**PathfinderTURB: an automatic boundary layer algorithm.**

Development, validation and application to study the impact on in-situ measurements at the Jungfraujoch.

Yann Poltera\(^1,3\), Giovanni Martucci\(^1\), Martine Collaud Coen\(^1\), Maxime Hervo\(^1\), Lukas Emmenegger\(^2\), Stephan Henne\(^2\), Dominik Brunner\(^2\) and Alexander Haefele\(^1\)

\(^1\)Federal Office of Meteorology and Climatology, MeteoSwiss, Payerne, Switzerland

\(^2\)Swiss Federal Laboratories for Materials Science and Technology, Dübendorf, Switzerland

\(^3\)Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

**Abstract.** Continuous observations of the vertical structure of the planetary boundary layer are invaluable for the validation of atmospheric transport models on the micro and meso scale. Lidar and ceilometer backscatter observations offer a robust technique with growing spatial coverage, but the obtained backscatter profiles need to be carefully translated into boundary layer parameters. Here we present the development of the PathfinderTURB algorithm for the analysis of ceilometer backscatter data and the real-time detection of the vertical structure of the planetary boundary layer. Two typical aerosol layer heights are retrieved by PathfinderTURB: the Convective Boundary Layer (CBL) and the Continuous Aerosol Layer (CAL). PathfinderTURB combines the strengths of gradient- and variance-based methods and addresses the layer attribution problem by adopting a geodesic approach. The algorithm has been applied to one year of data measured by two CHM15k ceilometers operated at the Aerological Observatory of Payerne (491 m, a.s.l.) on the Swiss plateau, and at the Kleine Scheidegg (2061 m, a.s.l.) in the Swiss Alps. The retrieval of the CBL has been validated at Payerne using two reference methods: (1) manual detections of the CBL height performed by independent human experts using the ceilometer backscatter data of the year 2014; (2) values of CBL heights calculated using the Richardson’s method from co-located radio sounding data. We found average biases as small as 27 m (53 m) with respect to reference method 1 (2). Based on the excellent agreement with the two reference methods, PathfinderTURB has been applied to the ceilometer data at the mountainous site of the Kleine Scheidegg for the period September 2014 till November 2015. At this site, the CHM15k is operated in a novel, tilted configuration at 71° zenith angle to probe the air next to the Sphinx Observatory (3580 m, a.s.l.) on the Jungfraujoch (JFJ). The analysis of the retrieved layers led to the following results: the CAL reaches the JFJ during 41% of the time in summer and during 21% of the time in winter for a total of 97 days during the two seasons. The season-averaged daily cycles show that the CBL height reaches the JFJ only during short periods (4% of the time) on 20 individual days in summer and never during winter. Especially during summer the CBL and the CAL modify the air sampled in-situ at JFJ, resulting in an unequivocal dependence of the measured absorption coefficient on the height of both layers. This highlights the relevance of retrieving the height of CAL and CBL in mountainous regions.

1 **Introduction**

During convective periods, particles and gases are mixed homogeneously within the convective boundary layer (CBL). The upper limit of the CBL corresponds to the interface between the well-mixed region and the free troposphere (FT) above it. This interface, also called entrainment zone (EZ), is a turbulent transition zone...
characterized by negative buoyancy flux. The description and detection of the EZ has drawn particular attention in studies about the CBL dynamics during the last decades. There are various methods to study the CBL and the EZ, based on profiles of temperature, backscatter or turbulence measured either by radio-sounding or by passive and active remote sensing or calculated by numerical models. Amongst the different observational methods, the remote sensing technique ensures the largest amount of profile data. Active remote sensing (acoustic or laser-based) provides the best vertical resolution allowing to resolve the multiple transitions (including the EZ) between different layers in the CBL and the FT. The instrument used for our study is a ceilometer, a low-power, compact and cost-effective version of a research LIDAR. A ceilometer is normally single-wavelength, and emits linearly-polarized laser light in the near-infrared spectral band (800-1100 nm), where the signal is highly sensitive to aerosols and cloud droplets. In the early 2000’s, the major manufacturers (e.g., Vaisala, Leosphere, MPL, Jenoptik) started producing ceilometers with the capability to store the full backscatter profile instead of the cloud base only. As a consequence, ceilometers have been rapidly recognized by the national meteorological services and research centres in Europe and worldwide as an efficient and affordable way to study the troposphere using aerosols as tracers (e.g., Münkel, 2007; Flentje et al., 2010; Martucci et al., 2010; Heese et al., 2010; Wiegner and Geiss, 2012; Wiegner et al., 2014). Ceilometers have increased significantly in number in Europe and the United States especially over the last decade, now reaching nearly 1000 units in Europe alone (http://www.dwd.de/ceilomap). If combined in a single large network, all ceilometers could provide helpful information on the vertical and horizontal distribution of aerosols and on the status of the CBL over a very large geographical domain in near real time.

In order to process automatically a large amount of data over a large and geographically-diverse domain, we need an algorithm capable of retrieving the vertical structure of the mixed layer (ML) during both convective (CBL) and stable (SBL) conditions and over both flat and complex terrain. The term “mixed” (different from “mixing”) indicates a layer in which the profiles of potential temperature and humidity do not vary much in height and the particles and gases are well-mixed, but are not necessarily still mixing. Whereas, within the CBL there is an ongoing active mixing normally related to the daytime updraft and downdraft cycle. Inside the ML (and to some extent also in the FT) several aerosol layers can be formed: an efficient retrieval method shall then solve the attribution problem (layer categorization), i.e. the unambiguous detections of the aerosol layers and the EZ. The attribution problem still remains one of the major sources of uncertainty affecting the ML height (MLH) retrieval. In order to respond to these requirements, we have further developed, validated and applied the pathfinder algorithm described by de Bruine et al. (2016) and optimized it for the real-time detection of the vertical structure of the CBL. The new version of pathfinder, called PathfinderTURB, is a gradient-based layer detection algorithm that, in addition to the traditional gradient detection, makes use of the aerosol distribution temporal variability (variance) to detect the MLH. The validated PathfinderTURB has been applied to the data of a ceilometer installed at the Kleine Scheidegg to probe the air that is sampled in-situ at the high Alpine station Jungfraujoch (JFJ). The fact of measuring the presence of the ML directly at the JFJ and in real time provides an unprecedented possibility to describe quantitatively the relation between the ML dynamics and the aerosols measured in-situ at JFJ. An additional motivation comes from the fact that the JFJ contributes with in-situ observations of greenhouse gases as a level-1 station to the ICOS (Integrated Carbon Observation System) infrastructure. At level-1 ICOS stations the detection of the CBL height in their vicinity is required to serve as a
validation dataset for atmospheric transport models. This is crucial when atmospheric concentration observations should be translated into greenhouse gas fluxes between the atmosphere and the land surface.

In Sect. 2, we present an overview of existing algorithms for the retrieval of the MLH from LIDAR backscatter profiles. The novel PathfinderTURB algorithm is then detailed in Sect. 3. In Sect. 4, we describe the configuration of the sites of Payerne (Swiss plateau) and Kleine Scheidegg (Swiss Alps) where ceilometer measurements are done. In Sect. 5, we present the validation of PathfinderTURB in Payerne against one year of human expert and co-located radiosonde reference MLH retrievals. In Sect. 6, the retrievals of PathfinderTURB from the recent ceilometer measurements at the Kleine Scheidegg are analysed and compared to the absorption coefficient at 637 nm at JFJ, highlighting the impact of both convective and injection layers on the in-situ measurements at the high Alpine station. Finally, conclusions are given in Sect. 7.

2 Overview of existing algorithms

Traditionally, the retrieval of the MLH from the backscatter profile of a LIDAR can be done using two types of methods: (i) the gradient-based algorithms that track gradients in the vertical distribution of aerosols (gradient of the backscatter profiles), (ii) and the variance-based algorithms that track fluctuations in the temporal distribution of aerosols (variance of the backscatter profiles). Some algorithms combine both techniques, which makes the MLH retrieval more robust especially in convective conditions when the ML dynamics change rapidly.

The gradient-based algorithms retrieve the MLH by tracking the well-marked drop in the aerosol concentration that often occurs at the base of (or within) the EZ in convective conditions or at the level of the temperature inversion capping the residual layer (RL), in neutrally-stratified conditions. All vertical negative gradients found starting from the ground are transitions between different aerosol layers and correspond to peaks along the LIDAR backscatter gradient profile. All peaks are labelled as possible candidates of the MLH (layer attribution) at each time step. The traditional approach using numerical approximations of the first or second derivatives of the LIDAR signal (e.g., Menut et al., 1999), has been improved by using the wavelet covariance transform and the fact that the strong gradient occurring at the top of a layer exists at both small and large scales allowing the wavelet-based methods to reduce the uncertainty when assigning the MLH (Davis et al., 2000; Cohn and Angevine, 2000; Brooks, 2003; Baars et al., 2008; Angelini et al., 2009; de Haij et al., 2010; Frey et al., 2010).

Alternatively, the Derivative of Gaussian (DOG) wavelet is used in Morille et al. (2007) or the Daubechies wavelets in Engelbart et al. (2008). The Canny edge detection method (Canny, 1986) also helps improving the retrieval of aerosol layers (e.g. STRAT2D: Morille and Haeffelin, 2010). Worth mentioning is also the method proposed in Steyn et al. (1999), which consists in fitting an idealized backscatter profile at the transition between the ML and the FT. In the more recent literature there are examples of different methods combining the LIDAR gradient-based retrievals with temporal height-tracking techniques, for example observational (Martucci et al., 2010a, b), predictive (Tomás et al., 2010) or model-based first guesses (Giuseppe et al., 2012), Pal et al. (2013), proposed a simplified bulk-model combined with surface turbulence measurements and atmospheric variance measurements, to help selecting the MLH amongst all candidates. Collaud Coen et al., (2014), used a gradient-based temporal continuity criterion to reduce the problem’s degeneration and improve the attribution skill. In the study described by de Bruine et al. (2016) presenting the pathfinder algorithm, the gradient field and guiding
restrictions are taken as core information to retrieve the MLH based on the identification of the most cost-effective path (geodesic) along the gradient lines in a graph.

The variance-based algorithms use the temporal fluctuations in the aerosol backscatter as a function of the height $z$ to retrieve the MLH. Within the EZ, cleaner, drier free tropospheric air is entrained repeatedly and mixed-in with the rising aerosol-laden, moister air coming from the ML. A variance-based algorithm can detect the MLH at the level where the backscatter variability reaches a maximum at the base or within the EZ. Variance-based algorithms calculate the temporal variance of the backscatter profile at each range bin, usually over periods shorter than one hour. Similarly to the gradient-based, the variance-based algorithms use peaks in the smoothed variance profile as candidates for the MLH (e.g., Hooper and Eloranta, 1986; Piironen and Eloranta, 1995; Menut et al., 1999; Martucci et al., 2007).

Because the transitions between different aerosol layers and between the ML and the FT are characterized by both a sharp gradient in aerosol concentration and by mixing of air through the interface, the variance- and gradient-based algorithms normally provide similar retrievals of the MLH. Still, the gradient-based and the variance-based algorithms have their specific advantages and disadvantages under different atmospheric conditions. Indeed, the depth and structure of the ML depend on non-linear interactions at different timescales, induced by mechanical and thermodynamic mixing. When retrieving the MLH it is then important to include in the algorithm more than one source of information (e.g. gradient, variance, a priori information) in order to account for the largest number of atmospheric conditions and then to minimize the attribution uncertainty. Combining the variance- and gradient-based methods allows to compare the two retrievals at each time step (Lammert and Bösenberg, 2006; Martucci et al., 2010a,b, Toledo et al., 2014). The retrieval method STRAT+ (Pal et al., 2013), based on STRAT2D (Morille and Haeffelin, 2010), uses the Canny edge detection applied to gradient profiles along with the information brought by the variance profiles and by the radiosoundings to detect the main MLH and internal boundaries as well as the growth rate.

3 PathfinderTURB

In order to retrieve the MLH operationally while minimizing the uncertainty due to the attribution problem and to assure the adaptability of the algorithm to diverse topographical conditions, we have extended the pathfinder algorithm proposed by de Bruine et al. (2016) by adding a variance criterion and the retrieval of the Continuous Aerosol Layer (CAL). The extended version PathfinderTURB has been applied to the ceilometer data at two sites: the Aerological observatory of MeteoSwiss at Payerne (PAY, 491 m a.s.l., 46.799° N, 6.932° E) and the Kleine Scheidegg (KSE, 2061 m, 46.547° N, 7.985° E). In the framework of de Bruine’s work, the pathfinder technique was applied to measurements at the tall-tower site of Cabauw in the Netherlands and successfully validated by radiosonde (RS) data. Compared to other algorithms, the advantage of the pathfinder (and of PathfinderTURB) is in the ability to solve directly the attribution problem by constructing a timeseries of boundary layer’s heights based on the geodesic approach between consecutive points minimizing a well-constructed cost function. Compared to pathfinder, PathfinderTURB uses the variance-based retrieval method in conjunction with the gradient-based to calculate the geodesic and to retrieve the boundary layer’s height while solving the attribution problem. The layers retrieved by PathfinderTURB are the CBL, retrieved only during daytime, and the CAL (defined in the next section).
3.1 Calculation of the Top of Continuous Aerosol Layer (TCAL)

The CAL is defined as the uninterrupted aerosol region along the backscatter profile starting from the ground and reaching the first discontinuity in the aerosol distribution. The top of the CAL (TCAL) is then defined as the height of the retrieved discontinuity. In terms of LIDAR backscatter profile, the CAL is the part of the profile that is not interrupted by a Rayleigh backscatter layer (purely molecular). The criteria to define the CAL are the following (see also Supplement S4):

1. The total (aerosol plus molecular) attenuated backscatter is larger than a threshold $Th$ that depends on the purely molecular backscatter profile at the ceilometer’s wavelength.

2. The signal-to-noise ratio (SNR) is larger than 0.6745.

Based only on condition 1 (signal condition), the TCAL is found at the height where the signal is larger than $Th$ and, over flat homogeneous terrain, it usually corresponds to the top of the RL during night and to the height of the CBL during day. In complex and mountainous terrain, during daytime, the TCAL corresponds rather to the top of the so-called injection layer. The injection layer was defined by Henne et al. (2004) as the layer formed by injections of CBL air to higher levels engendered by thermally-driven converging slope winds along the topography that is reaching higher than the average in-valley CBL. In contrast to the CBL, the injection layer is only sporadically mixed and indirectly connected to the surface. Condition 2 (SNR condition), imposes that the SNR is larger than the Gaussian-distributed 1-sigma value of the signal noise. In other words, because the background signal (dark current plus stray light) is range-independent and is considered to be Gaussian-distributed, the signal is considered noisy when it lies within the 50%-confidence interval of the background signal. The statistics of the background signal distribution are calculated in the far-range of the total signal. If the SNR condition is included, the retrieved TCAL can be shallower compared to when only the signal condition is taken into account. That happens especially during daytime when the SNR drops below the value 0.6745 at lower altitude compared with the night. In cases like this we cannot speak anymore of TCAL, but rather of maximum detected range. Depending on the sky conditions, if clouds are present, also the height of the first cloud layer detected by the ceilometer combined with the heights obtained by SNR and signal conditions determine the TCAL.

3.2 Selection criteria and the main assumption of PathfinderTURB

For a given day, the temporal evolution of the ceilometer signal is a matrix in time and space where each column represents a profile at time $t$ and constant range resolution. The noise level is calculated from the photon counting signal using the method described by Morille et al. (2007). The signal is binned at a resolution of (30 m, 1 min) at PAY and (45 m, 2min) at KSE, then smoothed in time and range. We provide here a description of the main selection criteria and the main assumptions on which PathfinderTURB is based. Further details of the algorithm, including the calculation of the atmospheric variability (signal variance at the micro-scale) and of the turbulence-enhanced zones, and the mathematical steps leading to the expressions of the measured variables is given in the Supplement (S1, S2).
3.2.1 Lower altitude limit

Close to the ground, for most of the industrial bistatic ceilometer, the overlap between the transmitter and receiver is missing or close to zero. In this region, called blind region, the returned signal is absent or extremely weak, dominated by the noise and it oscillates around zero. It is thus not possible to retrieve the CBL height (CBLH) in this region (low clouds or fog detections are however possible). Above this region, the overlap increases until becoming full and the noise component becomes negligible compared to the signal at least within the CBL. A positive gradient is then expected at the transition between the blind region and the region above. We thus define the lower altitude limit, \( \text{minH} \), as the first range where there is a transition from a zero to a positive gradient and we impose \( \text{minH} \) not to be higher than 350 m (where the overlap of the ceilometer is normally sufficiently large to allow physical measurements).

During the morning and until the end of the afternoon, we expect the CBLH to lie at least above the first region from the ground where turbulence is observed. Therefore, an additional inferior limit for altitude \( \text{minH}_{\text{TURB}} \) is set, which delimits the first appearance of turbulence, and that is derived from the calculation of the atmospheric variability. The lower altitude limit \( \text{minH} \) is replaced by \( \text{minH}_{\text{TURB}} \) whenever the latter is higher than the former. The selected minimum limit is called \( \text{liminf} \) in Figure 1.

3.2.2 Upper altitude limit

Different criteria are defined to calculate the upper altitude limit, \( \text{maxH} \). These criteria are based on the a priori knowledge of the climatological values of the CBLH at a specific site (climatological limit) and on the retrievals of different layers made before the CBLH retrieval. These retrievals are: the TCAL (which includes the SNR limitation), the cloud base height (cloud limit), and the mixing discontinuities (strong negative and positive gradients). The minimum altitude amongst the three determines the upper altitude limit, \( \text{limphys} \) in Figure 1.

Climatological limit

A climatological limit can be set based on visually-inspected ceilometer data from previous years and on model-simulated CBLH. The climatological limit depends on the site, and needs as input a maximal constant CBLH value for the early morning, the afternoon, and a maximal mean growth rate between the two periods. Hereafter, the early morning period is considered as the period starting at sunrise and ending 2.5 hours (PAY, 3h KSE) after sunrise. This period reflects a coarse estimate of the delay in convective plume formation and is considered constant through the year. The afternoon period is considered to end at sunset. For our study we used the limits 1500 m a.s.l, 3000 m a.s.l and 1 km/h for PAY and 3069 m a.s.l, 4069 m a.s.l and 1 km/h for KSE for morning maximum CBLH, afternoon maximum CBLH and maximum mean growth rate, respectively.

Cloud limit

Two types of clouds are considered: CBL clouds and non-CBL clouds. All cloud information (number of cloud layers, cloud bases, cloud depths) are provided by the ceilometer’s standard outputs. A CBL cloud is defined as a cloud detected by the ceilometer in the first (lower) layer, whose vertical depth is less than 500 m and whose top (cloud base + depth) is lower than the site-specific climatological CBLH limit set before. This criterion is purely...
mathematical as the cloud depth provided by the ceilometer just gives the depth of the not-totally attenuated part of the signal and not the real depth.

**Strong negative and positive gradients**

Strong positive or negative gradients indicate mixing discontinuities and can then be taken as limits of the CBLH. The sensitivity to positive gradients is higher than to negative gradients, because these indicate normally a change of layer from the FT to an aerosol layer or a cloud base. Strong negative gradients correspond to a signal drop of 25% (15% during the early morning period), whereas strong positive gradients correspond to a signal gain of 15% (5% during the early morning period).

### 3.2.3 Growth rate induced refinement

The validity of the limits is checked (e.g. the lower limit cannot be higher than the upper limit). Then the limits are recalculated backwards in time (i.e. from 23:59 to 00:00), taking into account a growth rate limited to \pm 0.625 m/s between two time steps (i.e. only jumps < 37.5 m at PAY and < 75 m at KSE in altitude are allowed between two time steps). This growth rate is larger than the climatological growth rate of 1 km/h, because it allows larger jumps over shorter time steps in order to account for the convective dynamics, e.g. updraft and downdraft cycle.

### 3.2.4 Weights

At each time step \( t \) and at each altitude \( z \), a weight is defined as the cost of passing by \((t, z)\), i.e. based on the values of the gradient and the variance at \((t, z)\), the weight says how much it costs to choose the CBLH at that altitude and time. The weights are calculated by PathfinderTURB as the product of the gradients weights and the variance weights. An offset is added to make the weights positive:

\[
weights = \log_{10}(weights_{\text{Grad}} + weights_{\text{Var}}) + \left[\min\left(\log_{10}(weights_{\text{Grad}} + weights_{\text{Var}})\right)\right]
\]

(1)

Where the minimum is taken for the whole day and all altitudes. The value of \( weights_{\text{Grad}} \) is given by the inverse negative of the gradient value \( \nabla S \). The weights corresponding to positive or zero values of \( \nabla S \) are set to 1000 times the largest weights of the inverse negative gradient values so that the cost of choosing a positive gradient is extremely high. The value of \( weights_{\text{Var}} \) is given by the inverse of the variance.

For KSE the weights are obtained by the same as equation (1), but without the contribution of the variance, because, due to the slant path and the higher noise already within the CBL, we preferred not to use the signal variance as a weight (statistically over the entire dataset the SNR < 3 already at 850 m a.g.l.).

\[
weights = \log_{10}(weights_{\text{Grad}}) + \left[\min\left(\log_{10}(weights_{\text{Grad}})\right)\right]
\]

(2)

### 3.2.5 Shortest path

From sunrise (midnight KSE) to the end of the time series, over consecutive windows of 30 minutes (overlapping at the first and last time steps), the shortest path in a graph (the geodesic in the metric space defined...
by the weights) are calculated with Dijkstra’s algorithm (Dijkstra, 1959), following the original method of de Bruine et al. (2016). The weighted graph is constructed from the matrix representation of the timeseries, the altitude restrictions and the weights. The graph only allows connections of one time step in the positive time direction and jumps of maximum 37.5 m (75 m KSE) amplitude in the altitude direction. A new shortest path starts where the previous shortest path ended. In case the previous shortest path failed, the time window is skipped, and the next shortest path starts at the range of the first local minimal weight. At the first time step (sunrise and midnight for PAY and KSE, respectively), the CBLH is set at the inferior graph limit and at the first local minimum weight at PAY and KSE, respectively. The CBLH time series calculated after sunset is discarded (for PAY).

3.2.6 Quality check

A binary quality index (0/1) is set for each time step of the retrieved CBLH time series. In case of rain or fog (information directly given by the ceilometer), the quality index is set to 0. Else, the ratio of the mean ceilometer signal over 150 m above the CBLH and the mean ceilometer signal over 150 m below the CBLH is calculated (the distance to the inferior graph limit is taken if it is less than 150 m). If this ratio is larger than 0.85 (i.e. the signal drop is less than 15%), the quality is set to 0, else it is set to 1. This quality check procedure is similar to what is done by de Haij et al., (2010).

3.3 Example of PathfinderTURB’s TCAL and CBLH calculation

For the retrieval of the CBLH, PathfinderTURB uses the backscatter profiles $S$ (Eq. 3, sect. 4.2) generated at each time step by the ceilometer. The retrieval’s procedure is the following: pre-processing of $S$, CBLH retrieval, and quality-check. In the pre-processing step, $\nabla \log_{10}(S)$, $\liminf$ and $\limphys$ are calculated. In step two, the timeseries of the range-restricted $\nabla \log_{10}(S)$ is transformed into a weighted graph and the CBLH is determined as the geodesic calculated using the Dijkstra’s algorithm (Dijkstra, 1959) within predefined successive subintervals during the temporal interval between sunrise and sunset. Finally, in step three, the quality of the CBLH retrieval is assessed based on the signal ratio criterion.

Based on the described retrieval of the geodesic in the metric space defined by the weights, PathfinderTURB follows the minima in the vertical aerosol gradients over a sequence of retrievals. The way the geodesic is calculated limits the problem of jumps between different layers when boundaries disappear or reappear. In this work, PathfinderTURB has been extended from the original pathfinder to use the signal variance in addition to the aerosol gradients. Based also on Equation (1), PathfinderTURB combines the information provided by the gradient with that provided by the variance. As a result, regions with high signal variance are favoured in the CBLH attribution process, which yields a physically more robust detection. This is accomplished by multiplying the gradient-based weights of the graph by the inverse of the atmospheric variance value. In this way, the influence of artificial and static aerosol gradients, present in some models of ceilometers and due to an incorrect overlap correction, is largely reduced. The different steps of the PathfinderTURB algorithm are illustrated in Figure 1, in four timeseries (panels a, b, c and d) for the case of 15 July 2014 in PAY.
In panel a), the logarithm of the range-corrected signal is displayed. The cloud base height (CBH) is directly provided by the ceilometer software and displayed in grey throughout all the panels. The TCAL is shown in green throughout all the panels. As already introduced above (Sect. 3.1), the TCAL is the combination of the altitudes determined by the signal and SNR conditions plus the height of the first cloud layer. The signal condition and the height of the first cloud layer play a critical role in this example. In panel (a), the CBL is restricted by the TCAL, but it can be restricted furthermore by additional physically meaningful altitude limitations (limphys), displayed in magenta. These limphys depend on climatology, the TCAL, the CBH and the height of strong negative and positive gradients (which indicate mixing discontinuities). Strong positive gradients can be observed in panel (a) at about 1500 m a.s.l. between 02:00 and 03:30 UTC; strong negative gradients occur at about 1750 m a.s.l. between 20:00 and 24:00 UTC.

In panel b), the variance is displayed. The variance is calculated using spectral analysis, more precisely is the result of integrating the spectrum of band-pass filtered one-hour-long $S$-timeseries at each altitude (as in Pal et al., 2013). The band-pass filter aims at removing mesoscale and noisy fluctuations so that only fluctuations due to short-lived aerosol load variability are taken into account (Supplement S2). The minimum altitude limitation (liminf) is calculated from there, and displayed in magenta. Indeed, the entrainment occurring at the top of the CBL prevents the CBLH to be located below this first region of enhanced variance.

In panel c), weights are displayed. Using the Dijkstra’s algorithm, PathfinderTURB calculates the CBLH as the most cost-efficient path from sunrise to sunset (black line), i.e. the CBLH-path tends to follow the deep-blue regions (low weights) in the timeseries of weights. The path can only follow the positive time direction and altitude changes are limited to 0.625 m/s. The weights are given by the product of the inverse magnitude of negative gradients and of the variance values (sect. 3.2.4). Thus, the top of the CBL, which is characterized by a drop in the aerosol concentration and high entrainment activity (and thus high variability), will most likely have a low weight.

In panel d), a final overview is given, with the retrieved CBLH (black line) displayed on top of the log10($S$) timeseries, together with the TCAL and the cloud bases.
Figure 1: Timeseries of the different processing steps of the PathfinderTURB algorithm for 15 July 2014 in PAY. Panel a) timeseries of log10(S) with the TCAL, the cloud bases and the superior limit (limphys) for altitude. Panel b) Atmospheric variance, cloud bases, limphys, inferior limit for altitude (liminf). Panel c) Weights, TCAL, cloud bases and retrieved MLH (geodesic from sunrise to sunset). Panel d) timeseries of log10(S) with the TCAL, the cloud bases and the retrieved MLH.

4 Remote sensing and in-situ observations at Payerne and Kleine Scheidegg

Two CHM15k ceilometers were installed at two different sites in Switzerland. The first at PAY in a rural and comparatively flat environment. PAY is equipped with numerous meteorological measurements allowing to interpret and to validate the measurements from the ceilometer. The most relevant measurements in the framework of the presented study are: a Raman LIDAR measuring humidity, temperature and backscatter at 355 nm and at the Raman-shifted wavelengths (Dinoev et al., 2013; Brocard et al., 2013), a wind profiler (Degreane, 1290 MHz), a RPG-HATPRO microwave radiometer measuring temperature and humidity using 14 channels
around the water vapour and oxygen absorption lines, the operational Meteolabor SRS-C34 radiosondes launched twice daily at 00 and 12 UTC (Philipona et al., 2013), the sun-photometers (direct, diffuse and global short and long wave, Vuilleumier et al., 2014) and the surface sensors of temperature and humidity.

A second instrument was installed at KSE, close to the JFJ. The JFJ is part of numerous global observation programs like GAW (Global Atmospheric Watch), EMEP (European Monitoring and Evaluation Programme), NDACC (Network for the Detection of Atmospheric Composition Change) and AGAGE (Advanced Global Atmospheric Gases Experiment). Most importantly, in the context of this work, JFJ participates with in-situ observations as a level-1 station in the ICOS project. In contrast to other ICOS sites located over flat terrain, it was decided to install the ceilometer at KSE to characterize the CBL below and above the JFJ. The presence of aerosols, detected by the ceilometer, and the frequency at which these reach the JFJ, can be directly compared with the chemical and physical in-situ measurements of aerosols and trace gases at the JFJ. In fact, several in-situ instruments are installed at the JFJ and operate continuously since many years to measure optical and chemical properties of aerosols and trace gases as well as diverse meteorological parameters (Bukowiecki et al., 2016).

Instruments of direct interest to our study are: a condensation particle counter (CPC; TSI Inc., Model 3772), which measures the particle number concentration and two instruments providing aerosol absorption coefficients: a Multi-Angle Absorption Photometer (MAAP) measuring at 637 nm and an aethalometer (AE-31, Magee Scientific) measuring at seven different wavelengths. The aerosol in-situ measurements are performed under dry conditions (relative humidity <20 %), while the ceilometer probe the unmodified atmospheric volume.

### 4.1 Sites description

The PAY site (Figure 2) is situated in the centre of the Swiss Plateau between the Jura Mountain range to the north-west (at a distance of 25 km) and the Alpine foothills to the south-east (20 km). The measurement site is characterized by a rural environment leading to biogenic aerosols sources combined with moderate urban emissions characterized by anthropogenic aerosol sources especially related to cars exhaust and house heating.

The KSE (Figure 2, KSE) is located in the Bernese Oberland Alpine region, at an altitude of 2069 m. It can be seen as a saddle point between the mountain peak Lauberhorn (2472 m) to the northwest and the Jungfrau (3465 m) to the southeast, and it is a pass between the villages of Wengen and Grindelwald. This topographic configuration has a considerable influence on the local wind circulation. Winds at the KSE are mostly blowing along the SW-NE axis (Ketterer et al., 2014), whereas the prevailing wind at JFJ are from NW toward SE. The JFJ itself is located on the ridge formed between the Mönch and the Jungfrau mountains and is 4.5 km to the south-east and 1.5 km higher than KSE. Most of the atmospheric observations at the JFJ are obtained at the Sphinx observatory (3580 m a.s.l.).
3.2 CHM15k Ceilometer data and settings

The measurements from a ceilometer of type CHM15k-Nimbus (hereafter referred to as CHM15k) manufactured by Lufft have been used for this study. The CHM15k is a bistatic LIDAR with a Nd:YAG solid-state laser emitting linearly-polarized light at a wavelength of 1064 nm. It has a repetition rate ranging between 5 and 7 kHz, a maximum vertical resolution of 5 m, a maximum range of 15 km, a first overlap bin at 80 m and a full overlap reached at 800 m (specific for KSE, and PAY ceilometers, Hervo et al., 2016). The standard instrument output is a function of the received power per laser pulse, \( P \), backscattered by the atmosphere at range \( r \) and time \( t \). More precisely, the standard output of the CHM15k is the background, range and overlap-corrected, normalized signal, \( S \) defined as:

\[
S = \frac{(P(r,t) - B(t))r^2}{C_{CHM}(r)O_{CHM}(r)}
\]  

(3)

Where, \( B \) is the background, \( C_{CHM} \) is a normalization factor accounting for variations in the sensitivity of the receiver and \( O_{CHM} \) is the temporally-constant overlap function provided by the manufacturer.
The measurements used for this study have been collected during the period January-December 2014 at PAY and during September 2014 - November 2015 at KSE. At PAY, the ceilometer was setup vertically with a slight tilt (5° zenith angle) to avoid the specular reflection effect on cirrus ice crystals, as suggested by the manufacturer. At KSE, the ceilometer was mainly operated in tilted setup (71° zenith angle) in order to point towards the JFJ. At both sites, the overlap function $O_{\text{CHM}}$ has been corrected for temperature variations following (Hervo et al., 2016).

4.2.1 Special instrument settings for KSE

The CHM15k ceilometer at KSE was installed in August 2014 on the roof of the maintenance centre of the train station, at an altitude of 2069 m (Figure 2). From September to November 2014 and from March to November 2015, the ceilometer was tilted at 71° zenith angle with the laser beam passing close above (~20 m) the JFJ. From the beginning of November 2014 till the end of February 2015, the ceilometer was set back in the vertical position (5° zenith angle) to prevent the sun shining directly onto the ceilometer’s telescope.

The tilted setup of the ceilometer was chosen to observe the injections of CBL air at the level of JFJ and to probe the same air that is measured by the in-situ instruments at the JFJ. The tilted configuration has required an adjustment of the heater to redirect the heat flow inside the CHM15k case and prevent the overheating of the laser. Moreover, when measuring in slant path the maximum vertical height, $R_{\text{max}}$, depends on the tilt angle and on the instrument’s maximum range (15 km for the CHM15k), then, when measuring at 19°elevation angle

$$R_{\text{max}} = 2.069 + 15 \sin(19^\circ) = 6.64 \text{ km a.s.l.}$$

At this relatively low altitude, the standard background correction that subtracts from the received power $P$ its own mean value over the far range (furthest ranges along the profile, where the signal is assumed to be entirely represented by noise), cannot be performed as the far range may still contain aerosols or clouds. In order to overcome this problem a new technique of background removal that depends on the variance of $P$ has been developed and applied to each profile. The variance is calculated within spatial windows of 120 m to 1600 m (in steps of 120 m) and computed for all range bins between 390 m and 14970 m. The background corresponds to the mean value of $P$ over an optimal window. The optimal window’s position is the one minimizing the average of its variance values. The optimal window’s width is the one corresponding to the 75th percentile of the variance values at the optimal window’s position. Another advantage of measuring in slant path is that, compared to the vertical setup, the JFJ is reached by the ceilometer’s laser beam at 4.8 km that is already in the full overlap region.

5 PathfinderTURB validation at Payerne

Although, gradient-based algorithms are easy to implement for automatic operations, the layer attribution remains the main limitation for the retrievals. Normally, most of the uncertainty affecting a CBLH retrieval depends almost solely on the choice of the best gradient. We believe that for methods based on aerosol gradients the visual identification of the correct gradient by human expert still solves the attribution problem with the least uncertainty. Therefore PathfinderTURB is validated here against independent detections by human experts as well as against the bulk-Richardson method applied to co-located radiosonde profiles. The aim of the validation is to create an as-accurate-as-possible reference without selecting only golden-cases, but filtering out those cases when fog and precipitation prevent to define the CBL.
5.1 Human expert CBLH retrieval

A graphical user-interface has been developed for the human experts to detect the CBLH manually by clicking on the time-height cross section of \( S \). Auxiliary information are available from the interface about \( \nabla S \) (central differences \( (S(r_{i+1}, t_j) - S(r_{i-1}, t_j))/(2\Delta r) \)), the temporal variance (over 10 minutes), the sunshine duration, the vertical heat flux at the surface, the trend of hourly-averaged surface temperatures \( \Delta T \), the hourly stability index (as defined in Pal et al., 2013), the sunset/sunrise time, estimations of the CBLH based on the Parcel method (PM, Holzworth, 1964) and the bulk Richardson method (bR, Richardson, 1920) from continuous remote sensing instrumental data (Microwave Radiometer, Wind Profiler, Raman LIDAR), and twice daily radiosounding data (at noon and at midnight). The experts perform a manual detection of the entire daily cycle with the support of all the ancillary data and information.

Four experts from the remote-sensing division of MeteoSwiss have processed the overlap-corrected, normalised signal, \( S \), data from the CHM15k for Payerne for the year 2014. The guidelines and the criteria of the manual CBLH detection are provided in the Supplement (S5).

5.1.1 Analyzed dataset

We compared the detections by three experts (test group) against one expert that is called the reference. For the year 2014, the analysed days were the 5th, 10th, 15th, 20th, 25th and 30th of each month and the whole months of January, March, July and October; the test group analysed the 5th, 10th, 15th, 20th, 25th and 30th of each month. Once the missing data (due to instrument disruptions) and fog or precipitation days had been removed from the dataset, the total number of days analysed was 174. Covering an entire year, the database inspected by the test group and the reference is comprehensive in terms of diverse synoptic conditions, sunshine duration, cloudiness, season, etc. The ceilometer profiles were analysed by the experts of the test group separately and with no possibility to interact with each other. The detections made by the reference ("reference data") and those made by the test group were compared at each time step so that a matching procedure was established between a CBLH point in the reference data and the test group’s detections for the same time step. Only the CBLH points that match in time were retained for the comparison. If needed, the test group detections were linearly interpolated in order to match exactly the time vector of the reference. When comparing the reference data with all the test group detections the two datasets showed an excellent agreement, with a coefficient of determination equal to 0.96 (total of 5097 points over 140 days) and RMSE of 92 m. Nevertheless, some large difference (\( >500\text{m} \)) in the CBLH detections occurred in less than 3% of all cases. In general, discrepancies occurred when there was more than one layer that could be reasonably followed as CBLH, for example when an advected aerosol layer enters the profile and gets mixed inside the CBL or during the often ambiguous separation between the RL and the decaying CBL in the afternoon after the convective peak.

5.1.2 Determination of the expert consensus

The CBLH detections by each expert are linearly interpolated over 30 s and truncated before sunrise and after sunset. Before comparing the CBLH values retrieved by PathfinderTURB with those detected by the experts, a number of additional restrictions are defined. Manually-detected CBLH values for a given day must fulfil the following conditions in order to be “valid”:

\[
S(r_{i+1}, t_j) - S(r_{i-1}, t_j) = 2\Delta r \times \nabla S
\]

...
• No precipitation (station measurement) and no fog ( ceilometer measurement) for more than 2 consecutive hours.

• Only time periods of at least 30-minute duration, containing interpolated data from at least 2 experts and with spread (i.e. difference between the maximum and minimum values) of less than 10% of the mean CBLH plus 100 m are taken into account. The allowed spread increases with altitude because the probability to lie on different layers decreases with increasing altitude. The offset of 100 m added to the maximum allowed spread is an empirical value that translates into a permitted 300-m spread for a CBLH at 2000 m, which is a conservative estimate of the depth of the entrainment zone at that altitude (EZ depth can be 40% of the CBLH, Stull, 1988).

The CBLH detections respecting the above criteria are retained for comparison with PathfinderTURB. The mean value of all the valid detections (reference and test group) for each time step was calculated along with the lower and upper error bounds determined as the minimum and the maximum CBLH detections. In total, an expert consensus was reached for 135 days, out of the initial 174, covering a total of 43914 minutes. On 6 days, no ceilometer data were available. On 13 days, poor weather conditions prevented all detections. On 20 days, either the spread was too large or the duration of the matched temporal interval was too short.

Because the number of CBLH detections decreases as the days become shorter in winter, midday is the period of the day with the highest availability of detections (and the best agreement). Both morning and afternoon periods present difficulties, when detecting the CBLH from the timeseries of S, nevertheless the morning provides better availability of detection than the afternoon. The limitations of morning detections are related to the fact that during the first 2-3 hours after sunrise the CBL top is still low above the ground (< 200 m a.g.l.) and the overlap-corrected, normalized signal, S, is affected by the incomplete overlap in that zone, which makes the detection in the first few hundred meters difficult and more uncertain. Another source of complexity in the morning detections is that the residual nocturnal layer may remain aloft and very close to the developing CBLH, which increases the uncertainty related to the layer attribution. During the late afternoon, the CBL transforms into the RL as soon as buoyancy weakens and it becomes neutrally stratified. Under these conditions, the CBL top drops rapidly, but usually without displaying a clear aerosol gradient, which leads to a large (and somewhat unphysical) spread among the experts’ detections.

5.1.3 PathfinderTURB validation against the expert consensus

After applying the quality check (sect. 3.2.6) to the PathfinderTURB retrievals, the total number of the accepted retrievals covers 34720 minutes out of the 43914 minutes of manual data, i.e. 79% of the human expert consensus. In other words, PathfinderTURB removes about the 20% of its retrievals because of weak gradients at the level of the retrieved CBLH. The PathfinderTURB retrievals that passed the quality check are spread over the same number of days, (i.e., 135). Figure 3 shows, for Payerne and the year 2014, the density scatter plot of the CBLH values obtained by the human experts’ detections meeting the consensus and by PathfinderTURB (top panel), and the boxplot plus the histogram (bottom panel) of the differences between the two data sets. A coefficient of determination of 0.96, an RMSE of 76 m and an interquartile range of the differences of 96 m are obtained. The median and mean difference are 27 m and 41 m, respectively. The overestimation is largest during the second half of the afternoon (not explicitly shown here), when PathfinderTURB tends to follow the top of the
residual layer instead of the decaying CBL. Furthermore, the error is smaller than 500 m for 98.6% of the PathfinderTURB retrievals, and 92% of the retrievals have a relative error (w.r.t to the manually determined CBLH) smaller than 10%.

Figure 3 top panel shows the density scatter plot of PathfinderTURB vs. manualPBL. The bottom panel shows the boxplot and histogram of the difference between PathfinderTURB and manualPBL. The comparison shows that PathfinderTURB is robust (no unphysical jumps, cloud presence taken into account) and can address the attribution problem adequately. Although PathfinderTURB combines both gradient and...
variance methods to improve the correctness of the retrieval in different atmospheric conditions, the retrieval’s
uncertainty grows larger during the afternoon due to the convective decay before sunset, the weak turbulence and
the lack of well-marked aerosol gradients. During this period, temperature or vertical wind variability profiles
may provide more valuable information than ceilometer data.

5.2 Comparison with radiosonde-estimated CBLH

Next, the PathfinderTURB retrievals of the CBLH were compared to two methods based on the thermal structure
of the atmosphere: the PM and the $bR$. The PM defines the upper boundary of the CBL as the height to which an
air parcel with ambient surface temperature can rise adiabatically from the ground, neglecting other factors
(entrainment/detrainment, advection, subsidence, air humidity). It relies on profiles of potential temperature ($\Theta$)
and therefore requires vertical profiles and surface values of temperature ($T$) and pressure ($p$). In Payerne, $\Theta$
profiles are generated every 10 minutes by a microwave radiometer (MWR), and at noon and midnight also by
RS. The bulk Richardson number ($R_i_b$) is a dimensionless parameter that can be seen as the ratio between the
buoyancy and the wind-shear generated turbulence. The upper boundary of the CBL is determined as the first
height where $R_i_b$ exceeds the critical threshold of 0.33 (unstable conditions) or of 0.22 (stable conditions). The
required input values are the profiles of $\Theta$ and the wind. The stability conditions, essential for choosing the
correct threshold value, are derived from the sign of the slope of the linear fit of $\Theta$ in the first 30 m. At Payerne,
wind profiles are provided every 30 minutes by the wind profiler (WP), and at noon and midnight also by RS. As
the threshold value is larger than zero, the $bR$ method retrieves higher CBLH values than the PM. We refer to
Collaud Coen et al. (2014) for a more detailed description of the operational CBLH retrievals at Payerne using
the $bR$ method.

PathfinderTURB is compared to the RS-based $bR$ retrieval of the noon CBLH during the year 2014. In order to
increase the robustness of the $bR$ retrievals, the comparison is performed only when both $bR$ and PM retrievals
are available. Based on the calculations of Collaud Coen et al. (2014) the uncertainty of the retrieved CBLH
using both methods is on the order of ±50 to ±250 m for the midday peak of the CBLH. Within their uncertainty
intervals, the two methods can then be considered providing the same retrievals when the difference between
them is equal to or less than 250 m. For this reason only the retrievals being closer than 250 m and having an
uncertainty less than 250 m have been retained for the comparison. That has resulted in a total of 175 days being
considered. Out of the 365 days considered, the ceilometer data at noon were not available on 16 days. Of the
remaining 349 days, PathfinderTURB has been able to retrieve the CBLH on 202 days (58%). Out of the 202
days, 115 days (31.5%) had both $bR$ and PM retrievals satisfying the robustness conditions listed above.

The median and mean difference between RS- and PathfinderTURB-retrieved CBLH was 53 m and 41 m,
respectively, indicating a slight overestimation of the $bR$ method with respect to PathfinderTURB. From the
comparison we obtain a coefficient of determination of 0.85, a regression slope of 1.02 (Fig. 4), an RMSE of 162
m and an interquartile range of the difference of 174 m, hence larger than the spread observed in the comparison
with the human expert retrievals. The difference distribution has a Gaussian shape, with slight positive offset
values (Fig. 4). About the 98% of the data have an error smaller than 500 m, and the 82% have an error smaller
than the 10% (plus 100 m) of the CBLH retrieval by $bR$. In general, the correlation between PathfinderTURB
(ceilometer-based) and the $bR$ retrievals (RS-based) is not quite as good as the one between PathfinderTURB and the experts’ retrievals (both ceilometer-based). When comparing the two methods, one using radiosonde and the other the ceilometer data, it should be remembered that the two methods rely on different physical processes, i.e. thermal structure (RS) versus actual state of mixing of the aerosols (ceilometer). An example is the slight overestimation of the $bR$ between the end of the morning and the beginning of the afternoon, i.e. when buoyancy-produced turbulence reaches a maximum. This is because the $bR$ indicates the depth of the layer where conditions are favourable for vertical mixing, whereas the aerosol gradient depicts the actual state of mixing. By visual inspection of the entire (not only 12 UTC) dataset of comparisons PathfinderTURB vs $bR$ using the MWR data we conclude that the $bR$-based CBLH rises generally faster than the aerosol gradient in the morning. The decay of the $bR$-based CBLH occurs also generally faster than that of the aerosol gradient in the late afternoon, resulting in $bR$ retrievals lower than the CBLH retrievals based on aerosol gradient. This is explained by the fact that the aerosols remain suspended in the near-neutrally stratified air (transition from CBL to RL) and that no detectable aerosol gradient forms at the top of the decaying CBL. The gradient remains thus at about the same altitude as its midday maximum leading to a significant overestimation by PathfinderTURB. For this reason, LIDAR and ceilometers using aerosols as tracers are not best suited to detect the CBL decay, but rather the RL. Nevertheless, although at the local noon the $bR$ still diagnoses slightly higher CBLH, the time when the peak of convection occurs is a suitable time interval when to compare the $bR$ with PathfinderTURB.
Figure 4: the upper panel shows the scatter plot of RS(bR 12H) vs. PathfinderTURB. The bottom panel shows the boxplot and histogram of the difference between RS(bR 12H) and PathfinderTURB.

6 PathfinderTURB at the Kleine Scheidegg

6.1 Transport of BL air masses to the JFJ

Updrafts and downdrafts (initiated and sustained by solar radiation received at the surface) is the main vertical transport mechanism of the CBL air above the Swiss Plateau (Collaud Coen et al., 2011). Air lifted from a sunlit mountain slope is often warmer than air at the same height over an adjacent valley even if the latter was lifted from the valley floor. Hence, next to the development of up-slope (anabatic) winds, thermals generated at a mountain slope may rise higher than those generated at the valley floor. When both topography and meteorological conditions are favourable, up-slope winds may develop strong enough to break through the CBL’s capping inversion and inject CBL air into the FT immediately above the local CBL (LCBL) resulting in the formation of an aerosol layer (AL) above the CBL (Henne et al., 2004). This complex mountain circulation is characterized by dynamics occurring at different spatial scales (Figure 5). The AL or injection layer is a near-neutral, partly-mixed layer that is more diluted than the LCBL being the result of LCBL air mixed with FT air. The LCBL normally follows the topography (scale of a few kilometres), especially in the morning, and is often topped by a temperature inversion that marks the transition with the above AL. The AL does not follow individual valley or ridges, but follows the large-scale topography (few tens of kilometres) and again maybe capped by a temperature inversion that marks the transition to the FT (Henne et al., 2004, de Wekker, 2002). In his work, de Wekker (2002) concludes that in mountainous regions, the mixing layer height corresponds to the
top of the AL rather than the top of the LCBL and he renames it “mountain mixing layer”, because the AL depicts the height up to which particles can be transported by the various venting processes. The combination of the LCBL and the AL forms the CAL.

Figure 5. Schematic view of the daytime atmospheric structure and vertical pollution transport in and above the KSE site. The red line shows the CHM15k line of sight towards the Sphinx. The annotations denote the different thermal transport and mixing mechanisms of boundary layer air.

6.1.1 Annual cycle of the transport of CBL and CAL air masses to the JFJ

At the JFJ, aerosols and gases have been sampled continuously since many years. Different sources and transport regimes towards JFJ have been studied by many authors (e.g., Lugauer et al., 1998; Zellweger et al. 2003, Balzani Lööv et al. 2008, Henne et al. 2010, Collaud Coen et al., 2011; Collaud Coen et al., 2014; Herrmann et al. 2015) and revealed that JFJ resides mostly in the undisturbed (“clean”) lower FT, but that, especially in summer, it is influenced by thermally-induced lifting of CBL air and in general by additional vertical lifting processes such as frontal passages and Föhn flows (Zellweger et al., 2003, Ketterer et al., 2014). The frequently occurring (~35 % of the time) thermally-induced transport of CBL air towards the JFJ observed by Zellweger et al. (2003) during summer merits a closer look. In this sense, we want to provide a statistical description of the mechanisms leading the air to be thermally injected above the actual CBL and to reach the JFJ. Based on all these studies, it appears in fact that the direct contact of the undiluted CBL air with the in-situ instruments at JFJ occurs only rarely and is limited to the summer period (e.g., Ketterer et al., 2014).

In the study by Lugauer et al., (1998), the authors provide a 9-year climatological analysis of the vertical transport of aerosols to the JFJ and the corresponding synoptic conditions. The thermally-induced transport is nearly absent in winter or under cyclonic conditions and it is strongest in summer under anticyclonic periods. During favourable conditions, the aerosol concentration increases at the JFJ during the afternoon with a peak at
around 18:00 UT and the peak is stronger in northern synoptic wind than in southern because of the difference in upwind topography. Collaud Coen et al. (2011) found as well that the JFJ is mainly influenced by free tropospheric air masses in winter and largely influenced by the LCBL (also during the night) in summer during subsidence periods. In order to understand the impact of the thermally-driven dynamics on the in-situ measurements at JFJ and in the attempt to quantify, by direct observations, the number of times that the LCBL and the CAL reach the JFJ throughout the year, the data from the CHM15k have been analysed using PathfinderTURB during the period August 2014 till November 2015.

6.2 Adaptation of PathfinderTURB in tilted configuration

A slant-probing version of the PathfinderTURB algorithm that does not use the atmospheric variance to calculate the weights (eq. (2)), but solely to retrieve the first transition to the enhanced turbulence zone (liminf), runs operationally on the KSE ceilometer since September 2014. The reason why, contrary to what happens at PAY, PathfinderTURB does not use the signal variance when calculating the weights to determine the geodesic path at KSE, depends entirely on the tilted configuration of the ceilometer. Altitudes higher than 1000 m a.g.l. correspond already to ranges farther than 3000 m along the laser’s line of sight. At KSE, the backscatter signal at these ranges is often characterized by very little aerosol concentration, and is contaminated by significant contribution from the solar background. The backscatter signal is then dominated by the noise with low SNR and the calculated variance is not reliable. At closer ranges, where the first transition to turbulence is usually to be found, the signal has a much higher SNR and the variance can be measured reliably. At KSE, the LCBL height (LCBLH) retrieved by PathfinderTURB, corresponds to the first discontinuity in the vertical mixing of aerosols and can be estimated also during nighttime.

6.3 Retrieval of aerosol layers at KSE

Installed at KSE at 2069 m a.s.l., the ceilometer can exclusively detect the aerosol layers that form in the surrounding lower-altitude valleys (e.g., 1034 m a.s.l. Grindelwald, 566 m a.s.l. Interlaken) and that reach higher than the altitude of KSE. As an exception local sources of aerosols may form a local aerosol layer that becomes visible to the ceilometer and that is independent from the contribution of other aerosols from the adjacent valleys. These dynamics are valid during both day and night, but lead to different scenarios: both TCAL and LCBLH can be detected during daytime at KSE, the LCBLH only during periods when convection lift the LCBL air into the ceilometer’s field of view; during nighttime, when there is no convection, only the TCAL can be detected (if it is present). Likewise for any aerosol layers observed by the ceilometer above KSE, also the observed nocturnal TCAL can form from the residual layer of the surrounding valleys (at least partially). Because PathfinderTURB is based on the same retrieval principle during day and night, i.e. it looks for the first discontinuity in the well-mixed aerosol region, we will refer to the retrieved nighttime layer also as to mixing layer or LCBL even when the mixing is not due to convection, but rather to mechanical mixing from surface and katabatic winds.
6.3.1 LCBLH retrieval

The seasonal-averaged daily cycles of the retrieved LCBLH and TCAL during spring, summer and autumn are shown in Figure 6. During winter, (December 2014, January and February 2015) the algorithm has retrieved only a negligible number of LCBL heights and they have not been taken into account due to their lack of statistical significance. During spring (Fig.6a), summer (Fig.6b) and autumn (Fig.6c), the LCBLH grows through morning till reaching a peak in the afternoon. From the systematic visual inspection and comparison of LCBLH timeseries at PAY and KSE, we can say that the LCBLH peak occurs later at KSE than at the PAY. During the night, the LCBLH decreases, due to the concurrent effects of aerosol gravitational settling, subsidence and katabatic drainage flows, which result from radiative cooling of the surface triggering katabatic drainage flows. A likely explanation of the delay in the onset of the LCBL and of the afternoon peak at KSE is the following: due to the nighttime katabatic winds, FT air is driven down into the valley. These katabatic winds also continue for one or more hours, depending on the season, after sunrise (especially from the shaded mountain side) and work against the formation of the LCBL.

Generally, the autumn LCBLH shows a less pronounced daily cycle than in spring and summer, this is probably due to the fact that the vertical transport of aerosol-rich air is reduced by the stabilization within the lower troposphere during this period (Lugauer et al., 1998). In summer, the LCBLH has been retrieved by PathfinderTURB every day with only few exceptions; from May to August the LCBLH retrievals have reached in the JFJ during short periods on 25 individual days, but these occurrences lay above the 75 percentile of the LCBLH dataset and, hence, are not represented by the blue-shaded area in Fig.6. All occurrences of when the LCBLH and TCAL have reached the JFJ during the different months are listed in the Table 1 and are shown in Figure 7 for winter and summer. The LCBLH temporal evolution follows the classical shape of a growing convective boundary layer like over flat terrain, but the growth and the duration of the LCBL occur over a shorter period. This is consistent with the delayed onset of the LCBL due to the katabatic winds and the earlier weakening of convection under the contrasting action of the afternoon onset of the katabatic winds.

6.3.2 TCAL retrieval

During spring and autumn, the daytime TCAL evolution is correlated with the LCBLH especially in spring during the first hours after sunrise (convective growth) and until the afternoon peak. The nighttime evolution of the TCAL in spring and autumn also shows a correlation with the LCBLH although weaker. In summer, the TCAL does not show any significant correlation with the temporal evolution of the LCBLH neither during the day nor during the night.
(a) SPRING

(b) SUMMER
Figure 6 Season-averaged daily cycle of the TCAL (red dots) and of the LCBLH (blue dots) at KSE. The size of the dots corresponds to the number of measurements available in each temporal bin. Panel a) Spring (March to May). Panel b) Summer (June to August). Panel c) Autumn (September to November). Shaded areas show the 25%–75% inter-quartile range for LCBL (purple) and TCAL (red). The altitude of JFJ is indicated by the black-dashed horizontal line.

As it will become clear in the next section, the CAL has a major impact on the aerosol measurements at JFJ, especially in summer. On the other hand, there is also a relationship between aerosol parameters observed at JFJ and LCBLH during the entire day.

6.3.3 Occurrence frequency of LCBL and CAL reaching JFJ

In Table 1 we show, for each month during the studied period, the number of hours (cumulative 2-minute data points over the month), the number of days (number of days with at least one data point), and the percentage of time (time when JFJ was inside LCBL or CAL as a percentage of the total time when the retrievals existed). On the left-hand side of the table, we show the statistics corresponding to when the JFJ is reached by or embedded into the LCBL, on the right-hand side we show the statistics corresponding to when the JFJ is either into the LCBL or the CAL (LCBL + AL). The statistics reported in Table 1 are shown in the Circle charts in Figure 7, specifically for the winter (December-February) and summer (June-August) seasons. The statistics show that during winter (light grey rows in Table 1) the aerosol measurements at JFJ are never modified directly by the LCBL air, which remains constantly below the JFJ. Moreover, the total duration of time when the PathfinderTURB has detected a LCBL above KSE during winter accounts for no more than 65.52 hours. On the other hand, the CAL reaches the JFJ about one fourth of the time (21.23%) corresponding to a duration of 109.44 hours (distributed over 26 days). The remaining three quarters of time (78.77%), corresponding to a duration of 406.32 hours, the JFJ is situated in the FT, i.e. the in-situ measurements are characterized by background (molecular) conditions. Although it is impossible to establish the exact origin of the air in the AL (injection...
layer), we can speculate that winter AL is composed of aerosols originating from long-range transport and synoptic-scale lifting, rather than LCBL injections.

During summer (dark grey rows in Table 1) the situation changes significantly with the LCBL reaching the JFJ during the 3.63% of time, corresponding to 34.56 hours (distributed over 20 days).

Table 1: statistics of frequency of LCBL and CAL reaching or embedding the JFJ.

<table>
<thead>
<tr>
<th>Date</th>
<th>Hours</th>
<th>Days</th>
<th>%</th>
<th>Date</th>
<th>Hours</th>
<th>Days</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/2014</td>
<td>4.87</td>
<td>2</td>
<td>1.94</td>
<td>09/2014</td>
<td>149.03</td>
<td>18</td>
<td>23.882</td>
</tr>
<tr>
<td>10/2014</td>
<td>8.00</td>
<td>4</td>
<td>5.81</td>
<td>10/2014</td>
<td>88.70</td>
<td>18</td>
<td>16.29</td>
</tr>
<tr>
<td>12/2014</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>12/2014</td>
<td>43.30</td>
<td>9</td>
<td>26.67</td>
</tr>
<tr>
<td>01/2015</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>01/2015</td>
<td>33.53</td>
<td>10</td>
<td>21.76</td>
</tr>
<tr>
<td>02/2015</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>02/2015</td>
<td>32.70</td>
<td>7</td>
<td>16.41</td>
</tr>
<tr>
<td>03/2015</td>
<td>0.2</td>
<td>1</td>
<td>0.12</td>
<td>03/2015</td>
<td>45.77</td>
<td>13</td>
<td>10.24</td>
</tr>
<tr>
<td>04/2015</td>
<td>5.67</td>
<td>3</td>
<td>3.59</td>
<td>04/2015</td>
<td>80.73</td>
<td>15</td>
<td>14.82</td>
</tr>
<tr>
<td>05/2015</td>
<td>5.43</td>
<td>5</td>
<td>2.21</td>
<td>05/2015</td>
<td>114.07</td>
<td>17</td>
<td>22.83</td>
</tr>
<tr>
<td>06/2015</td>
<td>0.50</td>
<td>2</td>
<td>0.16</td>
<td>06/2015</td>
<td>174.60</td>
<td>24</td>
<td>29.30</td>
</tr>
<tr>
<td>07/2015</td>
<td>18.60</td>
<td>12</td>
<td>5.61</td>
<td>07/2015</td>
<td>380.47</td>
<td>28</td>
<td>56.49</td>
</tr>
<tr>
<td>08/2015</td>
<td>15.50</td>
<td>6</td>
<td>5.17</td>
<td>08/2015</td>
<td>217.63</td>
<td>19</td>
<td>36.34</td>
</tr>
<tr>
<td>09/2015</td>
<td>0.97</td>
<td>2</td>
<td>0.51</td>
<td>09/2015</td>
<td>56.30</td>
<td>12</td>
<td>11.50</td>
</tr>
<tr>
<td>10/2015</td>
<td>0.00</td>
<td>0</td>
<td>0.00</td>
<td>10/2015</td>
<td>19.87</td>
<td>6</td>
<td>5.53</td>
</tr>
<tr>
<td>11/2015</td>
<td>0.2</td>
<td>1</td>
<td>0.36</td>
<td>11/2015</td>
<td>4.10</td>
<td>2</td>
<td>1.42</td>
</tr>
</tbody>
</table>

Although the relatively low percentage may induce to think of a marginal effect, the striking parameter is that during summer the undiluted, polluted air of the LCBL is able to reach the JFJ (and potentially strongly affect the in-situ measurements of particles concentrations and of their optical properties) on 20 different days. With regard to the frequency and duration when the CAL has reached or embedded the JFJ, the statistics are even more remarkable with the 40.92% of the time or 772.8 hours distributed over 71 days. Also for the summer statistics, no quantitative conclusions can be drawn about the origin and type of the aerosols inside the AL. The aerosols could be locally emitted and injected into the AL or transported from regional to continental scale sources and could have been formed secondarily in the AL (Bianchi et al., 2016). In any case, the convective conditions occurring frequently during the summer hint at a significant mixing of the LCBL air into the FT forming the AL. Also the measurements by the in-situ instrumentations at JFJ show that the absorption
coefficient (indirectly proportional to the black carbon concentration) is largest during the summer period (Fig. 1).

Figure 7. Circle chart of frequency of LCBL and CAL reaching or embedding the JFJ during summer (JJA) and winter (DJF). Red-shaded areas show the percentage of time when the LCBL (CAL) is reaching or embedding the JFJ with respect to the total time when the retrievals exist. The white-shaded areas show the percentage of time when the JFJ sits inside the FT with respect to the total time when the retrievals exist.

As mentioned in the previous sections, these results are in agreement and confirm the indirect measurements and model simulations of previous works, especially those by Zellweger et al. (2003), Collaud Coen et al. (2011), Ketterer et al. (2014) and Herrmann et al. (2015). In fact for the first time the occurrence of the frequency of the convective (LCBL) and injection (AL) layers reaching the in-situ measurements at JFJ has been quantitatively calculated thanks to the automatic retrieval of the two layers directly at the JFJ and not, like done in the past, by extrapolated measurements of the boundary layer taken 5 km afar from the JFJ and based on stringent homogeneity conditions of the atmosphere.

Next, we study the impact of the LCBL and the CAL on the JFJ measurement site by putting them in relation with the in-situ measurement of the absorption coefficient.
6.4 Comparison with in-situ instrumentation

Figure shows the relation per season between the amplitude (max/min) of the diurnal cycles of the LCBLH retrieved by the PathfinderTURB and the absorption coefficient at 637 nm measured by the MAAP at JFJ. In the graph, the marker types account for the fact that the LCBL reaches (empty circles) the JFJ or not (filled circles).

Figure 8 Scatter plot of the amplitude (max-min) of the diurnal cycles of the absorption coefficient at 637 nm measured by the MAAP at the JFJ vs the LCBLH retrieved by PathfinderTURB for the period September 2014 to November 2015.

A positive relation between LCBLH and absorption coefficient is thought to exist especially for times when the LCBL does not reach the JFJ (filled circles). In these situations, when the LCBL grows deeper and approaches the height of the JFJ, the injections of LCBL air into the AL are more likely to occur and bring more aerosols to the in-situ instruments. In this case the amplitude of the daily cycle of the absorption coefficient is greatest, as the difference between the convective peak of the LCBL (large injections) and its minimum (negligible injections) correspond to high and low absorption coefficients, respectively. On the contrary, when the LCBLH is higher than the JFJ for a few hours during the day, the residual AL is also richer in particles and the amplitude of the daily cycle of the absorption coefficient is much smaller (the absorption remains large all day). During summer and for LCBLH < JFJ (filled circles) convection causes significant injections into the AL leading to large amplitude of the daily cycle of the absorption coefficient (larger slope in Fig.8). On the other hand, during spring probably due to the thick layer of snow accumulated during winter, the convection is too weak to create
significant injections into the AL, and relatively deep LCBL do not correspond to large daily cycles in the absorption (smaller slope). Despite the low coefficients of determination of the seasonal fits, when all points are fitted together the dependency of the max-min cycle of the absorption coefficient on the LCBLH becomes clearer.

![Figure 9 Scatter plot of the height of the TCAL vs the absorption coefficient at 637 nm measured by the MAAP at the JFJ during September 2014 to November 2015. Two independent linear regressions were fit to the data when TCAL was below (magenta line) and above (green line) the JFJ.](image)

The relation between the TCAL and the absorption coefficient at 637 nm, shown in Figure 9, has been studied in a slightly different way than for the LCBLH case. The TCAL represents the upper boundary of the AL or of the LCBL (when the AL is not present), when the TCAL is below the JFJ (TCAL < JFJ) the in-situ instrumentation on the JFJ is located inside the FT showing very little absorption and a negligible daily cycle of the absorption. When the TCAL reaches or embeds the JFJ (TCAL > JFJ) the absorption grows accordingly. Because the aerosols injected into the AL do not undergo a convective mixing, they tend to settle under the gravity force leading to higher aerosol concentration at the bottom than at the top of the AL. Also for this reason the absorption grows larger accordingly to a higher TCAL. Instead of the daily cycle of the absorption coefficient, Figure 9 shows, the hourly values of absorption versus the hourly TCAL values. The relation is shown in Figure 9 separately for TCAL < JFJ (filled circles) and TCAL > JFJ (empty circles). In each region, the data have been linearly fitted. As expected, for the region TCAL < JFJ the absorption coefficient does not depend (or only negligibly) on the TCAL. The absorption remains very low during all seasons except for summer when sporadic
injections manage to push aerosols higher above the detected TCAL. In the region TCAL > JFJ, as the CAL grows beyond the altitude of JFJ, the absorption coefficient shows a linear dependence on the depth of the CAL. Data points, when TCAL was within the central region that spans 330 m around the JFJ’s station height (3580 m), were omitted from the regression in order to more clearly separate between representative FT and AL conditions.

7 Conclusions
A novel algorithm, PathfinderTURB, has been developed, validated and applied to the real-time retrieval of the vertical structure of the planetary boundary layer. Its main advantages can be summarized as it follows.
PathfinderTURB provides reliable estimates of the daytime convective boundary layer height (CBLH) and of the Top of the Continuous Aerosol Layer (TCAL). The retrieval of the two layers is performed operationally and without need of ancillary data or any a priori information (except for climatological limits) from a model.
PathfinderTURB can also be adapted to different probing line’s angles and types of instrument. For this study, two settings have been especially tested and applied to the data of a CHM15k, the vertical-pointing and tilted-pointing, but PathfinderTURB can be easily adapted to any other angles and types of ceilometer. In perspective, based on the adaptability of the algorithm to diverse topographic conditions and on the fact that it does not require real-time ancillary data, PathfinderTURB is best suited to treat a large dataset from networks of ceilometers in real time.

PathfinderTURB suffers anyway some limitations related to the instrument. Due to the incomplete transmitt- receiver overlap in the first few hundred meters, and the unphysical gradients occurring in this zone, PathfinderTURB, which is partly gradient-based, is affected by a larger uncertainty in the first few hundred meters. Another limitation is that during the late afternoon, the aerosols remain suspended in the air (transition from CBL to RL) showing no detectable aerosol gradient at the top of the CBL. The only detectable gradient remains normally at the same altitude as the maximum CBL height reached during the central hours of the day, and corresponds more to the RL top rather than to the decaying CBL top. In fact, any method using aerosols as tracers (e.g. LIDAR) is not best suited to detect the afternoon CBL drop, but rather the RL.

The algorithm has been applied to one year of data measured by two CHM15k ceilometers operated at the Aerological Observatory of Payerne, on the Swiss Plateau, and at the Kleine Scheidegg, in the Swiss Alps. The algorithm has been first thoroughly evaluated at the Payerne station. At Payerne, the CBLH retrievals obtained by PathfinderTURB have been compared with the manual detections by human experts that acted as reference for the CBLH values and with the noon CBLH values retrieved by two methods based on the thermal structure of the atmosphere and using the radiosounding data: the parcel method and the bulk Richardson method. The comparison against the human experts reference revealed a median difference of 27 m and a RMSE of 76 m. The median difference with respect to the radiosounding reference is 53 m with a RMSE of 162 m.

Based on the excellent agreement with the two reference methods, the PathfinderTURB has been applied to the complex-terrain site at the Kleine Scheidegg for the period September 2014-November 2015. There, the local CBL (LCBL) is retrieved based on the data of the CHM15k whose axis has been tilted by a zenith angle of 71° in order to probe the air volume next to the Sphinx Observatory (3580 m, a.s.l.) on the Jungfraujoch.
The results presented in Section 6 showed that the PathfinderTURB can be adapted to slant-path probing, thus providing real time and continuous LCBLH and TCAL data along the line of sight of the CHM15k. This has allowed to separate the contribution of these two layers and to understand their impact on the absorption coefficient measured in-situ at JFJ.

The season-averaged daily cycle shows that the CAL reaches or includes the JFJ for the 40.92% of the total time in summer and for the 21.23% of the total time in winter for a total of 97 days during the two seasons. The statistics suggest that the CAL modifies the physical and chemical properties of the air sampled at JFJ, especially during summer when the absorption coefficient at 637 nm at JFJ shows a distinct dependence on the CAL depth. The season-averaged daily cycles show that the LCBL reaches or includes the JFJ for short periods (3.94% of the total time) on 20 days in summer and never during winter. The statistics suggest that also the LCBL modifies the physical and chemical properties of the air sampled at JFJ, but exclusively during summer, as these measurements refer purely to the direct contact of undiluted LCBL air with the JFJ. During summer the amplitude of the daily cycle of the absorption coefficient at 637 nm at JFJ shows a distinct dependence on the LCBL depth.

As a general conclusion, we can state that the results obtained at the KSE site are in agreement and confirm the indirect measurements and model simulations of previous works, especially those by Zellweger et al. (2003), Collaud Coen et al. (2011), Ketterer et al. (2014) and Herrmann et al. (2015). In fact for the first time the impact of the convective (LCBL) and injection (AL) layers on the in-situ measurements at JFJ has been quantitatively calculated for a complete year thanks to the automatic retrieval of the two layers directly at the JFJ.

Acknowledgements
This study has been financially supported by the SNF through ICOS-CH. The authors would further like to thank Nicolas Bukowiecki for giving access to JFJ aerosol measurement data. The authors are grateful to Kornelia Pönitz and Holger Wille (Lufft) for technical information about the CHM15k.

References


Szintai, B.: Improving the turbulence coupling between high resolution numerical weather prediction models and Lagrangian particle dispersion models, Swiss Federal Institute of Technology, Lausanne, November., 2010.


