Estimating the urban atmospheric boundary layer height from remote sensing applying machine learning techniques

Gregori de Arruda Moreira [1,2,3], Guadalupe Sánchez-Hernández [1,4], Juan Luis Guerrero-Rascado [1,2] , Alberto Cazorla [1,2], Lucas Alados-Arboledas [1,2]

[1] Andalusian Institute for Earth System Research (IISTA-CEAMA), Granada, Spain
[2] Dpt. Applied Physics, University of Granada, Granada, Spain
[3] Federal Institute of São Paulo (IFSP), São Paulo, Brazil
[4] Department of Physic, University of Jaén, Jaén, 2371, Spain
Correspondence to: Guadalupe Sánchez-Hernández (guadalupesh@ugr.es)

Keywords Atmospheric Boundary Layer height, ceilometer, gradient boosting regression tree

Abstract

This study proposes a new methodology to estimate the Atmospheric Boundary Layer Height (ABLH), discriminating between Convective Boundary Layer and Stable Boundary Layer heights, based on the machine learning algorithm known as Gradient Boosting Regression Tree. The algorithm proposed here uses a first estimation of the ABLH derived applying the gradient method to a ceilometer signal and several meteorological variables to obtain ABLH values comparable to those derived from a microwave radiometer. A deep analysis of the model configuration and its inputs has been performed in order to avoid the model overfitting and ensure its applicability. The hourly and seasonal values and variability of the ABLH values obtained with the new algorithm have been analyzed and compared with the initial estimations obtained using only the ceilometer signal. Mean Relative Errors (MRE) between the ABLH estimated with the new algorithm and microwave radiometer show a daily pattern with their highest values during the night-time (stable situations) and their lowest values along the day-time (convective situations). This pattern has been observed for all the seasons with MRE ranging between -5% and 35%. This result notably improves those ABLH values derived by applying the gradient method to ceilometer data during convective situations and enables the Stable Boundary Layer height detection at night and early morning, instead of only Residual Layer top height. Finally, the model performance has been directly validated in three particular cases: clear-sky day, presence of low-clouds and dust outbreak event. In these three particular situations, ABLH values obtained with the new algorithm follow the pattern obtained with the microwave radiometer presenting very similar values, thus confirming the good model performance. In this way it is feasible by the combination of the proposed method with gradient method, to estimate Convective, Stable and Residual Boundary Layer height from ceilometer data and surface meteorological data in extended network that include ceilometer profiling.

1 Introduction

The Atmospheric Boundary Layer (ABL) is defined as the part of the troposphere that is directly influenced by the presence of the Earth's surface, and responds to surface forcings with a timescale of about an hour or less (Stull, 1988). The ABL is the atmospheric region directly affected by turbulent and evapotranspiration processes and where the air pollutants are dispersed (Stull, 1988). The characteristics of the ABL, and particularly, the ABL height (ABLH), play a fundamental role in numerous atmospheric areas such as weather forecasting, air quality and/or numerical modeling (e.g. Cheng et al., 2011).

The estimation of the ABLH with high temporal resolution is not an easy task, due mainly to its high variability throughout its daily cycle. Thus, based on an ideal scenario, some instants after the sunrise, the ground surface temperature begins to increase, due to the positive net radiative fluxes. Such a phenomenon causes the warming of air masses located at low heights favoring the convective process, and the heat transfer from the surface to upper atmospheric layers in the troposphere. This process generates a layer known as Convective Boundary Layer (CBL). Just before the sunset, the CBL becomes a layer called Residual Layer (RL), which is stably stratified and contains the characteristics from the previous CBL. In conjunction with this process arises from the ground a thermally stratified layer and endowed of lower heights (in comparison with RL and CBL), denominated Stable Boundary Layer (SBL).

In the last years remote sensing systems, such as elastic lidars (e.g. Toledo et al., 2017; Bravo-Aranda et al., 2017; Moreira et al., 2019; Vivone et al., 2021), ceilometers (e.g. Haeffelin et al., 2012; Caicedo et al., 2017; Lee et al., 2019; Uzan et al., 2020; Moreira et al., 2020a; Jiang et al., 2021), Doppler lidars (e.g. Manninen et al., 2018; Marques et al., 2018; Moreira et al., 2019) and microwave radiometers (e.g. Cimini et al., 2013; Bravo-Aranda et al., 2017; Moreira et al., 2020a; Jiang et al., 2021) have been widely used to characterize the ABLH. Among these remote sensing systems, ceilometers have the advantage to be a low-cost and low-maintenance system that monitors aerosol and clouds layers (Lee et al., 2019). Such characteristics have favored the creation of national (e.g., Automated LIdar-CEilometer network – ALICEnet (Haefele et al., 2016); Unified Ceilometer Network - UCN (National Research Council, 2009)) and international networks (e.g., EUMENET-Profiling Program [https://www.eumetnet.eu/]; E-PROFILE [https://e-profile.eu]; Iberian Ceilometer Network - ICENET) (Cazorla et al., 2017), which have been dedicated to standardize and expand the activities of ABL monitoring by ceilometer data.

Ceilometers have been applied in many previous works related to ABLH detection, which vary from short-term (e.g. Helmis et al., 2012; Bruine et al., 2017; Caicedo et al., 2017) to long-term studies (e.g. Stachlewska et al., 2012; Schween et al., 2014; Moreira et al., 2020), and various mathematical algorithms such as, vertical gradients (Emeis et al., 2008), wavelet covariance transform (Baars et al., 2008; Granados-Muñoz et al., 2012), STRAT (STRucture of the ATmosphere) [application of first derivative of the Gaussian filter on Range Corrected Signal (RCS) profile] (Morille et al., 2007), STRAT-2D [it has same structure of STRAT and includes an edge detection method based on both vertical and temporal gradients of RCS] (Haeffelin et al., 2012), STRAT+ [combination of radiosoundings information and Canny edge detection applied to gradient and variance profiles of RCS] (Pal et al., 2013), PathfinderTURB [it combines the strength points of gradient and variance of RCS methods and addresses the layer attribution problem by adopting a geodesic approach] (Poltera et al., 2017), or COBOLT (COntinuous BOundary Layer Tracing) [a time-height tracking procedure] (Geiß et al., 2017) have been developed in order to improve the ABLH values derived from them. In spite of this, there are still limitations in the application of ceilometers for ABLH monitoring. Special difficulties occur in cases considered as complex such as rainy situations, presence of low clouds, and dust outbreaks. In these situations, the abrupt changes of the aerosol vertical profile notably differ from the idealized profile on which most of the ABLH detection methods are based. Although some methods have been proposed to improve the ABLH detection from lidar data in the situations mentioned above (Bravo-Aranda et al., 2017; Liu et al., 2018), the use of only one wavelength and/or low signal-to-noise ratio make it difficult to apply such techniques in ceilometers. Another weakness in the application of ceilometers to obtain ABLH is the difficulty for discriminating between the RL top height (RLH) and SBL height (SBLH) during stable periods (Moreira et al., 2020a). This limitation comes from the basis of the detection procedure applied to ceilometers, based on the vertical profile of the atmospheric aerosol that prevents the detection of the top of the thermal inversion (SBLH) in this situation. Having in mind these facts, there is still room for some improvements in the ABLH retrieval with ceilometers on the basis of alternative data processing of the ceilometer's output.

Machine learning techniques have been widely applied in the environmental sciences during the last years (e.g. Cadeddu et al., 2009; McGovern et al., 2017; Bonnin et al., 2018; Vassalo et al., 2020; Moreira et al., 2021). These ML techniques can account for complex relationships on atmospheric processes and have been successfully applied in several atmospheric areas, ranging from the estimation of atmospheric parameters such as the mixing layer height (Bonin et al., 2018) or the analysis of sky-camera images to characterize the aerosol layer (Cazorla et al., 2008; Cazorla et al., 2009) to predict pollutants concentration (Moreira et al., 2021). Particularly for the estimation of the ABLH, Jiang et al. (2021) applied machine learning combined GPS radio occultation technology to build a simulation model to estimate the ABLH, providing reliable results for several months. Krishnamurthy et al. (2021) proposes a method based on Random Forest algorithm to estimate the ABLH from meteorological and Doppler lidar data. Such an algorithm provides an improvement of 50%, in the CBLH detection during clear sky or cloudy conditions, in comparison with a method based on vertical wind speed profiles.

Thus, the main objective of this study is to propose a machine learning algorithm to improve the ABLH estimations obtained from a ceilometer. To this aim, a Gradient Boosting Regression Trees (GBRT) algorithm has been trained using as input the ABLH values derived from a ceilometer, surface meteorological data and ABLH values derived from a co-located microwave radiometer (Moreira et al., 2018; Moreira et al., 2020a). From this algorithm it is possible to estimate the height of CBL and SBL combining ceilometer and meteorological surface data. Therefore, application of such methodology can expand ceilometer data applicability, so that it is possible to discriminate the three main ABL sublayers (CBL, SBL and RL) without acquiring expensive instruments. These ABLH values derived from the microwave radiometer have been the reference dataset for both the fitting and validation analysis. The model performance has been assessed analyzing its temporary and seasonal variability as well as the dependence of its residuals against several meteorological variables. Finally, the model has been directly validated in three particular cases (clear-sky day, presence of low-clouds and dust outbreak event).

The paper is organized as follows. First, the experimental site and the instrumentation used in this study have been described in Section 2. Then, in Section 3 the development and set-up of the machine learning algorithm here proposed are presented. Next the results of a seasonal analysis as well as the analysis of particular cases are shown in Section 4, while the main conclusions of this study are discussed in Section 5.

2 Experimental site and instrumentation

The measurements analyzed in this study were recorded at the University of Granada (UGR) station located on the roof of the Andalusian Institute of Earth System Research (IISTA-CEAMA) at Granada (37.164º N, 3.605º W, 680 m a.s.l.). These facilities are managed by the Atmospheric Physic Research Group (GFAT) and they are part of the observatory AGORA (Andalusian Global ObservatoRy of the Atmosphere) in the framework of

ACTRIS (Aerosol, Clouds and Trace Gases Research Infrastructure) and of the Iberian Ceilometer Network (ICENET) (Cazorla et al., 2017).

Granada is a medium sized non-industrialized city in the West Mediterranean region, at the Southeast of Spain, and presents a large seasonal temperature range associated with its Mediterranean-continental conditions. The city is characterized by cool winters and hot summers with the most humid period from late autumn to early spring (AEMET, 2015). This region is usually affected by mineral dust outbreaks from the Sahara desert in summer and spring (e.g. Guerrero-Rascado et al., 2008, 2009; Bravo-Aranda et al., 2015), and some extreme events have also occurred in winter (Cazorla et al., 2017; Fernández et al., 2019), but also for more local sources of aerosol particles such as traffic, domestic-heating or biomass burning in winter time (Titos et al., 2012, 2017). From roughly February to July, primary biological aerosol particles (pollen-type) are present in the region (e.g., Cariñanos et al., 2021). To a lesser extent, the area is also affected by advected fresh and aged smoke mainly from the Iberian Peninsula (Alados-Arboledas et al., 2011) and from North America (Ortiz-Amezcua et al., 2017), respectively. The combination of all these factors highly affect the local meteorology, and, therefore the ABL detection (Stull, 1988).

A ceilometer Jenoptik model CHM15k was operated at the UGR station. This instrument measures the backscattered signal of a pulsed Nd:YAG laser emitting at 1064 nm, with an energy per pulse of 8.4 μJ, a repetition frequency in the range of 5–7 kHz and a laser beam divergence less than 0.3 mrad. The backscattered signal is received by a telescope with a field of view of 0.45 mrad. The spatial and temporal resolution used were 15 m and 15 s, respectively. The complete overlap of the instrument is found around 1500 m above a.g.l. and its overlap is 90% at 555 m a.g.l., in accordance with the overlap function provided by the manufacturer (Cazorla et al., 2017). This equipment has been operating continuously since December 2012 and it is part of the Iberian Ceilometer Network (ICENET), an initiative of the Atmospheric Physics Group of the University of Granada (Cazorla et al., 2017). Measurements recorded with this instrument were employed to derive initial estimations of the ABLH, $ABLH_{CEIL}$, as input (feature) of the machine learning algorithm. The range of values to $ABLH_{CEIL}$ [200 – 4500 m] (Table 1) is based on a previous long-term study (2012-2016) presented in Moreira et al. (2020a).

The surface meteorological dataset consists of 1-min data of air temperature (T), relative humidity (RH), atmospheric pressure (P), and wind speed (WS) measured at the roof of the UGR station facilities and covering the whole analyzed period 2015-2017. T and RH at this station were monitored by a HMP60 probe manufactured by Vaisala. This probe has an accuracy of ±0.6 ºC and 2% for T and RH measurements, respectively. In this station WS was measured by an anemometer model 05103, manufactured by Campbell Scientific, with an accuracy of ±0.3 m/s. Simultaneously, P was monitored by a Vaisala PTB110 barometer with a silicon capacitive sensor specially designed to guarantee accurate (±0.3 hPa at +20 °C) and stable (±0.1 hPa/year) measurements. Global horizontal irradiance (G) in the range 280-2800 nm was measured by a CM-11 pyranometer manufactured by Kipp & Zonen while downward infrared irradiance (IR) in the range of 4000-50000 nm is measured by a Precision Infrared Radiometer (PIR) manufactured by EPPLEY. Both instruments comply with the specifications for the first-class WMO pyranometer classification with an accuracy below ± 5 W/m2 for daily measurements. All these sensors operated following the WMO standard protocols and procedures (WMO, 2013). These measurements and magnitudes derived from them were employed as input for the machine learning algorithm developed in this study as both input (features).

Co-located to the ceilometer, a ground-based passive microwave radiometer (MWR), model RPG-HATPRO G2 (Radiometer Physics GmbH) was operating in the scanning mode in automatic and continuous mode since November 2011 as part of MWRnet [http://cetemps.aquila.infn.it/mwrnet/] (Rose et al., 2005; Caumont et al., 2016). This instrument measures the sky brightness temperature with a radiometric resolution between 0.3 and 0.4 K root mean square error at 1 s integration time. The MWR uses direct detection receivers within two bands, 22-31 GHz (water vapor - K band) and 51-58 GHz (oxygen - V band), for deriving RH and T profiles, respectively, by inversion algorithms described in Rose et al. (2005). Both profiles have a range resolution varying between 10 and 200 m in the first 2 km and varying between 200 and 1000 m up to 10 km (Navas-Guzmán et al., 2014). This change in the profile resolution is associated with an exponential decrease with height of the MWR weighting functions (Spänkuch et al., 1996). The measurements recorded with this instrument were employed to derive the reference values (target) of the ABLH, $ABLH_{MRW}$. In the same way of the $ABLH_{CEIL}$, the range of values to $ABLH_{MWR}$ [200 – 4500 m] (Table 1) is based on Moreira et al. (2020a). The application of MWR data to ABLH detection have been extensively validated with other instruments such as: Doppler lidar (Moreira et al., 2018; Moreira et al., 2020b), elastic lidar (Granados-Muñoz et al., 2012; Bravo-Aranda et al., 2017; Moreira et al., 2018; Moreira et al. 2020a) and radiosoundings (Bedoya-Velásquez et al., 2019). Particularly, $ABLH_{MRW}$ has shown to be less influenced by presence of clouds (Moreira et al., 2020a) and decoupled aerosol layers (Bravo-Aranda et al., 2017) compared with other devices. Moreover, the MWR temporal resolution of 2 min guarantees the volume of data required for the development of the machine learning algorithm.

Finally, a database of hourly values of all these variables, listed in Table 1, were built for the period analyzed in this study, which encompasses three entire years from 2015 to 2017. This final dataset was split in two subsets: (1) a training subset, formed by measurements recorded along 2015 and 2016, and (2) a validation subset, composed by measurements taken along 2017.

3 Methodology

3.1 Gradient Boosted Regression Trees

The Gradient Boosting Regression Trees (GBRT) is a supervised non-parametric machine learning technique widely applied in classification and regression problems (e.g., Friedman 2001; Li et al. 2008; Ye et al. 2009; Chen et al. 2015; Baturynska et al. 2020). The idea behind boosting is to sequentially fit multiple 'weak learners', that is, simple models that perform relatively poorly with low accuracy (Friedman, 2001). In each iteration, a new model is proposed using information from the previous model trying to learn from its mistakes and improving iteration by iteration. In the case of GBRT, the 'weak learners' are decision tree models with very few branches. Because GBRT operates with small models training sequentially, it is a faster process and requires lower memory consumption than other machine learning techniques such as Random Forests. Additionally, GBRT does not require the application of advanced normalization techniques on its inputs and enables a combination of different numerical and categorical data as input. Similarly to the development of other types of atmospheric models, this machine learning technique requires both independent variables and reference data to build the model. Particularly in the machine learning vocabulary independent variables are named as features while the reference data is denoted as target. The variables used in this study will be detailed described in the next sections.

Figure 1 shows a flowchart that briefly describes the main steps of how an ensemble of trees is created by the GBRT algorithm. The ensemble consists of M trees built one-by-one. Thus, a first decision tree ($Tree_1$) is trained using the feature matrix, X ($ABLH_{CEIL}$ and meteorological variables), and the target variable y ($ABLH_{MRW}$). The predictions of Tree1 ($F_1(X)$) are used to determine the pseudo-residual errors (r_1) of the training set applying the loss function L (Root Mean Square Error in our case). After that, a second decision tree ($Tree_2$) is trained using X and r_1, which is the new target variable, as inputs. From the $Tree_2$ predictions ($F_2(X)$) the new pseudo-residual r_2 are computed and used as input to build an improved third tree. Such a process is repeated M times until the residuals are minimized and the improvement between consecutive trees is negligible. Finally, F_M (X) is obtained as the combination of the predicted values provided by each m-tree:

$$F_M(X) = F_0(X) + F_1(X) + \ldots + F_{m-1}(X) + F_m(X) \quad (1)$$

From Fig. 1 it is possible to observe that as more trees are added to the model, there is a progressive tendency to reduce errors in predictions. A more detailed description of this process, including the most relevant mathematical aspects, is given in Appendix 1.

3.2.- GBRT set up

3.2.1.- Inputs for the GBRT: features and target variables

The features initially selected to build the GBRT algorithm have been the ABLH obtained from the aerosol vertical profiles measured with the ceilometer, $ABLH_{CEIL}$, and set of near-surface meteorological variables which influence on the ABLH as reported in previous works (e.g. Stull, 1988; Georgoulias et al., 2009; Granados-Muñoz et al., 2012; Haeffelin et al., 2012 Allabakash and Lim, 2020; Rey-Sanchez et al., 2021)). $ABLH_{CEIL}$ have been obtained applying the gradient method (Flamant et al., 1997) on the 1-hour averaged range corrected signal ($\acute{RCS}$). Considering the existence of an intense reduction in the aerosol load in the transition region between the ABL and the Free Troposphere (FT), this methodology estimates the ABLH as the height (z) where the minimum in the gradient of the $\acute{RCS}$ profile is detected (Moreira et al., 2020a). Rain cases were flagged by an empirical threshold and removed (Moreira et al., 2020a). This methodology has been widely applied to numerous ceilometers belonging to national or international networks such as E-PROFILE (Haefele et al., 2016) or ICENET (Cazorla et al., 2017).

The initial near-surface meteorological dataset is composed of WS, T, P, RH, G and NR. On one hand, several authors have reported significant correlations between ABLH and near-surface values of WS, T, P and RH (Georgoulias et al., 2009; Wang et al., 2009; Allabakash and Lim, 2020; Krishnamurthy et al., 2021). On the other hand, G accounts for the total energy reaching the surface while NR is a proxy of the brightness temperature of the atmosphere, highly correlated with its composition. The solar zenith angle (SZA), the hour of the day (H) and the season (S), have been also included as inputs in order to account for the Sun position and possible daily and seasonal dependencies. Additionally, the clearness index (kt), estimated as the ratio between the solar radiation at the top of the atmosphere and the global solar irradiance on the Earth's surface, has been also initially considered as a proxy of atmospheric transmissivity and cloudiness, respectively. In addition to their influence of these variables on the ABLH they have been chosen because of their wide availability through national and international meteorological and radiation networks as well as from reanalysis and satellite databases.

In this study, the reference values or target, also included as input in the GBRT algorithm, are the ABLH values obtained from the MWR ($ABLH_{MWR}$). Such an ABLH is calculated from the potential temperature profile in an algorithm that combines gradient and parcel methods, for stable and convective situations, respectively. This technique has been previously validated with respect to co-located elastic lidar (Granados-Muñoz et al., 2012; Bravo-Aranda et al., 2017; Moreira et al. 2018) and Doppler lidar (Moreira et al. 2018; Moreira et al., 2020a) presenting in both comparisons reasonable correlations with a coefficient of determination, R^2, above 0.7. In a recent study, Bedoya-Velásquez et al. (2019) performed a validation of MWR data comparing them with 5 years of radiosonde data at Granada-Spain. Such analysis demonstrated a very low bias in MWR profiles respects radiosoundings, being this bias from 1.8 to −0.4 K with and standard deviation of 1.1 K for the temperature profiles and from 3.0 to −4.0% with and standard deviation around 135 for the humidity profiles, under all-weather conditions and below 2 km a.g.l..

Additionally, from the MWR potential temperature (θ) profiles, a feature to describe the atmospheric stability (A_tS_t) has been defined. Using the comparison criterion presented in Moreira et al. (2020), where each θ profile is classified as convective, the A_tS_t categorical feature has been obtained being $A_tS_t = 0$ for convective situations and $A_tS_t = 1$ for stable cases.

The initial Dataset is presented in Table 1. Hourly averages of all the relevant variables for the period 2015-2017 have been obtained from their original database, except for the values of H and S which were included as categorical variables. Additionally, continuous variables have been normalized with respect to their mean values, in order to homogenize their ranges of variability. Although this is not a required process in GBRT, Krishnamurthy et al. (2021) have pointed out slight improvements in ABLH detection, mainly at nighttime, when this normalization is applied. This final dataset was splitted in two subsets: (1) a subset with measurements recorded in the period 2015-2016 that will be used for the model set-up and training, and (2) a validation subset, composed by measurements taken along 2017.

3.2.2 Feature selection

In order to verify the relevance of each feature and to avoid data redundancy, as well as excessive complexity in the model, a selection of the most relevant features from the initial dataset has been performed (Guyon and Elisseeff, 2003). To this aim, the importance of each feature has been analyzed from two criteria. A first criterion, namely the Boruta algorithm, estimates the importance of each feature by comparing its influence on the predicted value with that of its randomly shuffled copies (Kursa et al, 2010). The second criterion, known as Recursive Feature Elimination (RFE), trains a predetermined model starting with all features in the training dataset, and after each iteration discards the least important features and refits the model (Yu and Liu, 2003). In this study, the variables that after being discarded did not cause a 2% reduction in coefficient of determination (R^2) were removed.

Both criteria have been applied on the entire database but also a specific feature importance analysis has been performed in order to account for possible differences in the feature relevance between day- and night-time situations. Figure 2 shows the relative importance of each feature, during day (a) and night (b), so that as higher the value obtained, greater is the influence of this variable on the results provided by the ML model. For daytime

$ABLH_{CEIL}$ and G appear as the most relevant features while T, RH, NR, WS, P and WS show a lower relevance and are sorted differently by each criteria. In the case of nighttime data, the most relevant feature is the hour (H), which explains how the model can identify nighttime situations, while the $ABLH_{CEIL}$ takes the second position and remains as one of the most important features for the model. On the opposite extreme A_tS_t, S, SZA, kt have been classified as irrelevant features. This result can be explained by the correlation of these variables with some of the features classified as relevant. Thus, i.e., all the near-surface meteorological features selected as relevant present some seasonal dependence making the use of the parameter S redundant. Similarly, A_tS_t, appears in both cases, nighttime and daytime, as one of the less relevant features. In the case of our location, this is explained because nighttime/daytime classification is mostly equivalent to a stable/convective classification, making the variable A_tS_t a redundant input. Thus, in a deep analysis of the entire database no stable cases during the daytime while the 95.5% of nighttime cases are convective. These irrelevant (A_tS_t, S, SZA, kt) features have not been included as input in the final GBRT model in order to avoid redundancy in the dataset.

3.2.3 Hyperparameters

GBRT algorithm requires a thorough setup of the so-called hyperparameters (parameters that cannot be updated during the training process) in order to avoid overfitting in the training dataset. The most relevant hyperparameters involved in the GBRT proposed in this study are: (1) the maximum depth of each tree, which represents the maximum number of leaves in each tree, (2) the maximum number of features, which indicate the maximum number of features inputted in each tree, (3) the learning rate, which indicates the influence of the previous decision-trees on its successors, and (4) the minimum sample leaf, which represents the minimum number of samples required to be at a leaf node in the tree.

In this study, the hyperparameters of the baseline model have been obtained from a large group of values randomly selected over our setup-training subset, over which a cross validation and Bayesian optimization processes (Frazier, 2018) have been applied using the Python library Scikit-learn (Pedregosa et al., 2011). Then, an empirical fine-tuning was performed in order to detect the values that provide the best results. From this analysis, the most suitable value for the maximum depth of each tree has been estimated as 5 while for the maximum number of features a value of 4 has been selected. These low values of the hyperparameters contribute to reducing the potential overfitting. A low value has been also obtained for the learning rate (0.0573), which ensures the improvement of the correction under ceilometer data during stable periods. The optimal minimum sample leaf value was indicated as 3, avoiding higher values of this parameter that can generate greater smoothing in the predicted values.

3.3 Model training

Once the inputs and hyperparameters have been determined, the GBRT algorithm has been trained (stage where the model is fitted) and tested (stage where the model performance is analyzed in terms of accuracy/precision). As indicated in Section 2, this training has been performed using a two-year dataset (2015-2016) with 5.153 cases. In order to reduce possible bias, the k-fold cross-validation methodology (James et al. 2013) has been applied. In this methodology, the dataset is randomly shuffled and divided into k parts, approximately equal. Then, k iterations are performed and, in each one of them, one group is selected as a test while the others k-1 are used for training. After k iterations, the chosen performance parameters obtained from each iteration and mean absolute error are averaged, and such values are considered as the performance parameters of the model. In this work k is

5 and, consequently, in each iteration an 80% and 20% of the data subset were employed for training and testing, respectively. Figure 3 illustrates this process.

In the training stage, the model reached a R^2 of 0.97, which indicates a satisfactory performance and that the overfitting was avoided. The Mean Absolute Error (MAE) obtained was 127 m. During the test stage, although a reduction of around 20% in R^2 (0.76) was observed, the variation of MAE was lower than -2%, resulting in 129 m.

3.4 Analysis

The GBRT algorithm proposed in this study has been validated using data recorded in our station along the entire 2017. Thanks to that, different aspects have been analyzed. On one hand, the general performance of the algorithm has been assessed analyzing the temporal and seasonal variability of the Mean Relative Error (MRE) among the $ABLH_{GBRT}$ and $ABLH_{MRW}$ values. This statistic quantifies the mean relative deviation between the target value ($ABLH_{MRW}$) and that one provided by the model ($ABLH_{GBRT}$). The MRE has been estimated by the following equation:

$$MRE_{GBRT}(\%) = 100 \cdot \sum \left(\frac{ABLH_{GBRT} - ABLH_{MRW}}{ABLH_{MRW}} \right) \quad (2)$$

This statistic has been also calculated for the $ABLH_{CEIL}$ values in order to assess the improvement of the algorithm proposed in this study with respect to the use of the ceilometer alone.

The statistical analysis has been completed with the estimations of the relative Root Mean Squared Error (rRMSE) defined as:

$$rRMSE_{GBRT}(\%) = 100 \cdot \sqrt{\frac{1}{N} \sum \left(\frac{ABLH_{GBRT} - ABLH_{MRW}}{ABLH_{MRW}} \right)^2} \quad (3)$$

where n is the number of samples.

In order to identify possible limitations of the proposed algorithm under different atmospheric conditions, cloudless, stable and convective situations have been differentiated and the MRE values for these situations have been analyzed. Day- and nighttime have been separated in terms of the solar zenith angle values (SZA), with SZA < 80º for daytime and SZA > 100º for nighttime. As mentioned above, because of the results of the convective/stable analysis performed from the A_tS_t feature, in this study nighttime is equivalent to stable and daytime is equivalent to convective situations. Additionally, cloudy and cloudless conditions have differentiated. In this study, clouds have been detected from the intensity of the RCS measured by the ceilometer, which notably increases in presence of clouds over the instrument. Clouds are detected when the RCS reaches values above 107, which is the empirical threshold estimated for our station as representative of cloud presence (Moreira et al., 2020a). Day- and nighttime have been separated in terms of the solar zenith angle values (SZA), with SZA < 80º for daytime and SZA > 100º for nighttime. As mentioned above, because of the results of the convective/stable analysis performed from the A_tS_t feature, in this study nighttime is equivalent to stable and daytime is equivalent to convective situations.

Moreover, to the MRE values, the analysis of these situations has been performed analyzing R^2 with respect to the reference measurements and defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \hat{x})^2} \qquad (4)$$

where n number of samples, $x_i$is the reference values ($ABLH_{MWR}$), $\hat{x}_i$ is the estimated value ($ABLH_{GBRT}$) and, $\hat{x}$ is the average of the reference values ($ABLH_{MWR}$).

Finally, the $ABLH_{CEIL}$, $ABLH_{GBRT}$ and $ABLH_{MWR}$ were intercompared for three days endowed with specific atmospheric situations: a) cloudless day, b) low-cloud day, and c) a day under the influence of a Saharan dust outbreak. These situations have been chosen due to the limitations observed in the ceilometer estimations of the ABLH, mainly under low-cloud scenarios (Coen et al., 2014) and decoupled aerosol layers (Caicedo et al., 2017), and will contribute to analyze the improvement of the methodology proposed in this study.

4 Results

4.1 General performance

4.1.1 Temporary and seasonal variability

Figure 4 presents the MRE hourly-averaged $ABLH_{CEIL}$ and $ABLH_{GBRT}$ values for all analyzed cases. Both $ABLH_{GBRT}$ and $ABLH_{CEIL}$ overestimate the $ABLH_{MWR}$ values although with notably lower percentages in the case of the GBRT estimations (note the different scales in Figure 4). Thus, the MRE_{GBRT} values (black line) do not exceed 36% while MRE_{CEIL} (magenta line) always has values higher than 30%. In both cases, the MRE values present a diurnal pattern with their lowest values during the first hours in the afternoon (0% for MRE_{GBRT} vs. 30% for MRE_{CEIL}) and their highest values during nighttime (around a 30% for MRE_{GBRT} and up to a 200% for MRE_{CEIL}). The higher differences between $ABLH_{CEIL}$ and $ABLH_{MWR}$, observed during the night and early morning, occur because the $ABLH_{MWR}$ estimates the SBLH, while the $ABLH_{CEIL}$ detects the RLH (Moreira et al., 2020). From these results, it is possible to observe the possibility of estimating the SBLH from ceilometer data in combination to surface meteorological information. In addition, from the combination of gradient method and GBRT is possible to detect SBLH, RLH and CBLH.

Figure 5 presents a comparison among the hourly averaged $ABLH_{GBRT}$ (black line), $ABLH_{MWR}$ (red line), and $ABLH_{CEIL}$ (magenta line) values. These plots show the expected pattern with lower ABLH values from sunrise to sunset and higher values during daytime, following a delayed solar cycle pattern (Moreira et al., 2020a). For all seasons, the $ABLH_{GBRT}$ and $ABLH_{MWR}$ values present high similarity, especially between 09 to 18 UTC, being all $ABLH_{MWR}$ values within the $ABLH_{GBRT}$ error range (grey shadow). The R^2 values between $ABLH_{GBRT}$ and $ABLH_{MWR}$ in each season are always greater than or equal 0.88, in contrast to the seasonal R^2 values between $ABLH_{CEIL}$ and $ABLH_{MWR}$, which are always lower than 0.30. Such values occur because during nighttime and early morning notable differences are observed between $ABLH_{CEIL}$ and $ABLH_{MWR}$ values because the methods based on the gradient of aerosol concentration tend to monitor the RLH in these situations. However, in the central hours of the day $ABLH_{CEIL}$ and $ABLH_{MWR}$ estimates the CBLH. On the other hand, the GBRT method is well trained to detect the SBLH, in a similar way as the $ABLH_{MWR}$ detects it. As the central hours of the day approach, the difference is gradually reduced, being minimal at the point where the maximum height of the ABLH is reached.

Table 2 summarizes the rRMSE values of the GBRT for all cases, as well as for each season, for both day- and nigh-time (convective/stable) situations. These results confirm the good performance of the model proposed here with an average rRMSE of 20% for all cases. Summer is the season with the lowest values of rRMSE for both day- and night-time. The rest of the seasons show similar behavior with rRMSE ranging from a 14% for daytime in Spring and 26% for nighttime in Autumn.

The daily patterns of the hourly averaged MRE_{GBRT} and MRE_{CEIL} values, per season, are shown in Fig. 6. The scales evidence that in all cases MRE_{CEIL} are larger than MRE_{GBRT}, mainly during night and early morning, as expected, due to differences in ABLH definition considered by each algorithm. In the case of MRE_{GBRT} similar patterns for the different seasons have been found although with larger errors (in the range 15% to 45%) between 19 - 08 UTC while lower errors (< 15%) occur between 09 to 18 UTC. The highest MRE_{GBRT} values are observed in autumn (Fig. 6d), while the lowest values are estimated in summer (Fig. 6c). On the other hand, the MRE_{CEIL} values show a seasonal pattern, with values higher than 140% between 19 - to 08 UTC, for all seasons, excluding summer (Fig. 5c). A result that is associated with the detection of the RL top height by the ceilometer processing (Moreira et al., 2020a). Between 09 to 18 UTC (period predominantly convective), the MRE_{CEIL} has low values, underestimating the $ABLH_{MWR}$ in some situations (13 to 14 UTC in summer and 14 UTC in autumn). The highest MRE_{CEIL} values are observed in winter (Fig. 6a), while the lowest ones occur in summer (Fig. 6c).

4.1.2 Dependence on atmospheric/meteorological conditions

Table 3 summarizes the R^2 and MRE_{GBRT} values for cloudless, stable and convective situations. In general, cloudless cases present only a variation of around a 1% for the R^2 and MRE_{GBRT} values with respect to the all-sky situations, indicating a low dependence of the GBRT algorithm on cloudiness. Such a result is in accordance with low relative importance of kt presented in section 3.2.2. When stable and convective cases are not differentiated, the GBRT model shows its highest values of R^2=0.91) and MRE_{GBRT} values are around 20%. When convective and stable cases are differentiated, the highest R^2 values (0.89) and the lowest values of MRE_{GBRT} (11%) were observed during the convective periods. In daytime situations, and mainly under cloudless conditions, the top of the aerosol layer coincides with the CBL height due to the absence of the RL (e.g. Eresma et al., 2006; Caicedo et al., 2017; Moreira et al., 2020a). Stable cases show a slightly lower performance, with $R^2 = 0.75$ and MRE_{GBRT} values around 28%, where the lower SBLH values are partially responsible of the rather large MRE_{GBRT}.

4.2 Case studies

4.2.1 Case 1: A clear-sky day (24th January 2017)

Figure 8 shows the evolution of the ABLH for a clear-sky day, characterized by the absence of low clouds and a thick and well-defined aerosol layer. It is observed that $ABLH_{CEIL}$ (magenta stars) represents the RL, from the beginning of measurement, until around 09:00 UTC. Thus, as CBL begins to increase, the difference between $ABLH_{CEIL}$ and $ABLH_{MWR}$ decreases, so that they are coincident at 10:00 UTC, and have a difference lower than 350 m between 11:00 and 18:00 UTC. At 19:00 UTC the $ABLH_{MWR}$ presents values clearly decoupled of the top of the aerosol layer ($ABLH_{CEIL}$) detecting the SBLH, consequently the differences between $ABLH_{CEIL}$ and $ABLH_{MWR}$ increase, reaching the maximum at 23:00 UTC (around 1200 m).

For its part, the $ABLH_{GBRT}$ values (black stars) show a very similar behavior with respect to the $ABLH_{MWR}$ values (red stars). Their higher agreement occurs between 04:00 to 06:00 UTC and 18:00 to 22:00 UTC, showing that the GBRT model provides appropriate estimates of the SBLH in the presence of the RL. The highest differences between $ABLH_{MWR}$ and $ABLH_{GBRT}$ are observed between 10:00 and 17:00 UTC, nevertheless, they are always lower than 100 m.

4.3.2 Case 2: A day with presence of low clouds (7th February 2017)

Figure 9 shows a case with the presence of low clouds (altitude < 2000 m), which can directly influence the ABLH detection when using the gradient method with the ceilometer data (Moreira et al., 2020). From 01:00 to 06:00 UTC due to low RL height, $ABLH_{MWR}$ and $ABLH_{CEIL}$ present, in general, a similar behavior with differences lower than 300 m. At 07:00 UTC, when the first clouds appear, the gradient method tends to overestimate the $ABLH_{MWR}$ (which is situated at 410 m), increasing the $ABLH_{CEIL}$ values up to 1500 m (cloud base). Due to the presence of low clouds throughout the day, the $ABLH_{CEIL}$ is estimated at the cloud base overestimating the $ABLH_{MWR}$, mainly during the stable period, where the difference between them reaches up to 2000 m at 23:00 UTC. Similar results were observed by Coen et al. (2014) and Caicedo et al. (2017).

In the case of the $ABLH_{GBRT}$ values, from 01:00 to 07:00 UTC $ABLH_{GBRT}$ and $ABLH_{MWR}$ are almost coincident (differences lower than 20 m). However, between 08:00 and 16:00 UTC the $ABLH_{GBRT}$ overestimates the $ABLH_{MWR}$, so that the maximum difference (300 m) is observed at 12:00 UTC. Due to the presence of clouds, radiative cooling occurs in the region near the base of the cloud, affecting the temperature profile and, consequently, decreasing the $ABLH_{MWR}$. On the other hand, the ML model is a combination of a group of variables, which are not totally affected by the clouds, therefore higher ABLH values are estimated. From 18:00 to 22:00 UTC the $ABLH_{GBRT}$ underestimates the $ABLH_{MWR}$, with maximum difference (-100 m) being observed at 20:00 UTC. Despite $ABLH_{GBRT}$ values present differences between -100 and 300 m during the cloudy period compared to the MWR estimations, such results demonstrate a remarkable improvement in these situations in comparison $ABLH_{CEIL}$, which is strongly affected by low clouds.

4.3.3 Case 3: Sahara dust outbreak (21st February 2017)

Estimating ABLH during dust outbreaks is a challenge to methods based on stand-alone vertical aerosol profiles, mainly when the dust layer is advected in the ABL region (e.g. Granados-Muñoz et al., 2012; Bravo-Aranda et al., 2017). In these situations, methods to estimate the ABLH based on stand-alone vertical aerosol profiles are unable to distinguish aerosol layering and the ABLH values tend to be overestimated (Granados-Muñoz et al., 2012), requiring the use of more sophisticated methods such as POLARIS based on depolarization measurements (Bravo-Aranda et al., 2017). However, due to the technical limitation of ceilometers (described previously in section 1), applications of techniques that require the use of more than one wavelength cannot be applied. Resulting in a lack of studies that address the detection of ABLH from ceilometer data during dust outbreaks.

Figure 10 presents the second day of an extreme Sahara dust outbreak registered over the Iberian Peninsula from 20 to 23 February 2017, resulting in values of aerosol optical depth around 2.3 (at 675 nm) in Granada (Fernández et al., 2019). These values correspond to level 2.0 data provided by AERONET.

In this situation, notable problems can be observed in the $ABLH_{CEIL}$ values. Thus, from 01:00 to 12:00 UTC, $ABLH_{CEIL}$ is estimated on the top of the dust layer due to the high gradient between this layer and FT. Consequently, the $ABLH_{CEIL}$ is overestimated, mainly between 01:00 and 07:00 UTC. Between 13:00 and 14:00 UTC, due to a reduction in the height of the aerosol layer, $ABLH_{CEIL}$ is situated close to the $ABLH_{MWR}$, underestimating it by around 200 m. From 15:00 UTC until the end of day, the height of the aerosol layer increases again, so that $ABLH_{CEIL}$ returns to overestimate the $ABLH_{MWR}$, resulting in a maximum difference of 2400 m at 23:00 UTC.

Despite the complexity of the situation, $ABLH_{GBRT}$ values present a very high agreement with respect to the $ABLH_{MWR}$. A slight overestimation occurs from 02:00 to 14:00 UTC (less than 100 m) and after 19:00 UTC (with maximum at 22:00 UTC of 200 m). On the other hand, $ABLH_{GBRT}$ underestimate the $ABLH_{MWR}$ from 15:00 and 18:00 UTC, so that the maximum difference (-100 m) is observed at 16:00 UTC.

These results confirm the possibility of estimating reliable ABLH during cases of dust outbreaks using a ceilometer combined with near-surface meteorological data as input of the machine learning algorithm proposed in this study.

5 Conclusions

A new methodology to estimate the Atmospheric Boundary Layer Height (ABLH), detecting the Stable Boundary Layer Height (SBLH) in stable cases, based on the machine learning algorithm known as Gradient Boosting Regression Tree (GBRT) has been proposed. This algorithm uses as features (independent variables) estimations of the ABLH derived applying the gradient method to a ceilometer signal ($ABLH_{CEIL}$) and several surface meteorological variables. The target (reference) ABLH values in this study have been those estimated from a microwave radiometer ($ABLH_{MWR}$). A detailed study of the features and the hyperparameters involved in the model set-up have been developed in order to avoid the model overfitting and guarantee its good performance during the training (R^2 = 0.97; MAE = 127 m) and test (R^2 = 0.76; MAE = 129 m) stage.

The proposed new algorithm has been validated using the entire year 2017. The model performance analysis has shown a daily pattern in the MRE_{GBRT} values, with their highest values during the night-time (stable situations) and their lower values along the day-time (convective situations). Minimum differences between $ABLH_{GBRT}$ and $ABLH_{MRW}$ appears, during the central hours of the day and first hours in the afternoon, when the ABL presents is higher height. This pattern has been observed for all seasons with MRE_{GBRT} ranging between -5% and 35%. A remarkable improvement is observed with respect to the MRE_{CEIL} values, which show similar daily patterns but range between 36% and 190%.

The new model has been analyzed under different atmospheric conditions revealing no dependence of the algorithm on cloudiness conditions. Small differences have been observed between stable and convective situations. Thus, while MRE_{GBRT} is around 11% in convective situations, these values increase up to 28% in the case of stable situations. Nevertheless, for both cases R^2 values are above 0.75 for stable and convective atmospheres and take a value of 0.91 when all conditions are considered. These results confirm the robustness of the GBRT algorithm presented in this study.

Three particular cases, namely a clear-sky day, a day with presence of low-clouds and dust outbreak event, have been chosen due to analyze and overcome the limitations observed in the $ABLH_{CEIL}$, particularly in the dust-outbreak events for which the gradient method is highly inefficient. In general, in these three particular situations $ABLH_{GBRT}$ shows very similar values and behavior than the $ABLH_{MRW}$, confirming the good model performance and a remarkable improvement with respect to the $ABLH_{CEIL}$ in complex situations, and enabling the SBLH detection. Therefore, the combination of GBRT and gradient method enables the detection of SBLH, CBLH and

RLH from ceilometer data together with surface meteorological information. Such results can be easily applied to well-established ceilometers networks over the world, just adding low cost surface meteorological sensors, which typically are available in these stations.

Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness through projects CGL2015-73250-JIN, CGL2016-81092-R, CGL2017-83538-C3-1-R ,CGL2017-90884-REDT and PID2020-120015RB-I00 and by the University of Granada through "Plan Propio. Programa 9 Convocatoria 2013. The financial support for EARLINET in the ACTRIS Research Infrastructure Project by the European Union's Horizon 2020 research and innovation program through project ACTRIS-2 (grant agreement No 654109). The authors thankfully acknowledge the FEDER program for the instrumentation used in this work and the University of Granada that supported this study through the Excellence Units Program. COST Action TOPROF (ES1303), supported by COST (European Cooperation in Science and Technology), is also acknowledged.

References

Alados-Arboledas, L., Müller, D., Guerrero-Rascado, J.L., Navas-Guzmán, F., Pérez-Ramírez, D., Olmo, F.J., 2011. Optical and microphysical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star-and sun-photometry. Geophysical Research Letters, 38 (1), https://doi.org/10.1029/2010GL045999.

Agencia Estatal de Meteorología (AEMET), 2015. Calendario Meteorológico 2015. Información meteorológica y climatológica de España. Ministerio de Agricultura. Alimentación y Medio Ambiente. ISSN: 02 1 3-3849.

Allabakash, S., and Lim, S., 2020. Climatology of Planetary Boundary Layer Height-Controlling Meteorological Parameters Over the Korean Peninsula. Remote Sens., 12, 2571; doi:10.3390/rs12162571.

Baars, H., Ansmann, A., Engelmann, R. and Althausen, D., 2008. Continuous monitoring of the boundary-layer top with lidar. Atmospheric Chemistry and Physics, 8, 7281–7296. https://doi.org/10.5194/acp-8-7281-200.

Baturynska, I., Martinsen, K., 2021. Prediction of geometry deviations in additive manufactured parts: comparison of linear regression with machine learning algorithms. J. Intell. Manuf. 32, 179–200. https://doi.org/10.1007/s10845-020-01567-0.

Bedoya-Velásquez, A.E., Navas-Guzmán, F., de Arruda Moreira, G., Román, R., Cazorla, A., Ortiz-Amezcua, P., Benavent-Oltra, J.A., Alados-Arboledas, L., Olmo-Reyes, F.J., Foyo-Moreno, I., Montilla-Rosero, E., Hoyos, C.D., Guerrero-Rascado, J.L., 2019. Seasonal analysis of the atmosphere during five years by using microwave radiometry over a mid-latitude site. Atmos. Res. 218, 78–89.

Blanco-Muriel, M., Alarcón-Padilla, D.C., López-Moratalla, T., Lara-Coira, M., 2001. Computing the solar vector, Solar Energy, 70 (5), 431-441, https://doi.org/10.1016/S0038-092X(00)00156-0.

Bravo-Aranda, J.A.; de Arruda Moreira, G.; Navas-Guzmán, F.; Granados-Muñoz, M.J.; Guerrero-Rascado, J.L.; Pozo-Vázquez, D.; Arbizu-Barrena, C.; Olmo Reyes, F.J.; Mallet, M.; Alados-Arboledas, L., 2017. A new methodology for PBL height estimations based on lidar depolarization measurements: analysis and comparison against MWR and WRF model-based results. Atmospheric Chemistry and Physics, 17, 6839-6851.

Bonin, T. A., Carroll, B. J., Hardesty, R. M., Brewer, W. A., Hajny, K., Salmon, O. E., & Shepson, P. B., 2018. Doppler lidar observations of the mixing height in Indianapolis using an automated composite fuzzy logic approach. Journal of Atmospheric and Oceanic Technology, 35(3), 473-490.

Bruine, M.D.; Apituley, A.; Donovan, D.P.; Baltink, H.K., 2017. Pathfinder: Applying graph theory for consistent tracking of daytime mixed layer height with backscatter lidar. Atmos. Meas. Tech., 10, 1–26.

Cadeddu, M.P., D.D. Turner, and J.C. Liljegren., 2009. A neural network for real-time retrievals of PWV and LWP from Arctic millimeter-wave ground-based observations. IEEE Trans. Geosci. Remote Sens., 47, 1887-1900.

Caicedo, V., Rappenglück, B., Lefer, B., Morris, G., Toledo, D., Delgado, R., 2017. Comparison of aerosol lidar retrieval methods for boundary layer height detection using ceilometer aerosol backscatter data, Atmos. Meas. Tech., 10, 1609–1622, https://doi.org/10.5194/amt-10-1609-2017.

Cariñanos, P., Foyo-Moreno, I., Alados, I., Guerrero-Rascado, J.L., Ruiz-Peñuela, S., Titos, G., Cazorla, A., Alados-Arboledas, L., Díaz de la Guardia, C., 2021. Bioaerosols in urban environments: Trends and interactions with pollutants and meteorological variables based on quasi-climatological series. Journal of Environmental Management, 282, 111963, https://doi.org/10.1016/j.jenvman.2021.111963.

Cazorla, A., Olmo, F.J., Alados-Arboledas, L., 2008. Using a sky imager for aerosol characterization**.** Atmospheric Environment 42 (11), 2739-2745.

Cazorla, A., JE Shields, ME Karr, FJ Olmo, A Burden, L Alados-Arboledas, 2009. Determination of aerosol optical properties by a calibrated sky imager. Atmospheric Chemistry and Physics 9 (17), 6417-6427.

Cazorla, A., Casquero-Vera, J. A., Román, R., Guerrero-Rascado, J. L., Toledano, C., Cachorro, V. E., Orza, J. A. G., Cancillo, M. L., Serrano, A., Titos, G., Pandolfi, M., Alastuey, A., Hanrieder, N., and Alados-Arboledas, L., 2017. Near-real-time processing of a ceilometer network assisted with sun-photometer data: monitoring a dust outbreak over the Iberian Peninsula, Atmos. Chem. Phys., 17, 11861–11876, https://doi.org/10.5194/acp-17-11861-2017.

Chen, F., Kusaka, H., Bornstein, R., Ching, J., Grimmon, C.S.B., Grossman-Clarke, S., Loridan, T., Manning, K.W., Martilli, A., Miao, S., Sailor, D., Salamanca, F.P., Taha, H., Tewari, M., Wang, X., Wyszogrodzki, A.A., Zhang, C., 2011. The integrated WRF/ urban modelling system: development, evaluation, and applications to urban environmental problems. Int. J.Climatol. 31 (2), 273–288.

Chen, T., Singh, S., Taskar, B., Guestrin, C., 2015. Efficient second-order gradient boosting for conditional random fields. In Proceeding of 18th Artificial Intelligence and Statistics Conference (AISTATS'15), vol. 1.

Cimini, D., De Angelis, F., Dupont, J.-C., Pal, S., Haeffelin, M., 2013. Mixing layer height retrievals by multichannel microwave radiometer observations, Atmos. Meas. Tech., 6, 2941–2951, https://doi.org/10.5194/amt-6-2941-2013.

Coen, M.C., Praz, C., Haefele, A., Ruffieux, D., Kaufmann, P., Calpini, B., 2014. Determination and climatology of the planetary boundary layer height above theSwiss plateau by in situ and remote sensing measurements as well as by the COSMO-2model. Atmos. Chem. Phys. 14, 13205–13221, https://doi.org/10.5194/acp-14-13205-2014.

Emeis, S., Schäfer, K., Münkel, C., 2008. Surface-based remote sensing of the mixing-layer height – a review. Meteorologische Zeitschrift, 17, 621–630.https://doi.org/10.1127/0941-2948/2008/031.

Eresmaa, N., Karppinen, A., Joffre, S. M., Räsänen, J., Talvitie, H., 2006. Mixing height determination by ceilometer, Atmos. Chem. Phys., 6, 1485–1493, https://doi.org/10.5194/acp-6-1485-2006.

Fernández, A.J., Sicard, M., Costa, M.J., Guerrero-Rascado, J.L., Gómez-Amo, J.L., Molero, F., Barragán, R., Basart, S., Bortoli, D., Bedoya-Velásquez, A.E., Utrillas, M.P., Salvador, P., Granados-Muñoz, M.J., Potes, M., Ortiz-Amezcua, P., Martínez-Lozano, J.A., Artíñano, B., Muñoz-Porcar, C., Salgado, R., Román, R., Rocadenbosch, F., Salgueiro, V., Benavent-Oltra, J.A., Rodríguez-Gómez, A., Alados-Arboledas, L., Comerón, A., Pujadas, M., 2019. Extreme, wintertime Saharan dust intrusion in the Iberian Peninsula: Lidar monitoring and evaluation of dust forecast models during the February 2017 event. Atmospheric Research, 228, pp. 223-241. DOI: 10.1016/j.atmosres.2019.06.007.

Flamant, C., Pelon, J., Flamant, P.H., Durand, P., 1997. Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer. Boundary-Layer Meteorol, 83, 247–284.

Frazier, P.I., 2018. A Tutorial on Bayesian Optimization. arXiv:1807.02811v1.

Friedman, J., 2001. Greedy boosting approximation: a gradient boosting machine. Ann. Stat. 29, 1189–1232. https://doi.org/10.1214/aos/1013203451.

Geiß, A., Wiegner, M., Bonn, B., Schäfer, K., Forkel, R., von Schneidemesser,E., Münkel, C., Chan, K.L. and Nothard, R., 2017. Mixing layer height as an indicator for urban air quality? Atmospheric Measurement Techniques, 10,2969–2988. https://doi.org/10.5194/amt-10-2969-2017.

Georgoulias, A.K., Papanastasiou, D.K., Melas, D., Amiridis, V., Alexandri, G., 2009. Statistical analysis of boundary layer heights in a suburban environment. Meteorol Atmos Phys 104, 103–111. https://doi.org/10.1007/s00703-009-0021-z.

Granados-Muñoz, M. J., Navas-Guzmán, F., Bravo-Aranda, J. A., Guerrero-Rascado, J. L., Lyamani, H., Fernández-Gálvez, J., and Alados-Arboledas, L., 2012. Automatic determination of the planetary boundary layer height using lidar: One-year analysis over southeastern Spain, J. Geophys. Res.-Atmos., 117, D18208, https://doi.org/10.1029/2012JD017524.

Guerrero-Rascado, J.L., Ruiz, B., Alados-Arboledas, L., 2008. Multi-spectral Lidar characterization of the vertical structure of Saharan dust aerosol over southern Spain. Atmospheric Environment, 42 (11), pp. 2668-2681. DOI: 10.1016/j.atmosenv.2007.12.062.

Guerrero-Rascado, J. L., Olmo, F. J., Avilés-Rodríguez, I., Navas-Guzmán, F., Pérez-Ramírez, D., Lyamani, H., and Alados Arboledas, L., 2009. Extreme Saharan dust event over the southern Iberian Peninsula in september 2007: active and passive remote sensing from surface and satellite, Atmos. Chem. Phys., 9, 8453–8469, https://doi.org/10.5194/acp-9-8453-2009.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157-1182.

Haefele, A., Hervo M., Turp, M., Lampin, J.-L., Haeffelin, M., Lehmann, V., 2016. E-PROFILE team, TOPROF team. The E-PROFILE network for the operational measurement of wind and aerosol profiles over Europe. Teco 2016 - Madri (Spain). Available at: https://www.eumetnet.eu/wp-content/uploads/2016/10/E-PROFILE_TECO_Madrid_2016.pdf.

Haeffelin, M., Angelini, F., Morille, Y., Martucci, G., Frey, S., Gobbi, G.P.,Lolli, S., O'Dowd, C.D., Sauvage, L., Xueref-Rémy, I., Wastine, B., Feist, D.G., 2012. Evaluation of mixing-height retrievals from automatic profil-ing lidars and ceilometers in view of future integrated networks in Europe.Boundary-Layer Meteorology, 143, 49–75. https://doi.org/10.1007/s10546-011-9643-z.

Helmis, C.G.; Sgouros, G.; Tombrou, M.; Schäfer, K.; Münkel, C.; Bossioli, E.; Dandou, A., 2012. A Comparative Study and Evaluation of Mixing-Height Estimation Based on Sodar-RASS, Ceilometer Data and Numerical Model Simulations. Bound.-Layer Meteorol., 145, 507–526.

Iqbal M., 1983. An Introduction to Solar Radiation, Academic Press, New York.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning with Application in R (Springer Texts in Statistics) 7th ed. Springer New York Heidelberg Dordrecht London.

Jiang, R., Zhao, K., 2021. Using machine learning method on calculation of boundary layer height. Neural Comput & Applic. https://doi.org/10.1007/s00521-021-05865-3.

Jiang, Y., Xin, J., Zhao, D., Jia, D., Tang, G., Quan, J., Wang, M., Dai, L., 2021. Analysis of differences between thermodynamic and material boundary layer structure: Comparison of detection by ceilometer and microwave radiometer. Atmospheric Research, 248, 105179, https://doi.org/10.1016/j.atmosres.2020.105179.

Krishnamurthy, R., Newsom, R. K., Berg, L. K., Xiao, H., Ma, P.-L., Turner, D. D., 2021. On the estimation of boundary layer heights: A machine learning approach, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2020-439.

Kursa, M.B., Jankowski, A., Rudnicki, W.R., 2010. Boruta - A System for Feature Selection. Fundamenta Informaticae, 101, 271–285. https://doi.org/10.3233/FI-2010-288.

Lee, J., Hong, J.W., Lee, K., Hong, J., Velasco, E., Lim, Y.J., Lee, J.B., Nam, K., Park, J., 2019. Ceilometer Monitoring of Boundary-Layer Height and Its Application in Evaluating the Dilution Effect on Air Pollution. Boundary-Layer Meteorol 172, 435–455. https://doi.org/10.1007/s10546-019-00452-5.

Li, P., Wu, Q., Burges, C. J., 2008. Mcrank: Learning to rank using multiple classification and gradient boosting. In Advances in Neural Information Processing Systems 20, pages 897–904.

Liu, B., Ma, Y., Gong, W., Yang, J., Zhang, M., 2018. Two-wavelength Lidar inversion algorithm for determining planetary boundary layer height. J. Quant. Spectrosc. Radiat. Transf., 206, 117–124.

Manninen, A. J., Marke, T., Tuononen, M. J., O'Connor, E. J., 2018. Atmospheric boundary layer classification with Doppler lidar. Journal of Geophysical Research: Atmospheres, 123, 8172– 8189. https://doi.org/10.1029/2017JD028169.

Marques, M. T. A., Moreira, G. de A., Pinero, M., Oliveira, A. P., Landulfo, E., 2018. Estimating the planetary boundary layer height from radiosonde and doppler lidar measurements in the city of São Paulo - Brazil. EPJ WEB OF CONFERENCES, v. 176, p. 06015. https://doi.org/10.1051/epjconf/201817606015.

McGovern, A., Elmore, K.L., Gagne, D.J., Haupt, S.E., Karstens, C.D., Lagerquist, R., Smith, T., Williams, J.K., 2017. Using artificial intelligence to improve real-time decision-making for high-impact weather. Bulletin of the American Meteorological Society, 98(10), pp.2073-2090.

Moreira, G.A., Guerrero-Rascado, J.L., Bravo-Aranda, J.A., Benavent-Oltra, J.A., Ortiz-Amezcua, P., Róman, R., Bedoya-Velásquez, A.E., Landulfo, E., Alados-Arboledas, L., 2018. Study of the planetary boundary layer by microwave radiometer, elastic lidar and Doppler lidar estimations in Southern Iberian Peninsula. Atmospheric Research, 213, 185-195. https://doi.org/10.1016/j.atmosres.2018.06.007.

Moreira, G.A., Guerrero-Rascado, J.L., Benavent-Oltra, J.A., Ortiz-Amezcua, P., Román, R., Bedoya-Velásquez, A.E., Bravo-Aranda, J.A., Olmo-Reyes, F.J., Landulfo, E., Alados-Arboledas, L., 2019. Analyzing the turbulent planetary boundary layer by remote sensing systems: the Doppler wind lidar, aerosol elastic lidar and microwave radiometer. ATMOSPHERIC CHEMISTRY AND PHYSICS (ONLINE), v. 19, p. 1263-1280.

Moreira, G.A., Guerrero-Rascado, J.L., Bravo-Aranda, J.A., Foyo-Moreno, I., Cazorla, A., Alados, I., Lyamani, H., Landulfo, E., Alados-Arboledas, L., 2020a. Study of the planetary boundary layer height in an urban environment using a combination of microwave radiometer and ceilometer. Atmospheric Research, 240, 104932, https://doi.org/10.1016/j.atmosres.2020.104932.

Moreira, G. A., Da Silva Lopes, F.J., Guerrero-Rascado, J.L., Ortiz-Amezcua, P., Cazorla, A., De Oliveira, A.P., Landulfo, E., Alados-Arboledas, L., 2020b. Comparison Among the Atmospheric Boundary Layer Height Estimated From Three Different Tracers. EPJ WEB OF CONFERENCES, v. 237, p. 03009, https://doi.org/10.1051/epjconf/202023703009.

Moreira, G.A., Andrade, I.S., Cacheffo, A., Yoshida, A.C., Gomes, A.A., Silva, J.J., Lopes, F.J.S., Landulfo, E., 2021. COVID-19 outbreak and air quality: Analyzing the influence of physical distancing and the resumption of activities in São Paulo municipality, Urban Climate, Volume 37, 100813, ISSN 2212-0955, https://doi.org/10.1016/j.uclim.2021.100813.

Morille, Y., Haeffelin, M., Drobinski, P., Pelon, J., 2007. STRAT: An Automated Algorithm to Retrieve the Vertical Structure of the Atmosphere from Single-Channel Lidar Data, Journal of Atmospheric and Oceanic Technology 24, 5 , 761-775, https://doi.org/10.1175/JTECH2008.1.

Müller, A.C., Guido, S., 2016. Introduction to machine learning with Python. O'Reilly Media, Inc., Sebastopol.

National Research Council. 2009. Observing Weather and Climate from the Ground Up: A Nationwide Network of Networks. Washington, DC: The National Academies Press. https://doi.org/10.17226/12540.

Ortiz-Amezcua, P., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., Benavent-Oltra, J. A., Böckmann, C., Samaras, S., Stachlewska, I. S., Janicka, Ł., Baars, H., Bohlmann, S., Alados-Arboledas, L., 2017. Microphysical characterization of long-range transported biomass burning particles from North America at three EARLINET stations, Atmos. Chem. Phys., 17, 5931–5946, https://doi.org/10.5194/acp-17-5931-2017.

Pal, S., Haeffelin, M., Batchvarova, E., 2013. Exploring a geophysical process-based attribution technique for the determination of the atmospheric boundary layer depth using aerosol lidar and near-surface meteorological measurements. Journal of Geophysical Research: Atmospheres, 118, 9277–9295.https://doi.org/10.1002/jgrd.50710.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A, Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12, pp. 2825-2830.

Poltera, Y., Martucci, G., Collaud Coen, M., Hervo, M., Emmenegger, L., Henne, S., Brunner, D., Haefele, A., 2017. PathfinderTURB: an auto-matic boundary layer algorithm. Development, validation and application to study the impact on in situ measurements at the Jungfraujoch. Atmospheric Chemistry and Physics, 17, 10051–10070. https://doi.org/10.5194/acp-17-10051-2017.

Rey-Sanchez, C., Wharton, S., Vilà-Guerau de Arellano, J., Paw U, K. T., Hemes, K. S., Fuentes, J. D., Osuna, J., Szutu, D., Ribeiro, J. V., Verfaillie, J., Baldocchi, D., 2021. Evaluation of atmospheric boundary layer height from wind profiling radar and slab models and its responses to seasonality of land cover, subsidence, and advection. Journal of Geophysical Research: Atmospheres, 126, e2020JD033775. https://doi.org/10.1029/2020JD033775.

Schween, J.H., Hirsikko, A., Löhnert, U., Crewell, S., 2014. Mixing-layer height retrieval with ceilometer and Doppler lidar: From case studies to long-term assessment. Atmos. Meas. Tech., 7, 4275–4319.

Stachlewska, I.S., Migacz, S., Szkop, A., Zielínska, A.J., Swaczyna, P.L., 2012. Ceilometer observations of the boundary layer over Warsaw, Poland. Acta Geophys. 60, 1386–1412.

Stull, R. B., 1988. An Introduction to Boundary Layer Meteorology, 666 pp., Kluwer Acad., Dordrecht, Netherlands.

Toledo, D., Córdoba-Jabonero, C., Adame, J.A., Benito, D.L.M., Gil-Ojeda, M., 2017. Estimation of the atmospheric boundary layer height during different atmospheric conditions: A comparison on reliability of several methods applied to lidar measurements. Int. J. Remote Sens., 38, 3203–3218.

Uzan, L., Egert, S., Khain, P., Levi, Y., Vadislavsky, E., Alpert, P., 2020. Ceilometers as planetary boundary layer height detectors and a corrective tool for COSMO and IFS models, Atmos. Chem. Phys., 20, 12177–12192, https://doi.org/10.5194/acp-20-12177-2020.

Vassallo, D., Krishnamurthy, R., Fernando, H. J. S., 2021. Utilizing physics-based input features within a machine learning model to predict wind speed forecasting error, Wind Energ. Sci., 6, 295–309, https://doi.org/10.5194/wes-6-295-2021.

Vivone, G., D'Amico, G., Summa, D., Lolli, S., Amodeo, A., Bortoli, D., Pappalardo, G., 2021. Atmospheric boundary layer height estimation from aerosol lidar: a new approach based on morphological image processing techniques, Atmos. Chem. Phys., 21, 4249–4265, https://doi.org/10.5194/acp-21-4249-2021.

Wang, W., Mao, F., Gong, W., Pan, Z., Du, L., 2016. Evaluating the Governing Factors of Variability in Nocturnal Boundary Layer Height Based on Elastic Lidar in Wuhan. International Journal of Environmental Research and Public Health. 13(11):1071. https://doi.org/10.3390/ijerph13111071.

Ye, J., Chow, J.H., Chen, J., Zheng, Z., 2009. Stochastic gradient boosted distributed decision trees. In Proceedings of the 18th ACM conference on information and knowledge management (pp. 2061–2064), ACM. https://doi.org/10.1145/1645953.1646301.

Yu, L., Liu, H., 2003. Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution. Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003), Washington DC.

Table 1. Group of input variables initially considered along with the instrument used to measured them and the range of variation of each variable during the period of study.

Instrument/Algorithm	Variable	Range
Ceilometer/Gradient Method	$ABLH_{CEIL}$	200 m - 4500 m
MWR/Gradient and Parcel Method	$ABLH_{MWR}$	200 m - 4500 m
HMP60	Temperature (T)	0 - 42 ºC
	Hour (H)	Categorical Variable
	Season (S)	Categorical Variable
	Stability (A_tS_t)	Categorical Variable (0-1)
HMP60	Relative Humidity (RH)	4.7 - 91 %
Barometer PTB110	Pressure (P)	920 - 952 hPa
Anemometer 05103	Wind Speed (WS)	0 - 5 m/s
Pyranometer CM-11	Global Radiation (G)	0 - 1016 W/m²
Pyrgeometer PIR	Net Radiation (NR)	-167 - (-1) W/m²
Blanco-Muriel et al. (2001)	Solar Zenith Angle (SZA)	14 - 165 º
Iqbal (1983)	Clearness Index (k_t)	0 - 1

Table2. rRMSE of the GBRT for all cases, as well as for each season, for both day- and nigh-time (convective/stable) situations.

	All cases		Winter		Spring		Summer		Autumn	
	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
$rRMSE_{GBRT}$ (%)	15	25	18	24	14	25	10	20	15	26

Table 3. MRE and R² of the GBRT algorithm under cloudless and all-cloud-type conditions. Additionally, for each category, stable, convective and all-stability conditions have been differentiated. Number of cases on each category have been included in order to prove their representativeness.

	All Cases			Cloudless Cases		
	Stable	Convective	All cases	Stable	Convective	All cases
Number of cases	1284	1579	2863	398	600	998
R^2	0.75	0.89	0.90	0.75	0.89	0.91
MRE_{GBRT} (%)	28.0	11.1	20.9	27.5	10.2	19.1

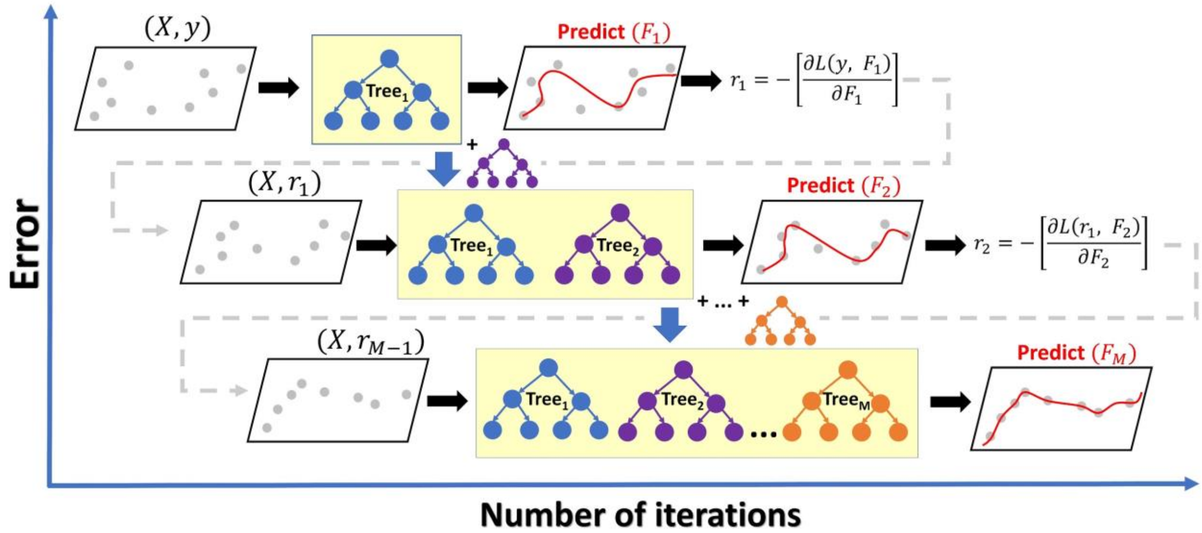


Figure 1 - GBRT flowchart. X, y, and L represent the feature matrix, target variable, and loss-function, respectively. r_N and F_N indicate de n-nth pseudo-residual error and prediction.

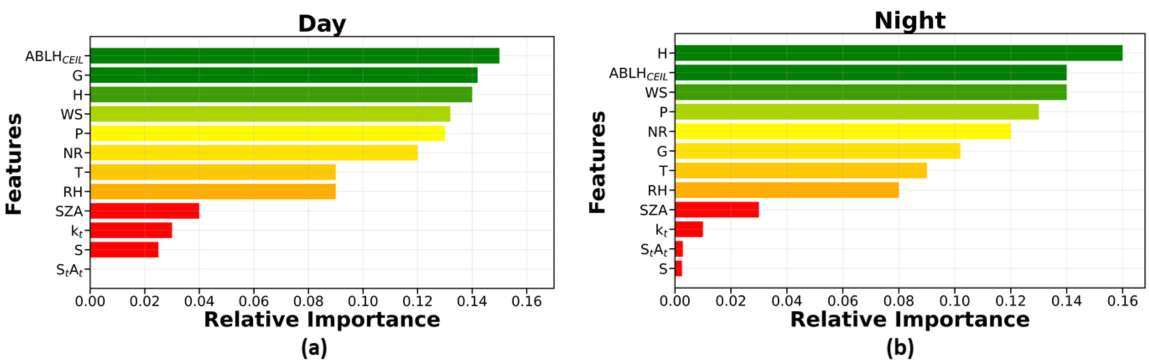


Figure 2 - Feature relative importance classification/ranking applying the Recursive Feature Elimination (RFE) method, for day (a) and night (b) situations.

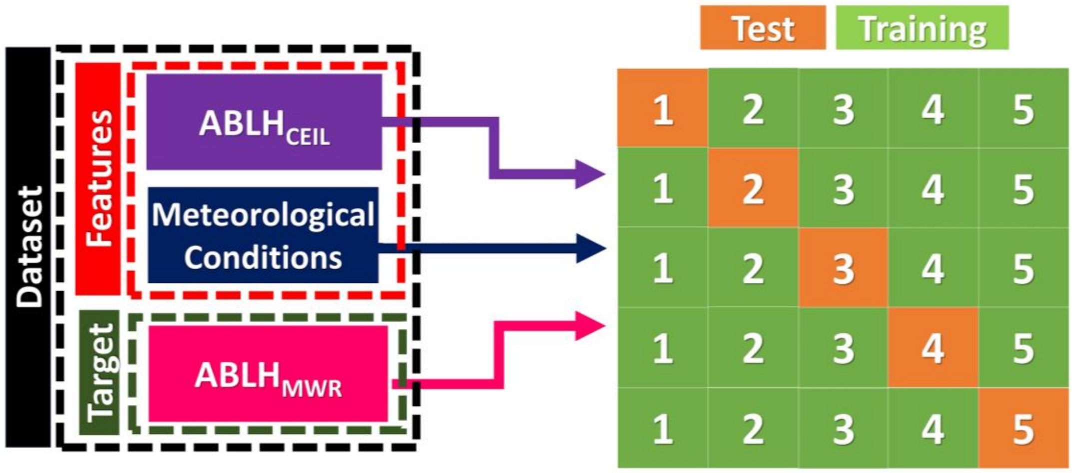


Figure 3 - Scheme of input dataset (left) and k-fold cross-validation methodology (right).

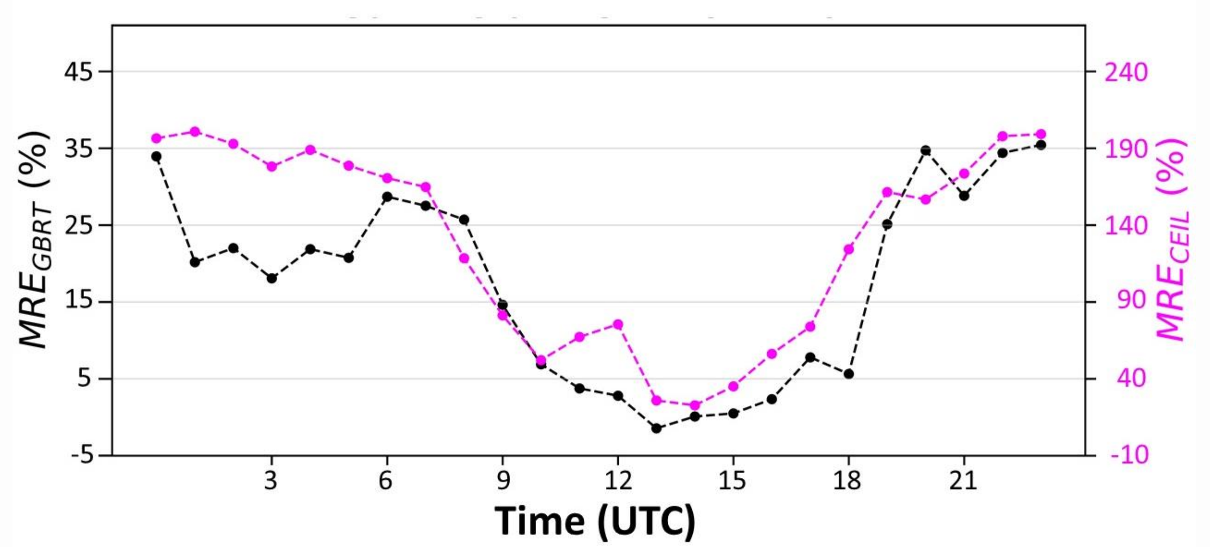


Figure 4 – Hourly Mean Relative Error for all the analyzed cases applied in the GBRT algorithm (black) and the gradient method to the ceilometer data (pink). It should be highlighted the important difference between the scales required for each methodology.

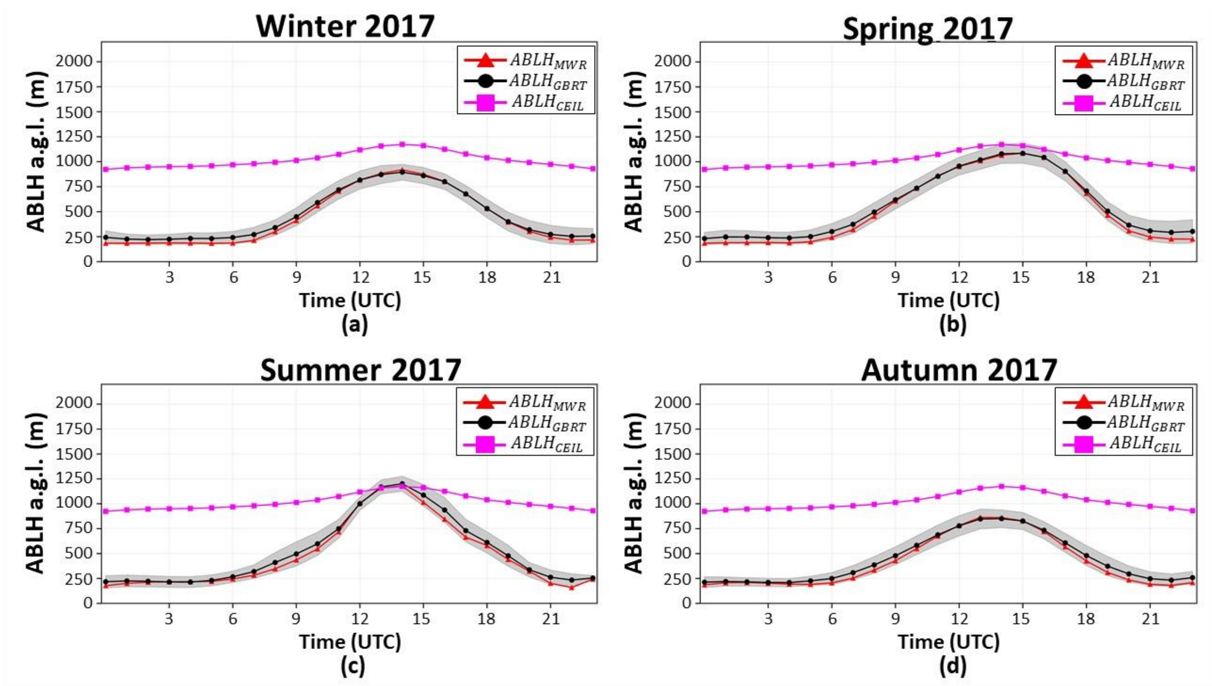


Figure 5 - Comparison between the hourly ABLH average measured (red line) and those predicted by the GBRT algorithm (black line) and the ceilometer (pink) for (a) winter, (b) spring, (c) summer, and (d) autumn. The dark shadow represents the GBRT model standard deviation.

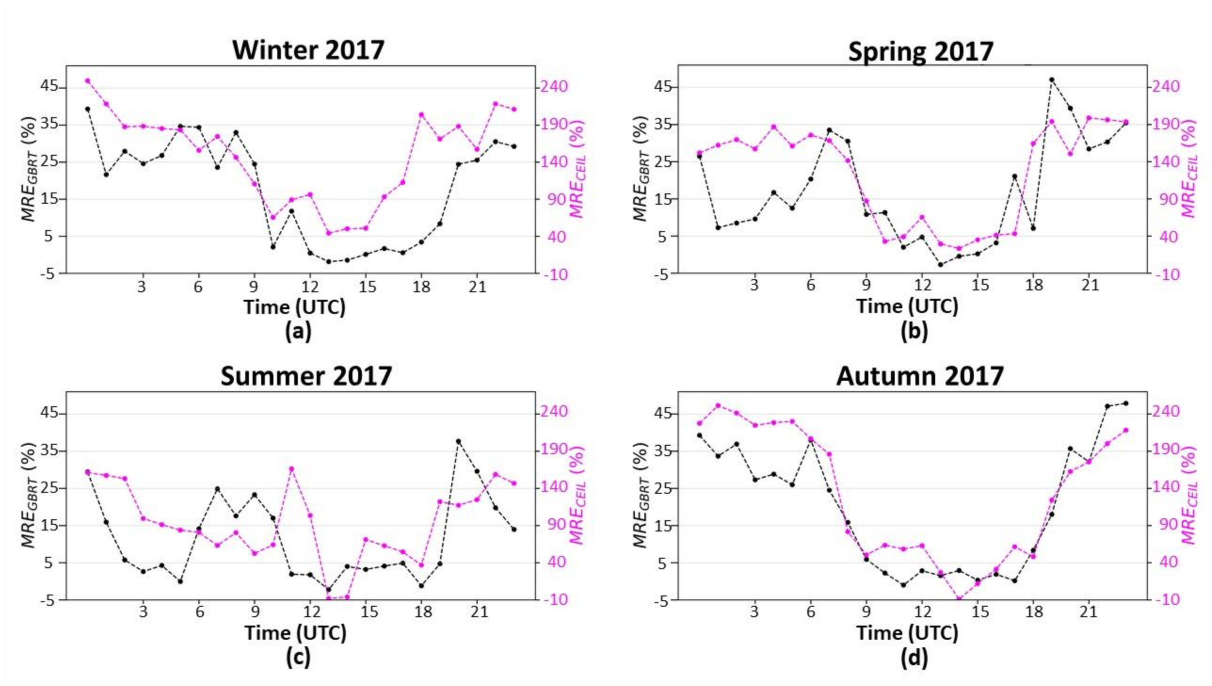


Figure 6 - Hourly Mean Relative Error for all the analyzed cases applied in the GBRT algorithm (black) and the gradient method to the ceilometer data (pink) during (a) winter, (b) spring, (c) summer, (d) autumn.

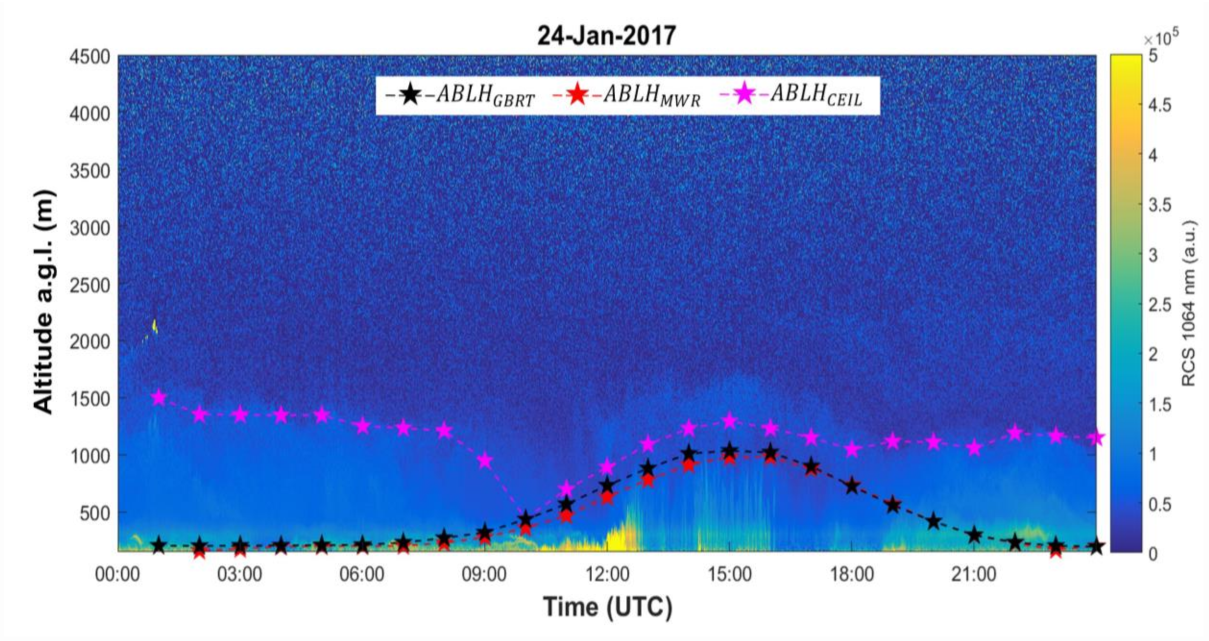


Figure 8 - Comparison among the hourly values of $ABLH_{GBRT}$ (black stars), $ABLH_{MWR}$ (red stars) and of $ABLH_{CEIL}$ (pink stars) at January 24, 2017.

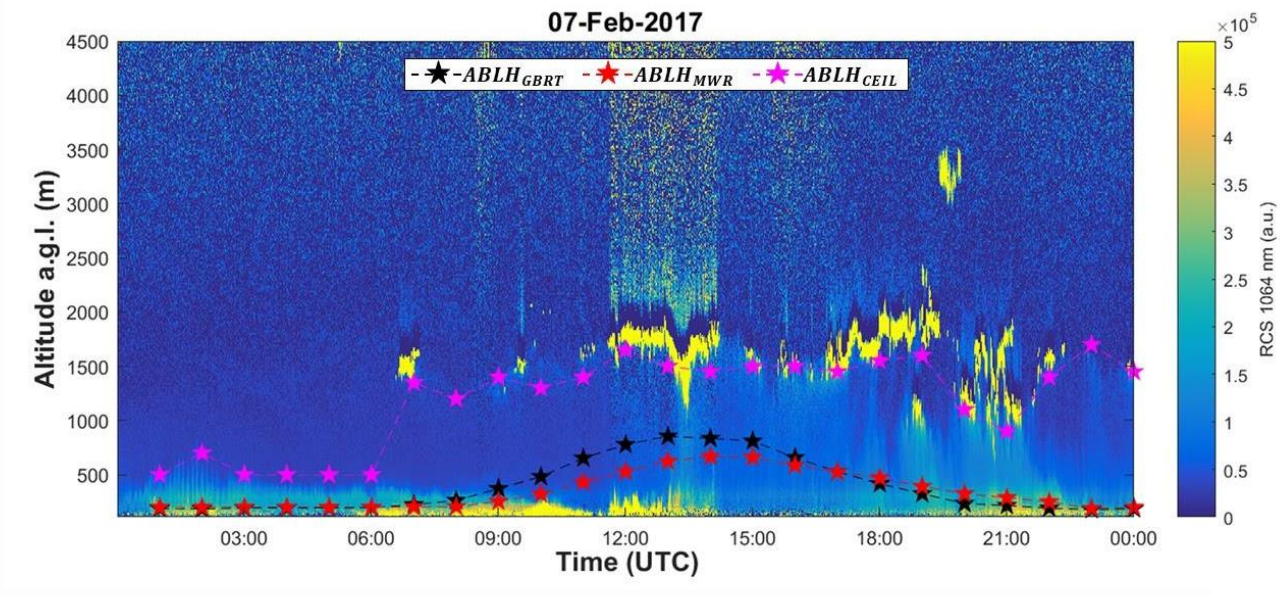
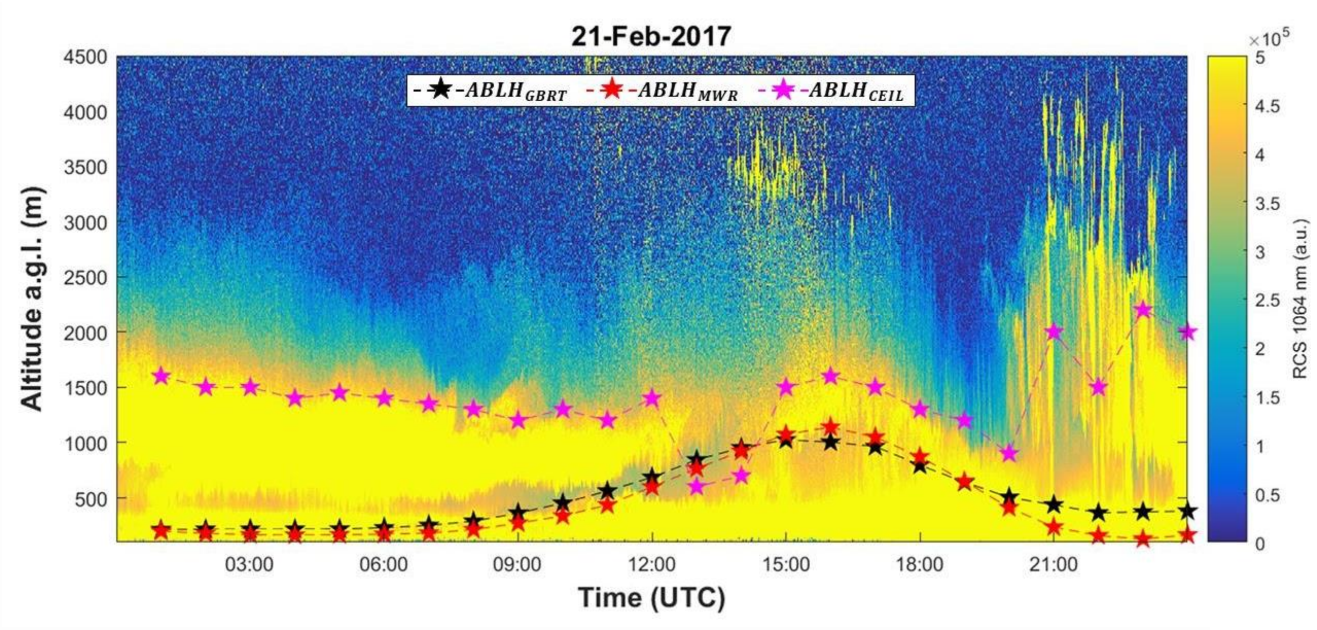


Figure 9 - Comparison among the hourly values of $ABLH_{GBRT}$ (black stars), $ABLH_{MWR}$ (red stars) and $ABLH_{CEIL}$ (pink stars) at February 7, 2017.

787



788 Figure 10 - Comparison among the hourly values of $ABLH_{GBRT}$ (black stars), $ABLH_{MWR}$ (red stars) and $ABLH_{CEIL}$
789 (pink stars) at February 21, 2017.
790