reduced, and VLa had not been exceeded in any station since 2016. $O_3$ VOPS superations are registered in Palau Reial in 2015 and in Sant Adrià de Besòs during 2015 and 2018.

Table I. Concentration thresholds for $NO_2$, $PM_{10}$ and $O_3$ in Spain.

| Pollutant | Name | Parameter | Value |
|---|---|---|---|
| $NO_2$ | VLd (Human health protection hourly threshold) | Hourly mean | 200 $\mu g/m^3$ * |
| | LLA (Alert threshold) | Hourly mean | 400 $\mu g/m^3$ |
| | VLa (Human health protection annual threshold) | Annual mean | 40 $\mu g/m^3$ |
| $PM_{10}$ | VLd (Human health protection daily threshold) | Daily mean | 50 $\mu g/m^3$ ** |
| | VLa (Human health protection annual threshold) | Annual mean | 40 $\mu g/m^3$ |
| $O_3$ | VOPS (Human health protection value) | 8h moving average day maximum | 120 $\mu g/m^3$ *** |
| | LLI (Information threshold) | Hourly mean | 180 $\mu g/m^3$ |
| | LLA (Alert threshold) | Hourly mean | 240 $\mu g/m^3$ |

* no more than 18 times within a civil year.
** no more than 35 times within a civil year.
*** no more than a mean of 25 days for every civil year in a 3 year period.

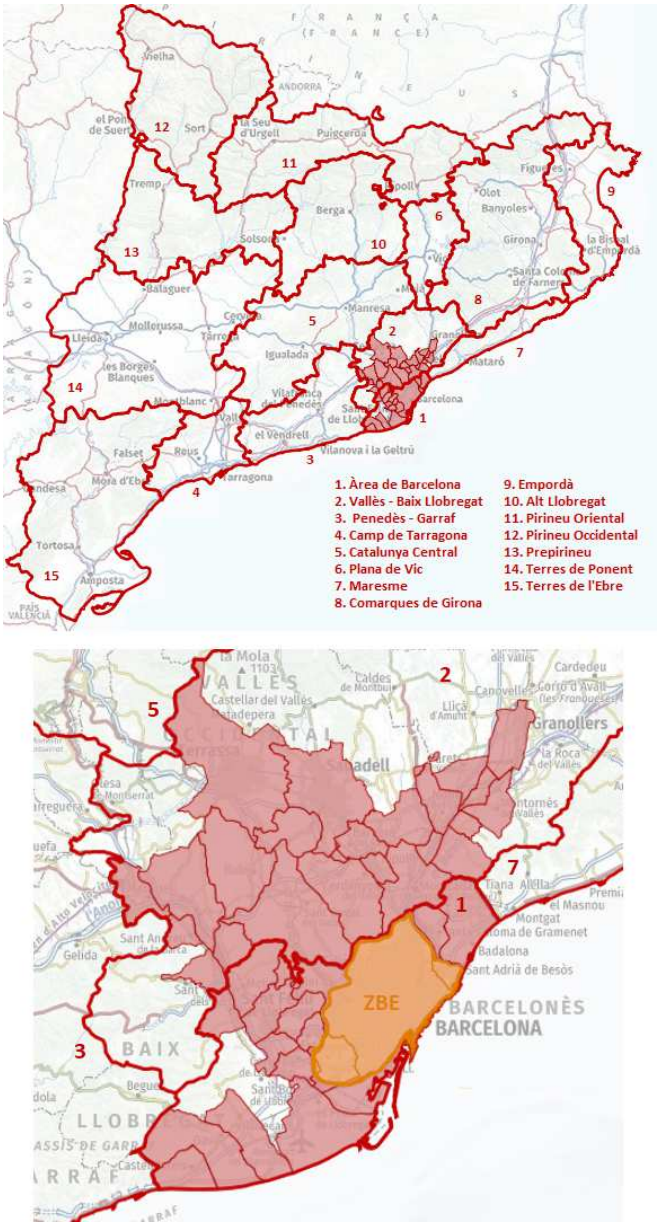


Figure 1. Catalan ZQA (red line), ZPE (red shade) and ZBE (orange shade).

Territory and Sustainability Department implements measures against atmospheric pollution. Until January 2020, two different procedures were applied for $PM_{10}$ and $NO_2$ episodes: preventive warning and environmental episode.

Preventive warning is the low pollution scenario and is the previous phase to an environmental episode. $PM_{10}$ preventive warning occurs when daily mean of the previous day is higher than 50 $\mu g/m^3$ in more than one ZPE station and the forecasting for the following day estimates a $PM_{10}$ daily threshold exceedance. $NO_2$ preventive warning is activated when hourly mean in more than one ZPE station is superior to 160 $\mu g/m^3$ and the 24-h forecast does not state a level decrease.

Environmental episode is the high pollution scenario and is activated when pollutant levels are high and there is an unfavourable prediction. For $PM_{10}$, two situations are considered: i) when daily mean of the previous day is higher than 80 $\mu g/m^3$ in more than one ZPE station and the forecast for the following day estimates a $PM_{10}$ daily 50 $\mu g/m^3$ threshold superations, or ii) more than consecutive 3 days with 50 $\mu g/m^3$ daily mean exceedance in more than one ZPE station and there is an unfavourable forecast. $NO_2$ environmental episode is activated when hourly mean in more than one ZPE station is superior to 200 $\mu g/m^3$ and the 24-h forecast does not signal a level decrease or when another justified circumstance requires it. $NO_2$ environmental episode declaration may lead to traffic restriction measures. Outside the ZPE, each case and measures applied are evaluated individually.

## B. Data

To develop this work, data from 4 sources are used: radiosounding, ceilometer, ZBE pollutant stations and data from a meteorological station.

Atmospheric radiosoundings (RSD) are instruments used to measure meteorological variables at different heights. They are equipped with temperature, humidity and pressure sensors, GPS position and time tracing to obtain wind, and a radio to transmit data to the ground station. The sonde is attached to an helium balloon which lifts it over 30 km height. They are launched around the world usually twice every day, reaching the top of the atmosphere at 00:00 UTC and 12:00 UTC, giving a global state of the vertical atmosphere.

Catalonia has one radiosounding station -model MODEM M10- run by the Catalan Meteorological Service and located at the Faculty of Physics of the University of Barcelona (41.385$^o$N, 2.118$^o$E) at 98 m above the sea level. Barcelona radiosounding is performed since 1998, and runs automatically since 2013. From December 2008, Barcelona radiosounding is integrated into the Global Meteorological Network (*SMC* 2017).

In this study, data from RSD launched at midday is used, considering the launching time at 11:10 UTC, from August 2015 to December 2018 -1248 days in total-.

Ceilometers are instruments used to estimate cloud base height and vertical visibility. Its functioning is based on sending a light pulse and measure the backscatter signal returned mainly by cloud droplets. Nevertheless, other particle contribution, like dust particles, is also reflected in the backscatter profile. This feature is used to compute PBLH (*Vaisala* 2018).

Ceilometer data has been obtained from a CL31 Vaisala Ceilometer set in the same location as the RSD. It is equipped with an InGaAs diode laser which sends 910 nm light pulses. It has a high measure range, from almost 0 km to 7.6 km, and a 10 m resolution. Data is reported every 16 s (*Vaisala* 2014). Barcelona Ceilometer data is available from August 2015, when it was installed.

Table II. ZBE pollutant stations (*XVPCA b* 2018). Classification is related to urbanization (U=urban, S=suburban, R=rural) and main pollution source (B=background, I=industrial and T=traffic).

| Station | Location | Address | Class | Automatic measurements |
|---|---|---|---|---|
| 1. Barcelona - Ciutadella | 41.386°N, 2.188°E | Parc de la Ciutadella | U/B | $NO_X$,$O_3$, CO, $SO_2$ |
| 2. Barcelona - Eixample | 41.385°N, 2.154°E | Av. Roma / c. Urgell | U/T | $NO_X$, $PM_{10}$, $O_3$, CO, $SO_2$ |
| 3. Barcelona - Gràcia - St. Gervasi | 41.399°N, 2.153°E | Pl. Gal·la Placídia | U/T | $NO_X$, $PM_{10}$, $O_3$, CO, $SO_2$ |
| 4. Barcelona - Observatori Fabra* | 41.418°N, 2.124°E | Observatori Fabra | S/B | $NO_X$, $O_3$ |
| 5. Barcelona - Palau Reial | 41.387°N, 2.116°E | Palau Reial de Pedralbes | U/B | $NO_X$, $PM_{10}$, $PM_{2.5}$, $O_3$, CO, $SO_2$ |
| 6. Barcelona - Poblenou | 41.404°N, 2.205°E | Pl. Doctor Trueta | U/B | $NO_X$, $PM_{10}$ |
| 7. Barcelona - Sants | 41.378°N, 2.133°E | Jardins de Can Mantega | U/B | $NO_X$ |
| 8. Barcelona - Vall d'Hebron | 41.426°N, 2.147°E | Parc de la Vall d'Hebron | U/B | $NO_X$, $PM_{10}$, $O_3$, CO, $SO_2$ |
| 9. L'Hospitalet de Llobregat | 41.370°N, 2.115°E | Av. Torrent Gornal | U/B | $NO_X$, $PM_{10}$ |
| 10. Sant Adrià de Besòs | 41.426°N, 2.222°E | c. Olímpic | U/I | $NO_X$, $PM_{10}$, $O_3$ |

* From 2018-03-09.

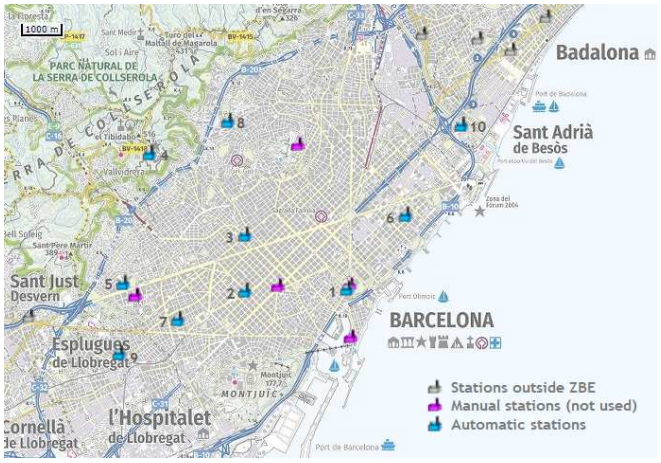


Figure 2. Pollutant stations in ZBE and surrounding area.

To collect, analyse and present backscattering data, CL31 Vaisala Ceilometer is equipped with Vaisala Boundary Layer View Software (BL-VIEW). BL-VIEW generates two files each day: "Level_2" files have backscattering information encrypted, and "Level_3" have the estimation of up to three planetary boundary layer and cloud heights using the Enhanced Gradient method (*Vaisala* 2010), which will be later explained in analysis methods section (section II C).

Pollutant data is obtained from XVPCA network. Hourly-mean automatic records are free public data, and can be downloaded from Open Data web-page (*Dades Obertes* 2020). In particular, this study uses XVPCA measurements from the 10 automatic stations located inside the ZBE (see Table II and figure 2). More specifically, pollution episodes are selected as cases when i) a legal threshold is exceeded or ii) a preventive warning or an environmental episode are declared.

To filter data depending on weather, meteorological data from the station located at Faculty of Physics of the University of Barcelona was used. It registers with a ten-minutes temporal resolution, temperature, relative humidity, maximum and mean wind force, wind direction mean and typical deviation, rain, pressure and short wave and long wave radiation. Of all available meteorological variables from the station only rain has been used, to distinguish between wet and dry days.

### C. Analysis methods

This study has initially used all available Barcelona radiosounding, ceilometer, pollutant and meteorological data from August 2015 to December 2018; that is 3 years and 5 months, 1248 days in total. Data is filtered using two criteria: i) days with precipitation between 6:00 UTC and 11:10 UTC are disregarded (93 days) to avoid possible effects due to the rain. ii) When there is a comparison between two algorithms, days when PBLH cannot be estimated either by one of the two algorithms compared are dismissed.

Next, the algorithms to calculate the PBLH from radiosounding and ceilometer are described.

#### *1. Determination of Radiosounding PBLH*

PBLH computed with radiosounding data is estimated using two different methodologies:

- Parcel Method (RSD0) (*Holzworth* 1964). Air parcels with a higher temperature than the environment are lifted until both temperatures are equal. RSD0 estimates PBLH for a convective dominated boundary layer as the height where air parcel evolution following a dry adiabatic evolution from surface intersects with the environment temperature profile. If the equilibrium temperature is located between two radiosounding heights, a linear adjustment is made.
- *Wang and Wang* (2014) "Gradient Detector" (hereafter WW), considering *Levi et al.* (2020) modifications. As the typical temperature and humidity profiles show sharp variations at BL top, PBLH can be identified as the height where maximum gradient of potential temperature ($\theta$) and

minima specific humidity ($q$), relative humidity ($U$) and refractivity ($N$) gradients are located. Specific humidity is calculated from RSD variables:

$$q \approx r = \frac{r_d}{r_v}\frac{e}{p-e} \quad (1)$$

where $r_d = 287.05\ Jkg^{-1}K^{-1}$ and $r_v = 461.50\ Jkg^{-1}K^{-1}$ are the specific constants for dry air and water vapour respectively, $e$ is the water vapour pressure, calculated using $e = U \cdot E$, where $E$ is the saturation water vapour pressure, computed from Magnus-Tetens expression, and $U$ is the relative humidity given by RSD data. Refractivity allows to consider joined effects of humidity, pressure and temperature, and it is calculated using:

$$N = 77.6\frac{p}{T} + 373000\frac{e}{T^2} \quad (2)$$

where $p$ is the pressure expressed in hPa and $T$ is the temperature in K. As they are needed for the computation, derivatives of temperature, refractivity and specific and relative humidity are obtained with a 4th order discrete approximation. Gradient peaks are located using an automated function from Scipy, a Python module, equipped with a quality filter to dismiss unreliable PBLH.

#### *2. Determination of Ceilometer PBLH*

PBLH determination using ceilometer data is performed using both ceilometer generated "Level_2" and "Level_3" files. In this work, 2 methods are considered:

- Function Adjustment (*Eresmaa et al.* 2006), hereafter ERES. In this method, 90 m height moving mean and 10-minutal mean backscattering profile is fitted to an idealized profile given by:

$$B(z) = \frac{B_m + B_u}{2} - \frac{B_m - B_u}{2}\ \mathrm{erf}(\frac{z-h}{\Delta h}) \quad (3)$$

where: $B_m$ is the mean backscattering in the BL, $B_u$ is the free atmosphere mean backscattering, $h$ is the PBLH, $\Delta h$, the entrainment zone width and erf is the Gauss error function. Variable adjustment is estimated using an automated Scipy function. In order to avoid noise dominance over backscattering signal, profiles are cut at 2250 m or when a strong response is detected, usually caused by clouds.

- BL-VIEW Enhanced Gradient method (*Vaisala* 2010) applying *Montolio et al.* (2018) algorithm, hereafter CEIL. BL-VIEW propose up to three candidates for PBLH indentified as the minima of the backscattering gradient, since the backscattering profile is characterized by a sharp decrease in the entrainment zone. Before the proper PBLH computation, the algorithm filters high backscatter values mainly caused by clouds and precipitation. Then, it averages data with a space range of 80 m (near the ground) to 360 m (over 1500 m) and with a time window between 14 and 52 minutes depending on the weather (*Vaisala* 2010) to minimize noise effects. BL-VIEW defines PBLH as the most pronounced minima. Instead of applying this criteria, *Montolio et al.* (2018) algorithm takes into consideration previous time evolution and analyse the most probable values to select PBLH.

All data obtained and PBLH methods are is statistically analysed in section III A. Days when there is an $NO_2$, a $PM_{10}$ or $O_3$ thresholds are exceeded, or when preventive warning or an environmental episode are activated, are identified and separately studied in section III B.

## III. RESULTS

### A. PBLH methods evaluation

Statistical data from PBLH obtained using all 4 methods is summarized in table III. We have to take into account that all days without rain are considered here, either clear sky or cloudy days. The mean PBLH in Barcelona from August 2015 to December 2018 is between 921 m (CEIL) to 1035 m (WW). Data availability is similar among RSD0, WW and CEIL algorithms, approximately 1100 cases, but it falls to 800 for ERES method. Table III also presents similar results for all 4 methods in the computation of the mean, 25th, 50th and 75th percentiles. Regarding 95th percentile, CEIL shows the lowest result (near 1500 m), and WW presents the highest one (over 1950 m) above all algorithms. This behaviour is also reflected in standard deviation, as CEIL computations produce the smallest value of all 4 methods, and WW exhibits the largest one.

To study possible time effects, data obtained is divided into the four seasons. In the results presented in figure 3, is it possible to see that RSD0 and CEIL compute low PBLH in winter and autumn, and slightly higher during spring and summer. This results are the expected ones, as forcings -such as convection- are stronger in warm seasons. On the other hand, WW and ERES do not perform this annual pattern: WW shows high PBLH during cold seasons and low in warm seasons, and ERES displays similar seasonal outcomes -medians varying from 850 m during autumn to 1000 m in winter-. Other remarkable results are the low 25th RSD0 winter percentile -lower than 500 m-, meaning that RSD0 calculates the lowest PBLH during winter. In contrast, spring and summer RSD0 75th percentile are the highest among all methods. For RSD0, are also remarkable the superior outliers during spring and summer, higher than 3000 m. ERES presents a lower whisker reaching almost 0 m, in contrast with the other algorithms.

Table III. Statistical rain-filtered data from PBLH computing algorithms

| Measurement | Valid data | Mean (m) | Std (m) | 5th (m) | 25th (m) | 50th (m) | 75th (m) | 95th (m) |
|---|---|---|---|---|---|---|---|---|
| RSD0 | 1101 | 961 | 462 | 247 | 631 | 921 | 1246 | 1778 |
| WW | 1109 | 1035 | 472 | 330 | 684 | 993 | 1264 | 1962 |
| ERES | 802 | 973 | 429 | 400 | 668 | 894 | 1215 | 1834 |
| CEIL | 1077 | 921 | 338 | 450 | 659 | 890 | 1143 | 1531 |

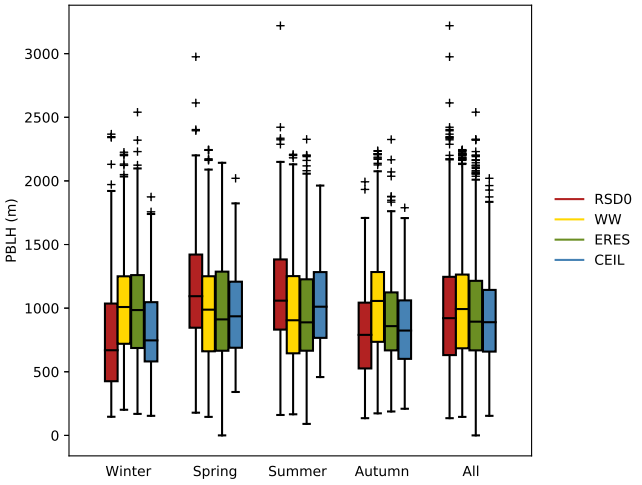


Figure 3. Boxplot of all 4 algorithms computed data labelled by station.

Figure 4 presents RSD0, WW and CEIL data in scatter plots identifying the season where the data was registered. ERES method has not been calculated due to the lack of confidence on the results as it will be explained at the end of this section. This figure is useful to confirm, in the first case, the seasonal percentile behaviour. There is a point accumulation for RSD0 in the section with values lower than 500 m during winter and, to a lesser extent, during autumn, which coincides with the results given by low RSD0 percentiles during this seasons. As explained before, WW may not detect the entrainment zone as PBLH, but other upper layer temperature inversions, disregarding the season. This can be seen by WW top values. Lastly, CEIL difficulties to detect very high PBLH are presented in the lack of points over 2000 m: just one value, measured in summer, exceeds this level.

In order to evaluate the correspondence between methodologies, figure 4 also presents the computation of a first order data adjustment and its correlation coefficient. The lines fitted equations are:

$$\mathrm{PBLH_{RSD0}} = (0.81 \pm 0.03)\mathrm{PBLH_{CEIL}} + (210 \pm 30), \quad r^2 = 0.348 \tag{4}$$

$$\mathrm{PBLH_{WW}} = (0.36 \pm 0.03)\mathrm{PBLH_{RSD0}} + (700 \pm 30), \quad r^2 = 0.121 \tag{5}$$

$$\mathrm{PBLH_{WW}} = (0.50 \pm 0.04)\mathrm{PBLH_{CEIL}} + (570 \pm 40), \quad r^2 = 0.127 \tag{6}$$

The results obtained do not allow to assure data follow a lineal pattern for any of the three plots. Yet, the slope -closer to 1 $m^{-1}$- and the $r^2$ coefficient obtained between RSD0 and CEIL (equation 4) suggest a better adjustment for this two algorithms. Although the R-squared results, there is an accumulation in the zone near the 1:1 curve in all of the subfigures. From all 1056 points available in the comparison between CEIL and RSD0, 579 of them (55%) show a difference between methods no larger than 200 m, while this value is reduced for WW-RSD0 and WW-CEIL scatter plots, 491 from 1088 data points (45%) and 495 from 1064 (47%) respectively. Even though, the adjustment continues being poor, influenced by non-compensated points due to non-ideal boundary layer conditions.

From this information and the study of some individual cases, we can comment about the behaviour and limitations of the different algorithms.

- RSD0 method PBLH results are influenced mainly by three different factors. i) Lower layer RSD data. RSD0 algorithm highly depends on surface temperature and its profile at low levels: the simulated adiabatic evolution can intersect the temperature profile in the first meters and compute a too low PBLH, as top figure 5 shows. Additionally, when surface potential temperature is the profile minima, this methodology does not give any valid result (for instance, with strong surface temperature inversions). This behaviour may explain lower values during the colder seasons using this algorithm, which explains the low 25th PBLH percentile in winter which translates to the results in 5th total percentile. In a lesser way, a secondary cause of incorrect early intersection may be the instabilities of the RSD launching, causing strong shifts on temperature profile in lower layers. The surface temperature have also influence in ii) Development of the BL. The adiabatic evolution might not intersect the entrainment inversion causing the detection of a PBLH candidate inside the boundary layer or in the free atmosphere (as seen in the center section of figure 5). iii) A definition problem, related with the assumption that an adiabatic evolution is always held. With cloud presence, PBLH is typically defined at its base, but RSD0 does not consider cloud formation. So, if there is a saturation before the adiabatic evolution is completed, it continues, ignoring the temperature difference caused by rain droplets formation. Because of that, higher values are expected when a cloud is formed within the boundary layer than other methodologies which

consider cloud formation. On the other hand, when the BL is strongly developed during warm clear sky days, this algorithm allows to correctly estimate very high PBLH. This may be the reason of the high 75th percentile results during spring and summer and the detection of up to 3300 m outliers during this seasons.

- As WW defines PBLH where it encounters a sharp temperature, humidity and refractivity variation in the RSD profile, it is able to correctly identify the entrainment zone. Nevertheless, it may not differentiate it when the profile shows more than one sharp variation or middle temperature inversions (as shown in figure 5 bottom). This is produced usually in colder months, when the boundary layer is not highly developed, and may force the algorithm to detect the residual layer top. This behaviour may cause the computed values of WW 95th percentile -near 1950 m, the largest value among all 4 methods-. At the same time, lower WW values than those yielded by RSD0 are accumulated during summer. This may be caused by the maximum gradient detector quality filter, that can ignore peaks that may be the correct PBLH option. Both factors may explain the unexpected seasonal results. Summer lower results compared with the other seasons may be justified by sea breeze producing a decrease in PBLH. Besides that, WW algorithm is able to detect strong temperature inversions at low heights even with surface inversions, situation known as one of the most problematic to pollutant dispersion.
- Despite ERES algorithm allows to a simplification of the backscattering profile, the restriction of the idealized profile to a Gaussian error curve limits it. When the ideal curve does not adjust to the profile, because its smoothness or variability, PBLH is not computed. This is the main cause of the difference of the number of valid data between ERES -approximately 800- and all other algorithms -nearly 1100-. Despite mean results are similar to all other algorithms, because of the nature of the computation ERES PBLH often differs from all other candidates. The poor profile-curve adjustment may justify both the small seasonal variation -median varies from 850 m during autumn to 1000 m in winter- and the lower whisker results. It detects with great accuracy, on the other hand, cloud presence due to the detection of a sharp backscattering decrease. This introduces another error in the algorithm scheme: if there is a PBLH candidate lower than the cloud base, ERES would ignore it and indicate the cloud base as PBLH.
- CEIL algorithm accurately detects backscattering gradients, as it can be suggested by the correct seasonal evolution and the analysis of individual results. But, on the other hand, it may have some complications when there is a smooth backscattering profile or when it has to detect very high PBLH, as noise may dominate over signal. This may explain the lack of values higher than 2000 m, since the algorithm is forced to select a PBLH candidate from lower levels. This may be problematic during warm cloudless days, with a strongly developed BL. If PBLH is near the cloud base, high backscatter filtering may introduce another error source because it deletes backscattering data from cloud levels, so the algorithm is unable to detect a valid PBLH at this heights. Above all, the main issue is that intermediate calculations -time and spatial means, high backscattering values dismissed, etc.- are unknown to the user, making the algorithm behave like a "black box".

Considering all the statements, it is possible to conclude that RSD0 algorithm use is not fully adequate in winter as it strongly depends on surface temperature, which may be influenced by radiation inversions, and, consequently, it may not detect the entrainment zone correctly. On the other hand, works well for clear sky days with mixed boundary layers with a strong entrainment inversion, usually in warm days, mostly during summer. WW is affected by the existence of secondary inversions. So, the algorithm behaves in a good manner when a strong or a unique inversion are present in the temperature and humidity profiles. ERES algorithm is not entirely reliable since the backscattering profile may not adjust the Gaussian error curve proposed, not giving a valid result or inferring a deviated PBLH. Nevertheless it considers the existence of clouds, detects them, and gives its base as a PBLH during cloudy days. Lastly, CEIL operates correctly if the BL is not very developed, since upper levels noise may interfere in the backscattering profile. This may cause some complications in warm clear sky days.

### B. PBLH during pollution episodes

To evaluate PBLH during high pollution days, ERES and RSD0 algorithms are dismissed. As presented, ERES does not always adjust to the backscattering profile, giving thus, untrustworthy results. Regarding RSD0, high pollution episodes are often concentrated when the temperature profile may show surface or low level temperature inversions. BL might not be fully convectively developed, making the adiabatic evolution not intercept the temperature profile at the entrainment zone. WW, on the other hand, is able to detect low level inversions, which may be adequate to pollution days.

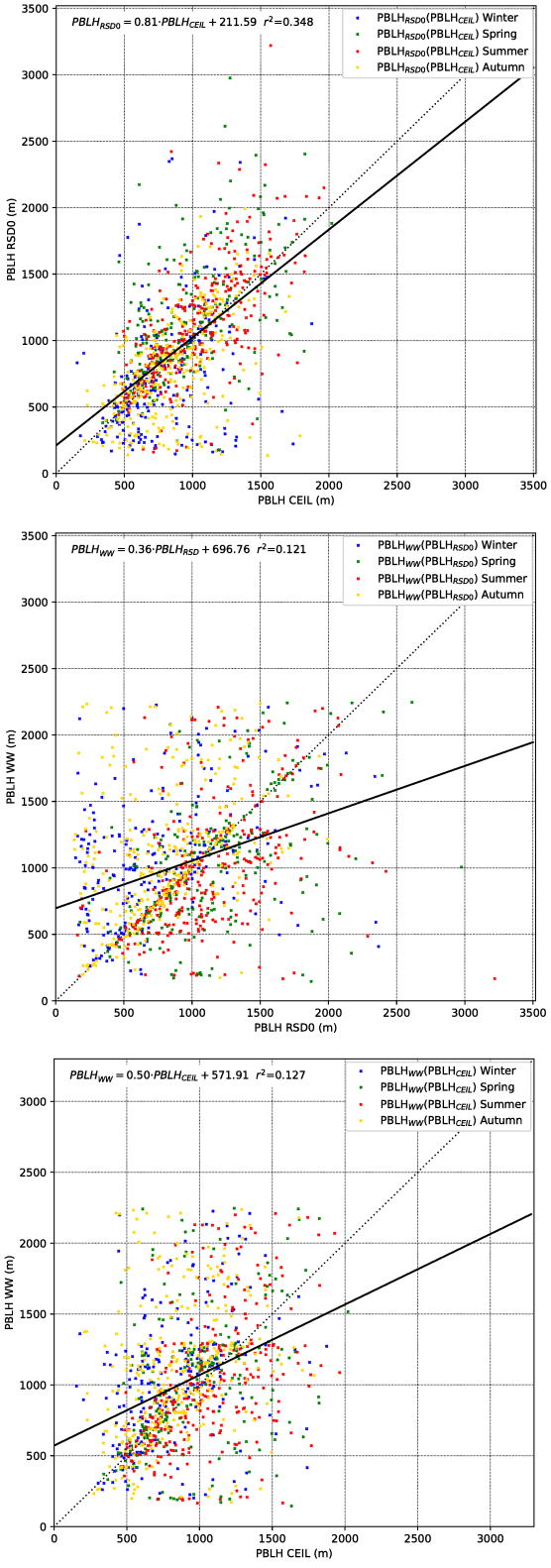


Figure 4. Scatter plots of PBLH computed by RSD0, CEIL and WW algorithms. Point colour indicate the season where the measurement was made: winter (blue), spring (green), summer (red) and autumn (yellow). The line is traced adjusting all points.

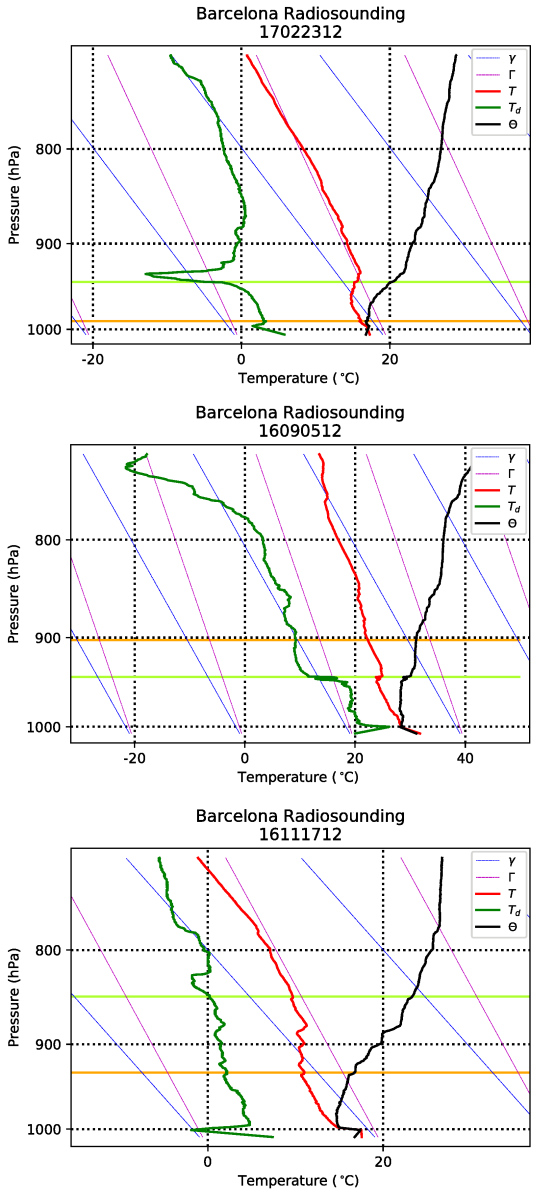


Figure 5. Details of 2016 September 5th ($NO_2$ episode), 2016 November 17th ($PM_{10}$ episode) and 2017 February 23rd ($PM_{10}$ episode) midday radiosoundings. There are represented temperature ($T$), dew point ($T_d$) and potential temperature ($\theta$) profiles, as well as curves marking the adiabatic evolution ($\gamma$) and an approximation to a saturated adiabatic evolution in low levels ($\Gamma$, with a slope of -0.6 K/hm). Orange horizontal line represents RSD0 PBLH determination and green horizontal line marks WW computation. Each one presents a problem relating to RSD0 or WW algorithms calculations: 2017-02-23 (top panel) presents 238 m for RSD0 and 646 m for WW (more reliable), since low surface temperature affects RSD0 results; 2016-09-05 (center) marks 1068 m for RSD0 and 685 m (more reliable) for WW, as RSD0 struggles finding the entrainment zone, and 2016-11-17 (last section) estimates a 767 m RSD0 PBLH (more reliable) and 1556 m for WW, because of the presence of various temperature inversions that hinder WW calculation.

It has been considered as "episode" all days in which a legal threshold is exceeded or if one of the high pollution scenarios -preventive warning or environmental episode- is activated. In total, there are 108 days with a total of 121 threshold superations, with episodes in the 3 years and 5 months analysed: 55 $PM_{10}$ 50 $\mu g/m^3$ daily mean superations, 37 superations of 160 $\mu g/m^3$ $NO_2$ hourly mean, 28 $O_3$ exceedance counting for VOPS threshold and 1 $O_3$ LLI superation. In this counting, there are included combined episodes, when 2 or more pollutant thresholds are exceeded the same day: 11 $NO_2$ and $PM_{10}$ episodes, 1 VOPS and LLI $O_3$, and 1 $NO_2$, $PM_{10}$ and $O_3$ thresholds exceedance.

Figure 6 shows the same scatter plot of PBLH calculation using WW and CEIL algorithms presented in last sector of figure 4, but without seasonality distinction. Instead, $PM_{10}$ and $NO_2$ episodes are highlighted. We can see most of the episodes (51 of 80, a 64%) located in the region where PBLH computed by both methods are lower than 1000 m. 23 measures (29%) show one of two methods with a value lower than 1000 m. This behaviour is related to the method difficulties detected in the section III A. The 6 remaining values, corresponding to WW and CEIL heights over 1000 m, are related to summer days with great boundary layer development.

Figure 7 displays a PBLH boxplot for CEIL and WW methods, distinguishing between: all cases, pollution days, high $PM_{10}$ and $NO_2$ pollution days studied together and individually, and $O_3$ episodes. Data obtained is shown in table IV. Pollution days present, as expected, a smaller PBLH than all data, but there are some differences in relation with the pollutant. $NO_2$ and $PM_{10}$ episodes show a low PBLH, with means of 673 m for CEIL $NO_2$ (a 27% less than all CEIL cases), 791 m for WW $NO_2$ (24% lower), 744 m for CEIL $PM_{10}$ (19% less than all WW data) and 746 m for WW $PM_{10}$ (28%). Additionally, it can be also seen that the median of all episodes is near the 25th percentile of all analysed valid days. This confirms an expected result: low PBLH favours high concentrations of $PM_{10}$ and $NO_2$, as the volume of air where pollutants are dispersed is reduced. Nevertheless, tropospheric ozone behaves in a different way than $NO_2$ or $PM_{10}$ episodes, as it can be seen by the mean results for both algorithms. High temperatures are required to force the reactions responsible to its formation, which allows the boundary layer to grow as convection is more thermically developed. This may explain why PBLH results for $O_3$ are higher than for the other pollutants. $O_3$ PBLH determination behaviour differs also in the methodologies: CEIL interquartile range (IQR) is higher than WW IQR, and influences all pollution days result. This may be explained by the impact of high developed BL in the algorithm calculation: WW may point small intermediate variations, and CEIL, although may ignore candidates masked by noise, can still signal a higher candidate than WW.

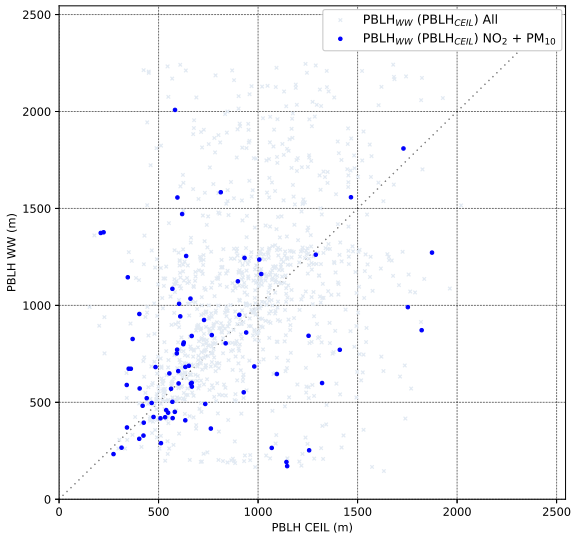


Figure 6. PBLH computed by WW and CEIL algorithms. Gray shaded represents all available days, blue for $PM_{10}$ and $NO_2$ episodes.

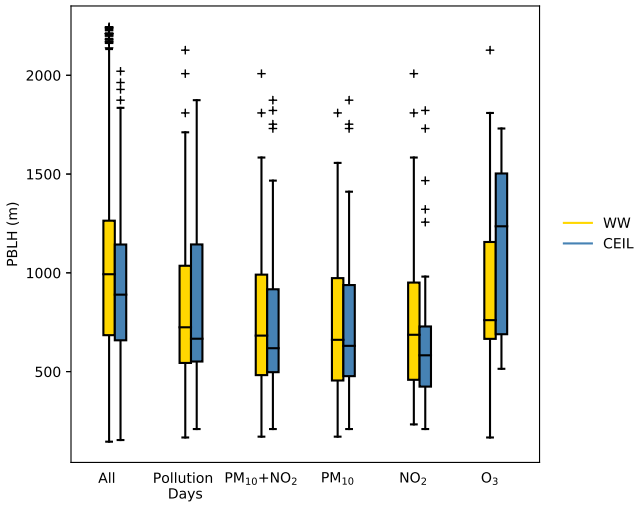


Figure 7. Boxplot for WW and Ceil algorithms. All available days and pollution days are shown.

Table IV. Statistical data for pollutant episodes for WW and CEIL PBLH computing algorithms.

| Measurement | Valid data | Mean (m) | Std (m) | 5th (m) | 25th (m) | 50th (m) | 75th (m) | 95th (m) |
|---|---|---|---|---|---|---|---|---|
| WW $PM_{10}$+$NO_2$+$O_3$ | 112 | 814 | 407 | 265 | 544 | 724 | 1036 | 1607 |
| WW $PM_{10}$+$NO_2$ | 81 | 777 | 394 | 265 | 482 | 682 | 991 | 1556 |
| WW $PM_{10}$ | 55 | 746 | 374 | 255 | 455 | 661 | 973 | 1375 |
| WW $NO_2$ | 38 | 791 | 443 | 264 | 459 | 686 | 951 | 1617 |
| WW $O_3$ | 32 | 938 | 446 | 374 | 666 | 761 | 1156 | 1755 |
| CEIL $PM_{10}$+$NO_2$+$O_3$ | 107 | 837 | 416 | 342 | 551 | 667 | 1144 | 1627 |
| CEIL $PM_{10}$+$NO_2$ | 79 | 733 | 374 | 337 | 497 | 618 | 917 | 1493 |
| CEIL $PM_{10}$ | 54 | 744 | 388 | 300 | 477 | 630 | 938 | 1523 |
| CEIL $NO_2$ | 37 | 673 | 379 | 306 | 424 | 583 | 729 | 1519 |
| CEIL $O_3$ | 29 | 1151 | 399 | 555 | 689 | 1236 | 1503 | 1680 |

Top and bottom panels in figure 8 present the evolution of PBLH in $NO_2$ or $PM_{10}$ pollution days for both methods during the 3 years 5 months period. Results disaggregated by season and pollutant are presented in table V. At first sight, they serve to identify periods of high pollution days accumulation (for instance December 2015, June 2016 and May 2018), whereas there are seasons where pollutant episodes are not registered (for example, 2016 spring or 2017 and 2018 autumns). Nevertheless, if favourable dispersion conditions are not met -such as the typical spring and autumn weather variability and rain accumulation-, cases like autumn 2015 may happen, with 25 high pollution days. Summer presents a lower number $NO_2$ episodes, maybe influenced by a lower emissions impact derived from vacation time, since traffic and mobility are reduced. The episodes distribution may suggest that seasonality does not have a strong effect on high contamination days -besides the behaviour of $NO_2$ summer episodes-, but they are influenced by large scale meteorological forcings. On the other hand, results show a low PBLH in most of $NO_2$ and $PM_{10}$ episodes. A 53% of all episodes for WW and 59% for CEIL, are under the 25th PBLH total cases percentile for each method without any significant differences between $NO_2$ and $PM_{10}$ episodes.

## IV. CONCLUSIONS

In this paper, 4 different algorithms -Parcel Method (RSD0), Gradient Detector (WW), Function Adjustment (ERES) and BL-View Enhanced Gradient (CEIL)- are tested to calculate the boundary layer height (PBLH) using the radiosounding and the ceilometer located in Universitat de Barcelona for the period from August 2015 to December 2018. PBLH mean in Barcelona is between

Table V. $PM_{10}$ and $NO_2$ episodes identified by season where they were registered. $PM_{10}$+$NO_2$ episodes are counted twice, one in each category, but once in the "Total" category.

| | Winter | Spring | Summer | Autumn |
|---|---|---|---|---|
| $PM_{10}$ | 6 | 14 | 12 | 23 |
| $NO_2$ | 9 | 9 | 3 | 16 |
| Total | 15 | 20 | 15 | 30 |

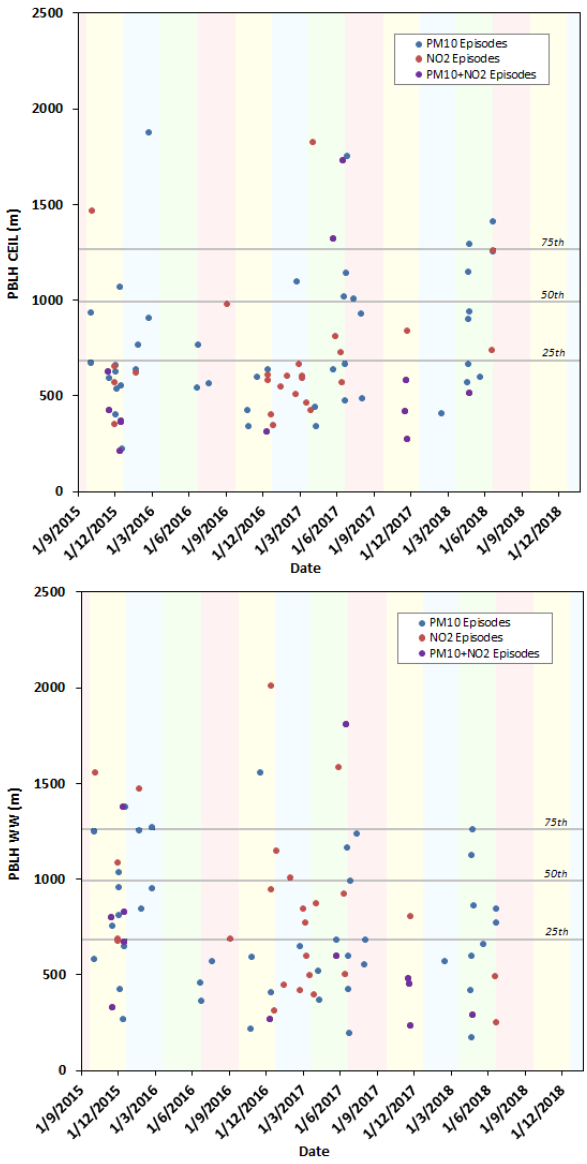


Figure 8. Time - PBLH plots for pollution days calculated with WW and Ceil algorithms. Background colour is related to the season: winter (blue), spring (green), summer (red) and autumn (yellow). Horizontal lines mark 25th, 50th and 75th percentile of total data of each method respectively.

920 m and 1035 m. This calculation is done for all days without rain, in all type of sky conditions and all type of large scale influence. The PBLH computed with this algorithms usually produced high disparity between results. Some causes related to this behaviour have been identified: RSD0 may not detect the entrainment zone inversion limiting the boundary layer because of its dependency of initial temperature values, which yields to poor results during winter as marked by the low median during this season; WW may point PBLH candidates from small intermediate inversions, complicating the PBLH determination if the boundary layer is not well developed or there is not a marked inversion in entrainment zone; ERES may give unsatisfactory results if the backscattering profile is not adjustable to a Gaussian error function, as showed by the decrease of analysed cases compared to other algorithms, and CEIL may produce inaccurate computations if the boundary layer is highly developed, as noise may dominate over backscatter signal at upper levels. This analysis suggests that RSD0 performs well during warm cloudless days, whereas CEIL may be useful to detect cold days PBLH. WW would be convenient to use during days with strong temperature inversions. ERES could be used as a support method during cloudy days.

The main advantages and weaknesses of all methods have been taken into consideration to select WW and CEIL as the analysis methodologies for pollution episodes. Results suggests the paper of low PBLH in $NO_2$ and $PM_{10}$ episodes, since more than 50% cases are lower than the 25th PBLH percentile for all data and for each algorithm (that is, lower than 684 m for the WW method). On the other side, $O_3$ photochemical origin and, consequently, the necessity of clear skies and high temperatures, produces a different behaviour in PBLH than $NO_2$ and $PM_{10}$. The conditions needed for its formation favours a high developed PBLH as showed by the values obtained, which are much higher than $NO_2$ and $PM_{10}$. With this BL conditions, CEIL and WW may not be appropriate. Future work may consist in identifying the adequate method in different meteorological situations and sea breeze development (*Arrillaga et al.* 2016).

Nevertheless, the differences between both methods used to compute PBLH during pollution days, and the possible errors involved in its calculations may difficult its application to determine PBLH with certitude. Only a 45% of the the computations using both algorithms show differences lower than 200 m. The linear adjustment done in the last sector of figure 4 and the poor R-squared obtained ($r^2$=0.127) point out that both algorithms must be improved in order to generate an automatic PBLH computation method reliable to its application during high pollution episodes. Once it is developed, WW method may be applied to modelled radiosoundings to predict the next day PBLH. In a similar way, RSD0 and ERES algorithms may be useful to the PBLH prediction, but they must be improved in order to assure its correct functioning.

PBLH is not the only factor to high pollution levels. As it can be seen in figures 4 and 6, a low PBLH does not immediately imply high pollution concentration. Important factors of horizontal dispersion, wind and terrain for instance, emissions and pollution transport, must also be considered. Future work may consist in developing a joint air ventilation variable considering together both vertical and horizontal dispersion. This may help to predict of high pollution episodes, and, consequently, to apply preventive measures to avoid heavy $PM_{10}$ and $NO_2$ concentrations.

## ACKNOWLEDGMENTS

I would like to thank Mireia Udina, my TFM advisor, for all her precious help and patience during this paper elaboration process. Many thanks also to Yolanda Sola for providing RSD data and her assist with backscattering files decoding, Joan Montolio for helping me to understand his algorithm, and to the immissions section of Territory and Sustainability Department for pollutant data and all I have learned during my stay there.

## REFERENCES

Arrillaga, J. A., C. Yagüe, M. Sastre, and C. Román-Cascón, A characterisation of sea-breeze events in the eastern Cantabrian coast (Spain) from observational data and WRF simulations, *Atmospheric Research*, *181*, 265–280, doi: 10.1016/j.atmosres.2016.06.021, 2016.

Arya, P., *Air Pollution Meteorology and Dispersion*, Oxford University Press, 198 Madison Avenue, New York, New York 10016, 1999.

BCN-CC, Barcelona City Council, Low Emissions Zone, `https://ajuntament.barcelona.cat/qualitataire/ca/afectacions-la-mobilitat/que-es-la-zona-de-baixes-emissions-de-barcelona`, (last access: Jan. 2020), 2020.

Dades Obertes, Generalitat de Catalunya Open Data Portal: XVPCA measurement points immission data, `https://analisi.transparenciacatalunya.cat/Medi-Ambient/Dades-d-immissi-dels-punts-de-mesurament-de-la-Xar/uy6k-2s8r`, (last access: Jan. 2020), 2020.

de Haij, M., W. Wauben, and H. K. Baltink, Continuous mixing layer height determination using the LD-40 ceilometer: a feasibility study, *Royal Netherlands Meteorological Institute (KNMI)*, 2007.

ED, European Directives on Air Quality, European Environment Agency, `https://www.eea.europa.eu/themes/air/air-quality-standards`, (last access: Jan. 2020), 2019.

Endlich, R., F. Ludwig, and E. Uthe, An automatic method for determining the mixing depth from lidar observations,

*Atmospheric Environment (1967)*, *13*(7), 1051–1056, doi: 10.5194/acpd-5-12697-2005, 1979.

Eresmaa, N., A. Karppinen, S. Joffre, J. Räsänen, and H. Talvitie, Mixing height determination by ceilometer, *Atmospheric Chemistry and Physics*, (6), 1485–1493, doi: 10.5194/acpd-5-12697-2005, 2006.

Holzworth, G. C., Estimates of mean maximum mixing depths in the contiguous United States, *Mon. Weather Rev*, *92*(5), 235–242, doi:10.1175/1520-0493(1964)092¡0235:EOMMMD¿2.3.CO;2, 1964.

IDESCAT, Statistical Institute of Catalonia, `http://https://www.idescat.cat/`, (last access: Jan. 2020), 2020.

Levi, Y., U. Dayan, I. Levy, D. M. Broday, et al., On the association between characteristics of the atmospheric boundary layer and air pollution concentrations, *Atmospheric Research*, *231*, 104,675, doi:10.1016/j.atmosres.2019.104675, 2020.

Montolio, J., Y. Sola, and M. Pandolfi, Inferring boundary layer time series from ceilometer data and its influence on air quality over Barcelona, TFM, Universitat de Barcelona, 2018.

Oke, T., *Boundary Layer Climates*, Halsted Press, New York, 1988.

RD, *Reial Decret* 102/2011, `https://www.boe.es/buscar/doc.php?lang=va&id=BOE-A-2011-1645`, (last access: Jan. 2020), 2011.

Seidel, D. J., C. O. Ao, and K. Li, Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, *Journal of Geophysical Research: Atmospheres*, *115*(D16), doi:10.1029/2009JD013680, 2010.

SMC, Catalan Meteorological Service, Radiosoundings, `https://www.meteo.cat/wpweb/divulgacio/equipaments-meteorologics/radiosondatge/`, (last access: Jan. 2020), 2017.

Stull, R., *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, P. O. Box 17, 3300 AA Dordrecht, the Netherlands, 1988.

Vaisala, BL-View User's Guide, 2010.

Vaisala, Ceilometer CL31 User's Guide, 2014.

Vaisala, Ceilometer CL31 Technical data, 2018.

Vidal, V., M. Udina, and J. Bech, Anàlisi dels nivells de contaminació per $NO_2$, $PM_{10}$ i $O_3$ a la Zona de Baixes Emissions de Barcelona al període 2015-2018, 2019.

Wang, X., and K. Wang, Estimation of atmospheric mixing layer height from radiosonde data, *Atmospheric Measurement Techniques Discussions*, *7*(2), doi:10.5194/amt-7-1701-2014, 2014.

WHO, World Health Organization Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide, Global update 2005, `https://apps.who.int/iris/bitstream/handle/10665/69477/WHO_SDE_PHE_OEH_06.02_eng.pdf?sequence=1`, (last access: Jan. 2020), 2005.

XVPCA a, XVPCA key concepts, `http://mediambient.gencat.cat/ca/05_ambits_dactuacio/atmosfera/qualitat_de_laire/avaluacio/xarxa_de_vigilancia_i_previsio_de_la_contaminacio_atmosferica_xvpca/conceptes_clau/index.html`, (last access: Jan. 2020), 2014.

XVPCA b, XVPCA measurement stations classified by Air Quality Zones, `http://www.gencat.cat/mediamb/qaire/web/mapa_nou_2005.htm`, (last access: Jan. 2020), 2018.