# Adaptive Estimation of the Stable-Boundary-Layer Height Using Combined Lidar and Microwave Radiometer Observations

Umar Saeed, Francesc Rocadenbosch, *Member, IEEE*

and Susanne Crewell, *Member, IEEE*

**Abstract**

A synergetic approach for the estimation of stable boundary-layer height (SBLH) using lidar and microwave radiometer (MWR) data is presented. Vertical variance of the backscatter signal from a ceilometer is used as an indicator of the aerosol stratification in the nocturnal stable boundary-layer. This hypothesis is supported by a statistical analysis over one month of observations. Thermodynamic information from the MWR-derived potential temperature is incorporated as coarse estimate of the SBLH. Data from the two instruments is adaptively assimilated by using an extended Kalman filter (EKF). A first test of the algorithm is performed by applying it to collocated Vaisala CT25K ceilometer and Humidity-and-Temperature Profiler (HATPRO) MWR data collected during the HD(CP)$^2$ Observational Prototype Experiment (HOPE) campaign at Jülich, Germany. The application of the algorithm to different atmospheric scenarios reveals the superior performance of the EKF compared to a non-linear least-squares estimator especially in non-idealized conditions.

U. Saeed is with the Department of Signal Theory and Communications (TSC), Remote Sensing Lab. (RSLab), Universitat Politècnica de Catalunya (UPC), Campus Nord, E-08034 Barcelona, Spain e-mail: (umar.saeed@tsc.upc.edu).

F. Rocadenbosch is with the Department of Signal Theory and Communications (TSC), Remote Sensing Lab. (RSLab), Universitat Politècnica de Catalunya (UPC) / Institut d'Estudis Espacials de Catalunya (IEEC/CRAE), Campus Nord, E-08034 Barcelona, Spain.

S. Crewell is with the Institute for Geophysics and Meteorology (IGMK), University of Cologne, 50969 Köln, Germany.

**Index Terms**

Laser radar, microwave radiometry, adaptive Kalman filtering, remote sensing, signal processing

## I. INTRODUCTION

Continuous estimates of the Atmospheric Boundary Layer Height (ABLH) are needed for several applications ranging from weather, avionics and air-quality and dispersion models. The development of Atmospheric Boundary Layer (ABL) over a diurnal cycle is a local as well as regional phenomenon and shows a typical behaviour under clear-sky and low synoptic conditions. During daytime, the ground-surface absorbs solar radiation, and as a result, near-surface air warms up and initiates convective motions leading to well mixed conditions in the Convective Boundary Layer (CBL). The CBL, also called Mixing Layer (ML), reaches its maximum height in the afternoon. At the top of CBL, the Entrainment Zone (EZ) acts as a buffer between the CBL and the Free Troposphere (FT) above. When the sun sets, turbulence decreases and radiative cooling causes the development of a Stable Boundary Layer (SBL) close to the surface. The remnants of the CBL with the associated well mixed pollutants form the Residual Layer (RL) on top of the SBL which is characterized by weak intermittent turbulence [1], [2], [3].

The structure of the Nocturnal Boundary Layer (NBL) mainly depends upon three underlaying physical processes namely, turbulent mixing, radiative cooling and heat exchange with the soil [4], [5]. The type of the NBL which develops at a particular location and time depends upon the relative strength of these processes and, therefore, there can be three types of the NBL: fully turbulent (also known as the night-time ML), intermittently turbulent, and non-turbulent (also known as the SBL). Fully turbulent NBL occurs when the wind shear becomes the dominant force, whereas in the case of the non-turbulent NBL or SBL [6], [7], radiation and heat exchange with the soil become dominant and the turbulence is almost non-existent resulting in horizontal stratification of the aerosols in the atmosphere through a process known as *fanning* [1]. The intermittently turbulent NBL occurs when there are alternating cycles of turbulence and non-turbulence.

From the modelling perspective, it is a big challenge to capture these three types of NBL. Whereas models do capture the turbulence and related mixing processes quite well, the modelling of stable boundary conditions is still poor [8]. Nevertheless, the accurate modelling of the SBL is highly important for correctly predicting night-time temperatures and the dispersion of pollutants and, therefore, it is of much interest and relevance in the ABL research community to study the SBL and its correct estimation [9].

Measurements of ABLH make use of the typical vertical structures of temperature, humidity, wind and aerosol. While aerosol distribution as probed by ground-based lidars is well suited for the estimation of the Convective Boundary Layer Height (CBLH), the estimation of the Stable Boundary Layer Height (SBLH) is more complex. Many of the techniques presented for the lidar data are a variant of the gradient detection method [10], [11] and, therefore, suffer from multiple-aerosol layers attribution problems. This means that under the conditions of multiple aerosol layers in the boundary-layer these methods are limited in terms of providing a consistent solution due to their non-adaptive nature and lack of physical basis for layer attribution. Some more advanced techniques such as the "Peaks" and "Wavelets" [12], which utilize advanced signal processing, and the Bayesian Selective Method (BSM) [13], which combines data from lidar, a physical boundary-layer model and a climatological data-set in a statistically optimal way, do improve the estimation performance. However, they are still limited by the starting height of full overlap of lidar and the unavailability of lidar data under cloudy and rainy conditions. Furthermore, cooling leads to a higher relative humidity and since the size of the aerosols varies with the moister content [10] due to hygroscopic effects, the higher backscatter coefficient is linked to a higher amount of humidity in the atmosphere.

Most of the previous work for SBLH estimation from temperature data are based on the measurements from Radiosonde (RS) [6], [7] mainly exploiting temperature profiles. Though Microwave Radiometer (MWR) can provide continuous time series of temperature profiles [14], few studies have exploited these data for SBLH estimation. Recently, an approach based on the gradient of the retrieved-potential temperature from a MWR been presented by [15]. One of the main limitation of MWRs is their low

vertical-resolution, thereby the large uncertainty associated with the estimates. Moreover, the temperature profile is over-smoothed especially at higher altitudes, hence missing important features within the retrieved profiles. As a result, MWRs are unable to give accurate indication of the SBLH and the residual boundary-layer height.

Departing from these previous efforts to estimate the SBLH, in this paper we present a combined lidar-MWR approach using an Extended Kalman Filter (EKF). The approach is based on the hypothesis, whereby stable aerosol layer corresponds to a minimum variance region (MVR) in the variance profile of the lidar backscatter signal. The hypothesis is based on the fact that during the stable conditions prevailing in the night-time, with minimal to nil convection, and in the absence of mechanical turbulence, which usually results in horizontal stratification of the aerosol layers [1], [13], [11], [16], the backscatter signal remains almost constant across the layer's vertical span. This minimum variance-behavior is opposite to the maximum variance behavior defining the EZ and CBL [17], [18].

As discussed earlier, MVRs are not unique due to the presence of multiple layers of aerosols. Moreover, MVRs also get corrupted by the instrument noise (added to the backscatter signal) essentially distorting its shape. In order to overcome the limitations associated with the use of MVRs from the lidar backscatter data alone, the proposed approach is based on the synergy between the lidar and the MWR, whereby the MWR plays a role of layer attributor. Nevertheless, the boundary containing the temperature inversion information should be sufficient to correctly segregate the most relevant MVR within the lidar backscatter variance profile. An estimation algorithm based on an EKF is then applied on the selected MVR to calculate the SBLH with low uncertainty. Unlike the Bayesian Selective Method mentioned earlier, the use of EKF provides statistically optimal estimates by minimizing the mean-square-error over time without the need of averaging the observation data.

This paper is organized as follows: Sect. II contains a brief summary of all the instruments and dataset used in this work. Sect. III formulates the SBLH estimation problem and related estimation algorithm. Sect. IV discusses the prototypic test cases and the retrieved SBLH estimation results. Finally,

the conclusions are presented in Sect. V.

## II. INSTRUMENTS AND DATA-SET

The instruments used in this work include a ceilometer, a MWR, and a RS. In the following, a summary of each instrument is given. In order to successfully cross-examine data and/or to assimilate it synergistically, collocated measurements taken in April/May 2013 at the Jülich ObservatorY for Cloud Evolution (JOYCE) [19] in Jülich, Germany are used for case studies and statistical analysis.

### A. Instruments

*Ceilometer* - A ceilometer is a single-wavelength optical instrument which measures the cloud-base height. It works on the principle of LIght Detection And Ranging (LIDAR), essentially measuring the backscattered light, after the emission of a laser pulse. The time of flight of the backscattered pulse is used to determine the distance to the target/height distribution of the atmospheric scatterers. Under relatively clear atmospheric conditions (typically, optical thickness, $\tau < 1$), the range-corrected intensity of the backscattered lidar signal is essentially proportional to the aerosol/molecular concentration of the atmospheric mixture. Though ceilometry refers to "cloud height and extent" and former ceilometer instruments were initially designed for cloud-base height detection only, today, modern ceilometers can detect the ABLH, and cloud base height. In case of semi-transparent clouds multiple layers can be observed. The ceilometer used in this work is the Vaisala CT25K [20], [19]. The transmission wavelength of this instrument is 905 nm with a pulse repetition frequency (PRF) of 5.6 kHz. Under clear-sky conditions the typical sounding range of the instrument is roughly from 60 to 7,500 m with a range resolution of 30 m and time resolution of 15 seconds. The receiver field-of-view (FOV) is 0.66 mrad. Since this ceilometer is a mono-axial system (laser and receiving telescope optical axis coincide) its minimum sounding height of approximately 60 m is caused by the near-range saturation effect. The instrument software provides profiles of the attenuated backscatter coefficient $\beta^{att}$ as an output.

*MWR: Humidity and temperature profiler (HATPRO)* - The Radiometer Physics GmbH (RPG) HATPRO [21] is a MWR, capable of profiling temperature and humidity with high temporal and limited spatial resolution. The receiver of HATPRO has two filter banks, each with 7 channels, in the $20 - 30$ GHz, and $50 - 60$ GHz bands, respectively. It works in two scanning modes, zenith-pointing mode for full tropospheric profiling (range up to 10 km, vertical discretization of $150 - 250$ m), and boundary-layer scanning mode (6 elevation angles, range upto $1000$ m, vertical discretization 50 m). The time resolution of the measurements is about $2.70$ minutes. The limited vertical-resolution of the MWR-retrieved quantities (e.g., physical temperature) is inherently due to having less degrees of freedom than the available measurement channels. Thus, for temperature retrieved profiles only about four pieces of independent information are available [22]. Therefore, the true vertical resolution on the inverted products ("clean-data" spatial resolution) is much lower than the vertical discretization of the retrieved temperature profile. Compared with radiosondes Löhnert and Maier (2012) show random differences between MWR and RS down to $0.5$ K in the lower boundary-layer increasing to $1.7$ K at $4$ km height.

*RadioSonde* - A RS is an *in situ* instrument which is capable of measuring temperature, pressure, and relative humidity (RH) in the atmosphere up to about $30$ km by vertical sounding. It is launched through a large balloon inflated with hydrogen or helium gas. During its flight up to $30$ km height, it can drift more than $200$ km away from above the point of its launch though the horizontal displacement within the ABL is not significant for practical purposes. Nevertheless, the vertical profiles of atmospheric parameters measured by the RS are still considered a *de facto* reference or physical truth for the remote sensing purposes. During the measurement period 226 soundings of Graw DFM-09 have been performed and used among others for a water vapor inter-comparison study [23].

Fig. 1 illustrates the development of the ABL on 24.04.2013 via the observed potential temperature and backscatter time-height structure over the full diurnal course. Because this day was characterized by weak synoptic forcing and nearly no clouds, ABL development is close to the idealized cases described in the introduction. Since the SBL prevailed until about an hour past sunrise, the aerosol mixing process

did not start until about 06:00 UTC when the convection became significant and the convective ML developed. This can well be seen by the neutral conditions, i.e. vertically constant potential temperature. Around sunset (18:30 UTC) mixing recedes and thus aerosol stratification occurs. The extent of aerosol stratification in the SBL is directly linked to the amount of thermal stability in the boundary-layer. Case studies from this day are shown in Sect. IV.

CT25k is an old first-generation ceilometer by Vaisala with low pulse energy (1.6 $\mu$J/pulse) and low pulse repetition rate (4369 Hz) as compared to modern second-generation systems such as CL31 with slightly lower pulse energy (1.2 $\mu$J/pulse) but with almost double pulse repetition rate (8192 Hz) [16]. This explains a poorer SNR from the CT25k side, which causes that the structure of the boundary layer at higher altitudes especially the top of the mixing layer during the day-time is not clearly identifiable (e.g., blurred noisy pattern between 5-18 UTC and 0.5-1.5 km in height).

## III. ADAPTIVE SBLH DETECTION METHOD

In Sect. III-A, the formulation of an inverted Gaussian-like model profile representative of an idealized MVR for stable aerosol layer in the night-time is presented. The model profile is used to adaptively fit the height-dependent variance profile of the measured lidar data. In Sect. III-B, the EKF is chosen as the adaptive estimator because it minimizes the mean-square-error over time and thus it assimilates the temporal information of the signal optimally. For comparison, a simplified non-linear least-squares (NLSQ) formulation is also presented. The methodology to obtain coarse SBLH estimates from the MWR is described in Sect. III-C. Sect. III-D presents the synergetic MWR-lidar approach to combine data from the two instruments.

### A. SBL problem formulation based on lidar data

During the night time, especially when there is minimum to nil convection and the turbulence due to mechanical wind shear is negligible, a SBL develops near the ground surface. As a result, in the absence of any external forces, aerosols in the atmosphere gets stratified in a layered fashion. This layering of

aerosols can result in a single or multiple layers depending on the location and type of the atmospheric aerosols.

Each layer of aerosols is seen in the backscattered lidar signal with time as a strong, and more or less constant, backscatter signature within the layer boundaries. As a result, regions with a relatively constant backscatter level in the height-dependent backscatter profile correspond to MVRs in the backscatter variance profile. MVRs mark a sharp decrease in the variance of the backscatter signal as compared to the variance levels below and above a layer.

For estimation purposes, a MVR is modelled by an inverted Gaussian-like function as shown in Fig. 2 [24]. The bulk of the inverted Gaussian bell lies in the height interval $\left[z_1', z_2'\right]$ whereas its ending tails lie in the range intervals $\left[z_1, z_1'\right]$, and $\left[z_2', z_2\right]$, characterized by an approximately constant high variance level. The inverted Gaussian-like profile shown corresponds to an idealized aerosol layer which is represented as a uniform backscatter signal across its vertical extent. The center of the idealized bell represents the height of the SBL and its spread or standard deviation roughly corresponds to the width of the aerosol layer. The constant variance levels of the Gaussian model correspond to the background variance outside of the aerosol layer. Mathematically, the backscatter model variance is formulated as

$$h(\mathbf{x}) = Be^{-\frac{1}{2}\left[b(\mathbf{z}-z_{SBL})\right]^2} + d, \tag{1}$$

where $\mathbf{x} = [z_{SBL}, b, B, d]^T$ is the state vector, $\mathbf{z}$ is the height vector, $z_{SBL}$ is the SBLH, $b = \frac{1}{\sigma}$ ($\sigma$ being the standard deviation of the Gaussian distribution) is the width parameter, $B$ is the variance amplitude, and $d$ is the background variance level. These four parameters will be estimated either adaptively by using an EKF or non-adaptively by using a NLSQ (see Sect. III-B). As customary, bold font is used to represent vectors.

*Lidar data pre-processing* - The range-dependent background-subtracted received lidar signal is given by

$$Q(z) = P(z) + n(z), \tag{2}$$

where $P(z)$ is the ideal (i.e., noiseless) lidar power return and $n(z)$ is the observation noise. The lidar power return is given by the single-scattering elastic lidar equation [25],

$$P(z) = \frac{C}{z^2}\beta(z)T^2(z), \tag{3}$$

where $\beta(z)$ is the range-dependent volume backscatter coefficient of the atmosphere ($m^{-1}sr^{-1}$), $T^2(z) = e^{-2\int_0^z \alpha(z')dz'}$ is the two-way path atmospheric transmittance, and $C$ is the system constant ($Wm^3$). The quantity $\beta(z)T^2(z)$ is known as the attenuated backscatter coefficient,

$$\beta^{att}(z) = \frac{z^2}{C}P(z). \tag{4}$$

Under moderate-to-clear atmospheres and lidar sounding paths roughly below 3 $km$, the transmittance term can be assumed close to unity and, therefore, $\beta^{att}(z) \cong \beta(z)$ [16], [26]. In the following, the term $\beta(z)$ will be used to refer to the ceilometer attenuated backscatter or simply "the backscatter".

The noisy attenuated backscatter coefficient, $\beta'(z)$, in response to a measured (i.e., noisy) lidar signal, $Q(z)$, can be obtained from Eq. 4 by substituting $Q(z)$ (Eq. 2) in place of $P(z)$. It follows that

$$\beta'(z) = \beta(z) + v(z), \tag{5}$$

where $\beta'(z)$ is the noisy attenuated backscatter coefficient, $\beta(z)$ is the noiseless attenuated backscatter coefficient, and $v(z) = \frac{z^2}{C}n(z)$ is the range-corrected noise scaled by the ceilometer system constant, $C$.

Fig. 3a shows the basic signal-processing block diagram to estimate the backscatter variance profile, $\hat{V}_\beta(z)$, from the noisy attenuated backscatter coefficient profile, $\beta'(z)$. The first step in the processing of $\beta'(z)$ is to denoise it by low-pass filtering (LPF). This gets rid of the high-frequency content of $\beta'(z)$, which is associated to instrumental noise, while retaining the low-frequency content. The later is associated to the noiseless atmospheric backscatter, $\beta(z)$, and related low-frequency atmospheric fluctuations of interest for this study. Formally,

$$\hat{\beta}(z) = \beta'_{LP}(z), \tag{6}$$

where a hat "^" indicates "estimate of" and subscript "LP" indicates low-pass filtering.

The LPF is implemented by using a moving-average filter with a rectangular window length, $w$. The appropriate window-length is obtained by monitoring kurtosis (K) of the residual high-frequency noise,

$$\hat{v}(z) = \beta'(z) - \beta'_{LP}(z), \tag{7}$$

as signal-processing indicator. Thus, the window-length yielding a kurtosis figure closest to $3$ (K=3 for a pure Gaussian random process) [27] is chosen as the filtering window-length.

Fig. 3b1 shows plots of the measured noisy backscatter, $\beta'(z)$, along with the estimated backscatter profile, $\hat{\beta}(z)$, for different window lengths. The estimated backscatter profile using the largest window length (300 m) becomes over-smoothed and, as a result, misses detailed atmospheric features. Likewise, the residual high-frequency noise, $\hat{v}(z)$ (Fig. 3b2) is far from the typical shape of a Gaussian process as evidenced by an asymmetric distribution of positive/negative noise spikes (K=2.2). On the contrary, the shortest window length tested (60 m) significantly leaks noise into $\hat{\beta}(z)$ and also yields K=2.2. A window length of $150$ m gives the best results for $\hat{\beta}(z)$ as it filters most of the high frequency instrumental noise along with preserving the atmospheric features of the signal and yielding a fairly symmetric noise distribution for the residual noise with K=3.1 (i.e., approximately Gaussian).

The next processing step to estimate the height-dependent variance profile in Fig. 3a is associated to the atmospheric backscatter, $\hat{V}_\beta(z)$, given the estimated backscatter profile, $\hat{\beta}(z)$. At this point we use that

$$\hat{V}_\beta(z) \cong V_{\hat{\beta}}(z), \tag{8}$$

where $V_{\hat{\beta}}(z)$ represents the vertical variance of $\hat{\beta}(z)$ (see block diagram in Fig. 3a). For simplicity, the window length associated to this moving-variance calculation centered at height $z$ has been kept the same as that of the denoising LPF. Fig. 3b3 shows the variance profile calculated using the selected window length of $150$ m. Example variance profiles (Fig. 4a) reveal the occurrence of up to two to three MVRs within one profile. During the course of two hours in the evening of 24.04.2013 MVRs between 300-600 m can be detected.

In order to further investigate and validate the existence the of MVRs and to prove their relation with

the nocturnal SBL, a data-set of $28$ days from HD(CP)$^2$ Observational Prototype Experiment (HOPE) campaign starting from 15.04.2013 until 14.05.2013 is used. Cloudy and rainy conditions are filtered out with a resolution of 30 min. Vertical variance of the filtered backscatter data in the first km in height is calculated and averaged over bins of 30-min in time and 200-m in height. MVRs within each time-height bin are determined by the local minima in the averaged vertical variance profiles. Considering up to two MVRs in one profile, the diurnal cycle of frequency of occurrence (Fig. 4b) reveals frequent MVR occurrence in the altitude range between 200 and 800 m during night-time which fades out towards daytime. The similarity of the MVR diurnal cycle with the typical behaviour of the SBL supports our hypothesis to use MVRs as proxy for SBL.

*B. The extended Kalman filter approach*

The use of EKF for the ABLH builds on the previous works of [28], [29] for the estimation of atmospheric optical parameters from the backscatter lidar signal. Later, [26], [30] have used an EKF based on an erf-like model to describe the ML-to-EZ transition for the estimation of the Mixing Layer Height (MLH). From this background, the four characteristics parameters of model Eq. 1 are assembled into the state vector,

$$\mathbf{x}_k = [z_{SBL,k}, b_k, B_k, d_k]^T, \tag{9}$$

which is to be estimated at each successive discrete time, $t_k$.

An EKF is essentially based on two models, the *measurement model*, and the *state-vector model*:

*Measurement Model* - The measurement model relates the atmospheric state vector, $\mathbf{x}_k$, to the measurement vector, $\mathbf{z}_k$,

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k, \tag{10}$$

where $h$ is the SBL measurement function given by Eq. 1, and $\mathbf{v}_k$ is the observation noise which consists of measurement noise as well as modelling errors. In Eq. 10 above, $\mathbf{z}_k$ refers to the observables formed

from the estimated backscatter variance (Eq. 8) at each time $t_k$, that is,

$$\mathbf{z}_k = \hat{V}_\beta(\mathbf{z})\Big|_{t=t_k} = \left[\hat{V}_\beta(z_1), \hat{V}_\beta(z_2), \ldots, \hat{V}_\beta(z_N)\right]_k. \tag{11}$$

At this point it is worth noticing that while $z$ stands for the height variable, $\mathbf{z}_k$ and $\hat{\mathbf{z}}_k$ are the actual and the estimated observation vector (also called *measurement* vector).

In the extended Kalman filter (EKF), the non-linear model function $h$ is linearized by calculating its Jacobian (or observation matrix $\mathbf{H}_k$) with respect to the state vector. The filter output at time instant $k$ can then be written as

$$\hat{\mathbf{z}}_k = \mathbf{H}_k\hat{\mathbf{x}}_k^-, \tag{12}$$

where

$$\mathbf{H}_k(z; \mathbf{x}) = \left[\frac{\delta h(z)}{\delta z_{SBL}} \; \frac{\delta h(z)}{\delta b} \; \frac{\delta h(z)}{\delta B} \; \frac{\delta h(z)}{\delta d}\right]\Bigg|_{\mathbf{x}=\hat{\mathbf{x}}_k^-},$$

$$= \left[\mathbf{H}_k^1 \; \mathbf{H}_k^2 \; \mathbf{H}_k^3 \; \mathbf{H}_k^4\right]_{N\times 4}, \tag{13}$$

and

$$\mathbf{H}_k^1(z_{SBL}, b) = \frac{\delta h(z)}{\delta z_{SBL}} \tag{14}$$

$$= Bb^2 (z - z_{SBL}) e^{-\frac{1}{2}\left[b(z-z_{SBL})\right]^2}, \; z \in [z_1', z_2'], \tag{15}$$

$$\mathbf{H}_k^2(z_{SBL}, b) = \frac{\delta h(z)}{\delta b} \tag{16}$$

$$= -Bb (z - z_{SBL})^2 e^{-\frac{1}{2}\left[b(z-z_{SBL})\right]^2}, \; z \in [z_1', z_2'], \tag{17}$$

and

$$\mathbf{H}_k^3(B, d) = \frac{\delta h(z)}{\delta B} \tag{18}$$

$$= e^{-\frac{1}{2}\left[b(z-z_{SBL})\right]^2}, \; z \in [z_1, z_1') \cup (z_2', z_2], \tag{19}$$

$$\mathbf{H}_k^4(B, d) = \frac{\delta h(z)}{\delta d} = 1, \; z \in [z_1, z_1') \cup (z_2', z_2]. \tag{20}$$

In Eq. 12, above, $\hat{\mathbf{x}}_k^-$ is the state-vector estimate prior to assimilation of the measurement at time $t_k$ and $\hat{\mathbf{z}}_k$ is the "projected" measurement estimate from the filter.

Range intervals $[z_1', z_2']$ and $[z_1, z_1') \cup (z_2', z_2]$ respectively define the measurement-model "fitting" ranges inside and outside the MVR (Fig. 2). In order to assimilate the thermodynamic information of the SBL, outer boundaries $z_1$ and $z_2$ are assessed synergetically from MWR estimates of SBLH (see Sect. III-C).

*State-vector model* - The state-vector model essentially describes the temporal projection of the state-vector at each successive time $t_k$ through the recursive equation,

$$\mathbf{x}_{k+1} = \mathbf{\Phi}_k \mathbf{x}_k + \mathbf{w}_k, \tag{21}$$

where $\mathbf{\Phi}$ is the transition state matrix ($4 \times 4$) and $\mathbf{w}_k$ is the state-noise vector with covariance matrix $\mathbf{Q}_k = E\left[\mathbf{w}_k \mathbf{w}_k^T\right]$. A simple Gauss-Markov model, with $\mathbf{\Phi} = \mathbf{I}$ ($\mathbf{I}$ is the identity matrix), can be set as the transition matrix.

The Kalman filtering recursive loop requires three inputs related to the state-vector model: (i) an initial guess of the state-vector, $\hat{\mathbf{x}}_0^- = [z_{SBL,0}, b_0, B_0, d_0]^T$, (ii) an estimate of the initial *a priori* covariance matrix, $\mathbf{P}_0^- = E\left[\mathbf{e}_0^- \mathbf{e}_0^{-T}\right]$ where $\mathbf{e}^- = \mathbf{x}_0 - \hat{\mathbf{x}}_0^-$ is the *a priori* error, and (iii) an estimate of the state-noise covariance matrix, $\mathbf{Q}_k = [w_k w_k^T]$.

The initial *a priori* error covariance matrix, $\mathbf{P}_0^-$, and state-noise covariance matrix, $\mathbf{Q}_k$, can be set as static diagonal covariance matrices of the form, $\mathbf{P}_0^- = diag\left[\sigma_{e,z_{SBL}}^2, \sigma_{e,b}^2, \sigma_{e,B}^2, \sigma_{e,d}^2\right]$ and $\mathbf{Q}_k = Q = diag\left[\sigma_{z_{SBL}}^2, \sigma_b^2, \sigma_B^2, \sigma_d^2\right]$ where $\sigma_X$ and $\sigma_{e,X}$, $X = (z_{SBL}, b, B, d)$, are the standard deviations of the state-vector parameters and of the *a priori* error on the initial guess, respectively. Subindex "e" stands for "error". These formulations for matrices $\mathbf{P}_0^-$ and $\mathbf{Q}_k$ are simple enough and conveniently model all the case examples (SBL or nocturnal) analyzed in Sect. IV. Moreover, $\sigma_X$ and $\sigma_{e,X}$ are represented as proportional to the initial guess $\hat{\mathbf{x}}_0^-$ via factors $\mu_Q$ and $\mu_P$, respectively. Therefore,

$$\left(\sigma_{e,z_{SBL}}, \sigma_{e,b}, \sigma_{e,B}, \sigma_{e,d}\right) = \mu_P\left(z_{SBL,0}, b_0, B_0, d_0\right),$$

$$\left(\sigma_{z_{SBL}}, \sigma_b, \sigma_B, \sigma_d\right) = \mu_Q\left(z_{SBL,0}, b_0, B_0, d_0\right) \tag{22}$$

or simply $\mathbf{Q}_0 = \text{diag}\left[\left(\mu_Q \hat{\mathbf{x}}_0^-\right)^2\right]$ and $\mathbf{P}_0^- = \left[\text{diag}\left(\mu_P \hat{\mathbf{x}}_0^-\right)^2\right]$. The scaling factors $\mu_Q$ and $\mu_P$ are user-defined parameters. For example, setting $\mu_Q$ and $\mu_P$ as $0.1$ means the margin of uncertainty associated to $\hat{\mathbf{x}}_0^-$ is $10\%$ and the change in $\hat{\mathbf{x}}_k$ from time instant $k$ to $k+1$ is expected to be within $10\%$ of the values of $\hat{\mathbf{x}}_k$ at time $k$.

*Observation-noise modelling* - The observation noise, $v_k$, is modelled by the noise covariance matrix, $\mathbf{R}_k = E\left[\mathbf{v}_k \mathbf{v}_k^T\right]$, where $E[.]$ is the expectancy or ensemble operator and $\mathbf{v}_k$ is the N-component vector associated to heights, $z_i, i = 1, \ldots, N$. This matrix, $\mathbf{R}_k$, informs the filter about the quality of the observables, $\mathbf{z}_k$, which is a crucial input to the filter in order to assimilate the information conveyed by each new observable at each time step $t_k$.

One difficulty in the estimation of noise covariance matrix is the need of an ensemble of measurements. In off-line processing applications this difficulty can easily be circumvented by accessing at each time $t_k$ a time window comprising past and future measurements, $\mathbf{z}_k$, from the recorded data with stationary statistics. In an on-line processor this implies an estimation delay equal to the time length of the "future" measurements accessed. In this work we resort to offline processing and stationary statistics. The hypothesis of stationary statistics is a realistic one in stable boundary-layer estimation since in the absence of any mixing process the aerosols tend to remain still aloft for longer periods and hence, the backscatter signal remains approximately constant with time. Therefore, temporal variations on the vertical profiles of $V_{\hat{\beta}}(z)$ on shorter time-scales (e.g., 2 min) provide an estimate of the noise covariance matrix. Formally,

$$\hat{\mathbf{R}}_k = diag\left[\sigma_n^2(z_1), \ldots, \sigma_n^2(z_N)\right]_{t_k} \tag{23}$$

with

$$\sigma_n^2(z_i)\Big|_{t_k} = V\left[\mathbf{z}_{i,k}\right]\Big|_{I_k}, \quad I_k = [t_{k-M}, \ldots, t_k, \ldots, t_{k+M}], \tag{24}$$

where $\mathbf{z}_{i,k}$ is the i-th component of the measurement vector, $\mathbf{z}_k$, at time $t_k$, which according to Eq. 8 represents the estimated backscatter variance at height $z_i$, that is, $z_{i,k} = \hat{V}_\beta(z_i)$. In Eq. 24 above, $I_k$ is the time interval defining the ensemble time window. For example, assuming ceilometer data with 15-s

temporal resolution, a time ensemble of $8$ records translates into a 2-min stationary time window (as is the case used here).

*Nonlinear Least Squares Approach (NLSQ)* - Alternative to the EKF, a non-adaptive NLSQ solution is also considered [31]. In the non-adaptive approach, the atmospheric state-vector, $\mathbf{x}_k$, is estimated at each time $t_k$ by using only the present-time measurement, $\mathbf{z}_k$, therefore, disregarding past information. For each lidar measurement, $\mathbf{z}_k$, at time, $t_k$, the NLSQ solution for the model parameters is found by minimizing the quadratic norm of the error function between the observation vector, $\mathbf{z}_k$, and the model output, $\hat{\mathbf{z}}_k = h\left(\hat{\mathbf{x}}_k^{LSQ}\right)$,

$$\epsilon\left(\mathbf{x_k^{LSQ}}\right) = \mathbf{z}_k - h\left(\hat{\mathbf{x}}_k^{LSQ}\right), \tag{25}$$

with respect to the state vector, $\mathbf{x}_k$. That is,

$$\min\left\{\left\|\epsilon\left(\mathbf{x}_k^{LSQ}\right)\right\|^2\right\}\Bigg|_{\mathbf{x}_k=\left[z_{SBL,k}, b_k, B_k, d_k\right]}. \tag{26}$$

### C. Coarse SBLH estimation from MWR data

In order to assimilate the thermodynamic information about SBL into the EKF formulation (Sect. III-B), potential temperature retrieved from the brightness temperature measured by the MWR is used for coarse SBLH estimates [24], [32]. Under stable conditions, the potential temperature increases until at a height where neutral conditions are met with a constant potential temperature. The transition from the SBL to the RL is typically rather smooth and thus a clear boundary between the two regions is challenging to define. Here, it is pertinent to mention that the top of the SBL is slightly higher than the surface-based temperature inversion which becomes evident in a potential temperature profile as the region where its slope is almost zero. The SBLH is therefore defined as the height where the temperature lapse-rate is adiabatic signifying neutral condition ($\frac{\partial \bar{\theta}}{\partial z} = 0$, see inset in Fig. 5a) [1], [33], [15].

In order to estimate SBLH and its uncertainty range from potential temperature profiles from MWR five idealized potential-temperature profiles [1] are fitted to them. The idealized profiles are based on two key parameters $\bar{\theta}_0$ and $\bar{\theta}_s$, which are the RL and near-surface potential temperature, respectively.

TABLE I: Five idealized physical profiles for the SBL. Potential temperature is in the units of Kelvins (K) and height in meters (m).

| Profile | Equation | Input Param. | Primary solving param. | Auxiliary solving param. |
|---|---|---|---|---|
| Stable mixed | $\theta(z) = \begin{cases} \theta_s, & \text{for } z \le h \\ \theta_0, & \text{for } z > h \end{cases}$ | $\theta_s$ | $h, \theta_0$ | - |
| Linear mixed | $\theta(z) = \begin{cases} \left(1 - \frac{z}{h}\right)\theta_s + \frac{z}{h}\theta_h, & \text{for } z \le h \\ \theta_0, & \text{for } z > h \end{cases}$ | $\theta_s$ | $h, \theta_0$ | $\theta_h$ |
| Linear | $\theta(z) = \begin{cases} \theta_s + (\theta_0 - \theta_s)\frac{z}{h}, & \text{for } z \le h \\ \theta_0, & \text{for } z > h \end{cases}$ | $\theta_s$ | $h, \theta_0$ | - |
| Polynomial | $\theta(z) = \begin{cases} \theta_0 - \left(1 - \frac{z}{h}\right)^{\alpha}(\theta_0 - \theta_s), & \text{for } z \le h \\ \theta_0, & \text{for } z > h \end{cases}$ | $\theta_s$ | $h, \theta_0$ | $\alpha$ |
| Exponential | $\theta(z) = \theta_0 - (\theta_0 - \theta_s)\, e^{-\frac{z}{H_{\Delta\theta}}}$ | $\theta_s$ | $h, \theta_0$ | $H_{\Delta\theta}$ |

The idealized profiles incorporate the SBLH, noted here as $h$, and $\theta_0$ as the main parameters along with auxiliary parameters. Tab. I summarizes a reformulation of these five idealized profiles in terms of the measured potential temperature and the parameters to be estimated.

From these profiles the problem of SBLH estimation from MWR data can be stated as a LSQ problem as follows: Given (i) the profile of the raw potential temperature (i.e., the retrieved potential temperature from MWR measurements), in what follows, $\theta_{MWR}(z)$ and (ii) the surface potential temperature $\theta_s$ from MWR, one wishes to estimate (a) the model SBLH, $h$, and (b) the model RL potential temperature, $\theta_0$, using the objective function

$$\min \left\{ \|\theta_{MWR}(z) - \theta(z, \theta_s, \vec{\mathbf{x}}_k)\|^2 \right\} \Big|_{\vec{\mathbf{x}} = [h, \theta_0, aux]}. \tag{27}$$

In Eq. 27, $\vec{\mathbf{x}}$ is the objective vector which is to be solved, and "aux" denotes the auxiliary solving parameters of Tab. I such as temperature jump at the top of the SBL, $\theta_h$, for the linear-mixed model, the model order, $\alpha$ (typical value of $2 - 3$), for the polynomial profile, and *integral depth scale* within the

SBL, $H_{\Delta\theta} = \frac{\int_0^h [\theta_0 - \theta(z)] d\theta}{\theta_0 - \theta_s}$, for the exponential profile.

*MWR data processing* - The MWR-retrieved potential temperature profiles, $\theta_{MWR,ret}(z)$, are non-uniformly sampled along the height dimension due to the coarse vertical resolution of the instrument, which decreases with height. As a first processing step, $\theta_{MWR,ret}(z)$, is interpolated to obtain a uniformly sampled profile, $\theta_{MWR}(z)$. It is pertinent to mention here that this interpolation step does not alter the inherent height-varying resolution of the potential temperature data retrieved from the MWR measurements. For interpolation purposes a cubic spline interpolation is used. The second processing step is fitting the five SBL profiles of Tab. I to the uniformly-sampled potential temperature, $\theta_{MWR}(z)$, profile by using the NLSQ approach discussed in Sect. III-B. The best fitting profile is selected based on a minimum mean-square-error (MSE) criterion.

Fig. 5a shows the MWR-retrieved potential temperature profile (blue crosses), $\theta_{MWR}(z)$, from the HOPE campaign data, 24.04.2013 at 21:00 UTC. The fittings from the five idealized profiles, of Tab. I, are also plotted. The polynomial and the exponential profiles provide best fits with a minimum root-mean-square-error of $0.15$ K and $0.29$ K, respectively. Note, that even the best-fit idealized profiles significantly deviate from the retrieved potential temperature at the heights of RL as the real atmosphere hardly ever behaves in an idealized way. Furthermore, the coarse vertical resolution and limited accuracy of the MWR-retrieved temperature profile contribute to the uncertainty $\Delta z_{MWR}$ especially at higher altitudes where the information content degrades.

The uncertainty associated to the MWR-derived SBLH estimate results from (i) the measurement uncertainty of the brightness-temperature and (ii) the ill-posed retrieval problem. Therefore typical instrument uncertainty is included in the development of the retrieval algorithm that is used to convert brightness temperatures into temperature profiles. Here a statistical retrieval algorithm trained on a long-term data set of representative atmospheric profiles following the procedure outlined by [14] and [22] is employed. By testing the performance of the algorithm on a synthetic test data set the altitude dependent uncertainty $\epsilon(z)$ can be derived which varies between $0.44$ K on the ground and $1.20$ K at 2 km. The increase in $\epsilon(z)$

TABLE II: Discretization of vertical grid.

| height [m] | discretization step [m] |
|---|---|
| $0 \leq z \leq 250$ | 50 |
| $250 < z \leq 700$ | 75 |
| $700 < z \leq 1000$ | 100 |

with height is a consequence of the low degrees of freedom in the measurement data [22]) and reflects that the true vertical resolution of the retrieved potential temperature profiles is coarser than the discretization grid selected -a subjective choice from the user's side.

To estimate the SBLH uncertainty, $\Delta z_{MWR,meas}$, associated to the retrieved potential-temperature profile, $\theta_{MWR}(z)$ (Fig. 5a), we apply the SBLH-model-fitting procedure described above, not only on $\theta_{MWR}(z)$ but also to the upper and lower error-bound profile for the potential temperature $\theta_{MWR}(z) \pm \epsilon(z)$, in order to estimate the impact on Fig. 5b).

The discretization grid of Tab. II, which has been used to retrieve the temperature data for this work, follows this resolution-decreasing behaviour with height and it is, therefore, proportional to the real resolution of the data albeit the true exact values are not known. Here, for simplicity, $\Delta z_{MWR,res}$, has roughly been approximated by the discretization step of Tab. II at the estimated SBLH.

After the perturbational procedure and taking into account the uncertainty due to the low vertical resolution, the upper and lower bounds of the estimated SBLH can be written as

$$\hat{h}_{MWR}^{u,l} = \hat{h}_{MWR}^{fit} \pm \Delta z_{MWR}, \tag{28}$$

where superscript "u" and "l" stand for "upper" and "lower" error bounds, respectively, and $|\Delta z_{MWR}| = |\Delta z_{MWR,res}| + |\Delta z_{MWR,meas}|$ is the estimated error including the height discretization uncertainty.

Fig. 5c shows four potential temperature profiles with a time-spacing of 30-min for the time interval 21:00-22:30 UTC from real measurements. For each profile the error-bar is based on the $\Delta z_{MWR,meas}$ meaning the actual error-bars will be even bigger when $\Delta z_{MWR,res}$ is taken into account for each profile. The error-bars thus define the EKF fitting ranges in the form of coarse SBLH estimates.

TABLE III: Statistics relating the thermodynamic SBL determined from potential temperatue retrieved from MWR and MVRs.

| - | total count | percentage |
|---|---|---|
| Total 30-min bins | 488 | 100 |
| Only one MVR inside MWR error-bar | 264 | 54.1 |
| Two MVRs inside MWR error-bar | 26 | 5.3 |
| No MVR inside MWR error-bar | 62 | 12.7 |
| MVRs do not exist | 136 | 27.9 |

### D. Lidar-MWR synergy for SBLH estimation

In order to study how the SBLH height range given by the MWR-retrieved potential temperature is related to MVRs within the variance profile of the lidar backcatter data a statistical assessment was performed (Tab. III). Over the study period (15.04.2013-14.05.2013) $488$ 30-min averaged variance profiles with corresponding MWR observations exist during night-time, i.e. between 20:00 and 06:00. By counting the number of MVRs within the uncertainty range defined by the MWR error-bar, $\Delta z_{MWR}(z)$, it is evidenced that in almost $54\%$ of cases a single MVR falls within the MWR error bar. For $26\%$ of cases two MVRs lie within the MWR error-bar. No MVR was present inside the MWR error-bar for about $12\%$ and MVRs did not exist for about $28\%$ of cases.

When a single MVR falls within the MWR error-bar, which is the case for $54\%$ of cases, it becomes straightforward to determine a fine estimate of the SBLH without the need to perform further processing. However, the situation gets complicated when more than one MVRs lie within the MWR error-bar or no MVR lies in the MWR error-bar or MVRs do not exist at all. In such cases, previous time estimates of SBLH can be utilized to come-up with an SBLH estimate for the present time. The Kalman filter provides a convenient framework for such kind of estimation problems where already small fluctuations (the MVRs) must adaptively be estimated and time track under random environments (i.e., the vertical profile of the variance evolving with time as a random process) and where information from previous

time records must be assimilated. In the following a formulation, to determine the EKF fitting ranges, for the synergy between the MWR coarse estimate of the SBLH and the width of MVR from the previous time instant is presented.

At each time instant, $t_k$, estimation range boundaries $z_{1,k}$, $z'_{1,k}$, $z'_{2,k}$ and $z_{2,k}$ (Sect. III-B and Fig. 2) are determined by using a combination of the MWR SBLH estimates as well as the shape of the minimum-variance region from the previous time step, $t_{k-1}$. The following adaptive search boundaries are proposed

$$
\begin{cases}
z_{1,k} = \hat{h}^l_{MWR,k} \\
z'_{1,k} = \hat{z}_{SBL,k-1} - \Delta h^l_k \\
z'_{2,k} = \hat{z}_{SBL,k-1} + \Delta h^u_k \\
z_{2,k} = \hat{h}^u_{MWR,k}.
\end{cases}
\tag{29}
$$

In Eq. 29 the MWR coarse SBLH search interval (refer to Fig. 2) is defined as

$$
I_{MWR,k} = \left[ \hat{h}^l_{MWR,k}, \hat{h}^u_{MWR,k} \right],
\tag{30}
$$

and $z_{SBL}$ is the *fine* SBLH estimated from the EKF/NLSQ, and hence, the solution of the estimation problem.

Besides, the time-resolution of the MWR is lower than that of the ceilometer (a factor of 10 in this case), which means that $I_{MWR,k}$ changes every 10 ceilometer time records. Therefore, to match the time-stamps and to obtain a one-to-one correspondance, $I_{MWR,k}$ are interpolated according to the ceilometer time-stamps. In Eq. 29 the MVR-search range, $\left[ z'_{1,k}, z'_{2,k} \right]$, is estimated from the SBLH estimate at previous-step plus/minus an incremental height $\Delta h^u_k / \Delta h^l_k$, respectively, which define the upper ("u") and lower ("l") MVR search bounds. In order to comply with the constitutive relation, $z'_1 > z_1$ and $z_2 > z'_2$ depicted in Fig. 2, the upper and lower search bounds $\left[ z'_{1,k}, z'_{2,k} \right]$ are computed at the $1\sigma$ width of the MVR (parameter $b$ in Eq. 1) and constrained by the maximum allowable upper/lower heights $b^{l/u}_{max}$ (consequence of the

geometry of Fig. 2). Formally,

$$\Delta h_k^l = \max \left( b_{k-1}, b_{max}^l \right),$$

$$\Delta h_k^u = \max \left( b_{k-1}, b_{max}^u \right), \qquad (31)$$

where

$$b_{max}^l = \hat{z}_{SBL,k-1} - h_{MWR}^l,$$

$$b_{max}^u = h_{MWR}^u - \hat{z}_{SBL,k-1}. \qquad (32)$$

In Eq. 32 above $b_{max}^{l/u}$ define the maximum distance (i.e., height in Fig. 2) from the estimated SBLH, $\hat{z}_{SBL}$, to the lower search bound, $h_{MWR}^l$, and the maximum distance from the upper search bound, $h_{MWR}^u$, to the estimated SBLH, $\hat{z}_{SBL}$. A pictorial representation of Eq. 31 and Eq. 32 variables is given in Fig. 2.

The recursive scheme of Eq. 29 thus allows to conveniently merge thermodynamic information about the SBL from the MWR-derived potential temperature with information about aerosol stratification provided by the lidar. Thus, whereas the MWR plays two important roles: 1) providing the correct aerosol layer attribution, and 2) defining coarse estimation search ranges ($z_{1,k}$ and $z_{2,k}$), the lidar provides a highly-resolved estimate of the SBLH corresponding to the center of the aerosol layer identified by the thermodynamic stability information from the MWR.

## IV. RESULTS AND DISCUSSIONS

In this work, data collected during the HOPE campaign at Jülich conducted during April 02-July 24, 2013 is used. Measurements from CT25K ceilometer and Humidity And Temperature PROfiler (HATPRO) MWR (Sect. II) are used to estimate the SBLH under different atmospheric scenarios. First, the algorithm is applied to night-time data from 24.04.2013 (Fig. 1), i.e. a clear-sky day with classical boundary layer. Second, the performance is evaluated for a two hour case study from 29.04.2013 with a weakly stable nocturnal boundary layer and low aerosol amount. SBLH estimates from the EKF and NLSQ estimates are compared to assess their different performance.

*24.04.2013* - At midnight the temperature profile (Fig. 1) reveals stable conditions. Together with the high aerosol backscatter signal in the SBL that reaches heights up to 600 m (Fig. 6a) this indicates a well developed stable boundary layer between mid-night and sunrise (around 4:30 UTC). Both EKF and NLSQ estimate SBLH between 400 and 600 m with good agreement among each other though EKF SBLH estimates are much smoother due the assimilation of the information from past estimates.

The benefit of EKF becomes more pronounced when the period past sunset (around 19:30) is considered from 21:00-24:00 UTC. As the surface of the Earth becomes gradually cooler, convection ceases and atmospheric stability increases affecting successively higher altitudes. Initially aerosol is not accumulated in the SBL and the backscatter return from the RL is still significant when compared to SBL. This is in contrast to the previous period between midnight and early morning when aerosol had accumulated in the SLBH during the course of the night and the backscatter signal in the RL was lower. Thus, Fig. 6b shows that although there is a higher aerosol backscatter signal from the lower heights, a considerable amount of aerosols are still trapped in the RL. Fig. 6b also compares the SBLH estimates obtained from the synergetic lidar-MWR approach of Sect. III-D from EKF and NLSQ estimators. Comparison between EKF and NLSQ SBLH estimates shows that the NLSQ fails most of the time. This is due to the non-adaptive behaviour of the NLSQ estimator which causes that when it is confronted with different MVRs within the "coarse search" boundaries marked by the MWR, the NLSQ can not disambiguate which one to choose. In these situations, the NLSQ just provides the least-squares-error (LSQ) solution averaged over all the MVRs in the estimation range. In contrast, the EKF conveniently provides a reasonable solution averaged under a criterion of minimum MSE over time due to its assimilation of past temporal information (covariance). Furthermore, good agreement with the radiosonde at 23:00 is evident.

*29.04.2013* - The period between 22:00-24:00 UTC provides a more challenging case from the signal-processing point of view in which the variance (an already small quantity as mentioned in Sect. III-A) must be estimated from the backscatter returns (the signal component) in response to an atmospheric scene nearly depleted of aerosols (weakly stable nocturnal boundary layer). Thus, Fig. 6c shows that aerosols

in the lower height, where usually a stable boundary layer is developed by this time of the day, provide much less backscatter signal as compared to higher heights belonging to the residual layer (such kind of situations usually prevail when thermal emission from the Earth surface is slow or there is a cloud cover trapping the surface heat in the atmosphere or a cleaner airmass is advected due to synoptic conditions at the location of measurements). In the case of Fig. 6c the spatial variance of the backscatter signal (estimated via the approximation of (Eq. 8) becomes a weak "tracking" indicator of the SBLH because the variance is in fact estimated from signal samples approaching zero (i.e., the backscatter returns in a SBL virtually depleted of aerosols) and the existence -and correct detection- of MVRs is inherently linked to the stratification of aerosols. Therefore, this is a "complex" estimation problem characterized with very low SNRs, where the role of the MWR becomes even more critical and where SBLH estimates largely benefit from "a priori" information coming from potential temperature data. Obviously, NLSQ estimator (non adaptive) fails almost at all the times. However, the EKF still provides reasonably good SBLH estimates thanks to the assimilation of MWR and past temporal information.

Main limitations of the technique presented in this work encompass both instrumental and environmental limitations: (i) Concerning *instrumental limitations*, the partial overlap of the lidar (also known as laser-telescope cross-over function), which is the case for bi-axial lidar systems, distorts the attenuated backscatter profile at low heights. Therefore, the technique only works for SBL heights which are above the range of full overlap of the system. Since the CT25k is a mono-axial system with its first range gate starting at around 60 m, this instrumental limitation does not arise in this work. (ii) Concerning *environmental conditions*, it must be said that typically, the SBL height is not more than one km [1], which means that the technique presented here will, almost always, be of application along the first km of the vertical lidar profile. Yet, the existence of stable atmospheric conditions is always a pre-requisite since the stratification of aerosols occurs only under SBL. In fact, aerosol load will ultimately condition the quality of filter convergence since aerosols are needed as tracers of the atmospheric phenomenon under study. Though this is still a matter of research, comparatively, the EKF has successfully estimated

the SBLH in the study case of Fig. 6c (low aerosol load) with a contrast backscatter level in the SBL as low as 0.3-0.6 a.u. (heights interval 274-574 m) as compared to a free-troposphere level of almost 0 a.u. at 2 km. In the study case of Fig. 6b (high aerosol load, nocturnal case) the backscatter level lies between 0.7-1.0 a.u. In terms of thermodynamic stability detected by the MWR, elevated inversions are a challenge for detection.

## V. CONCLUSIONS

An adaptive solution based on synergetic use of data from a lidar-ceilometer and a MWR has been presented. Vertical variance of the attenuated backscatter signal from a (Vaisala CT25K) ceilometer has been used as an indicator of the aerosol stratification in the nocturnal stable boundary layer. Minimum variance regions within the vertical variance profiles have been modelled by an inverted Gaussian-like function and model parameters including the sought-after SBLH have adaptively been estimated by using an EKF. Coarse SBLH estimates from MWR-retrieved potential temperature observations have been assimilated for aerosol layer disambiguation and to incorporate information about the thermodynamic stability of atmosphere.

Physical/signal-processing steps can be summarized as follows: First, the vertical variance of the ceilometer backscatter signal is estimated by using a moving-average filter ($150$ m rectangular window) as de-noising step. The correct filter window length is obtained by monitoring statistical properties of residual instrumental noise. By this means, minimum variance regions (MVRs), which are indicative of the stratification of aerosol layers in the nocturnal stable boundary-layer, are evidenced in the vertical moving variance profiles of the denoised backscatter signal. The existence of MVRs and their correlation with the SBL has been further investigated by processing one month data from HOPE campaign. After screening for clouds and rain, the vertical variance calculated and averaged for 30-min in time and 200-m in height bins. Counting the MVRs falling inside the error-bar defined by MWR, it was observed that for about $54\%$ of cases a single MVR and for about $5\%$ of cases two MVRs were present inside the MWR error-bar. Whereas, no MVRs lied inside MWR error-bar for $13\%$ of cases, the MVRs did not exist at all

for about $28\%$ of cases.

MVRs are modelled by using an inverted Gaussian-like function with the SBLH as key parameter of the state-vector model. The state vector is adaptively estimated by using an EKF whose search boundaries are defined from the MWR "coarse" SBLH estimates and the $1\sigma$ width of the time-adaptive Gaussian model.

The synergetic approach has been applied to measurements from the HOPE campaign at Jülich, Germany. Three atmospheric scenarios have been presented: 1) An early morning scenario of SBL with deep stratification of aerosols and thermodynamic stability in the atmosphere; 2) An evening case where the SBL just starts developing from the Earth surface; 3) And a complex case with shallow stratification of aerosols in the atmosphere. Results from these three cases have shown that the proposed synergetic approach performs well for the different time intervals of the day as well as under different "nocturnal" atmospheric conditions.

Future work of this prototype algorithm is to involve long-term measurement data as well as more complex atmospheric scenarios whereby the nocturnal boundary layer is intermittently turbulent. More sensitive ceilometer instruments such as the CL31 and CL51 can provide better information about the aerosols stratification and could reveal fine structures in the SBL. Moreover, better processing of the MWR data for SBLH estimates with lower uncertainty could provide better results for complex cases. The ultimate goal of this preliminary study is to develop a synergetic retrieval algorithm for full diurnal cycle of the ABLH over the course of the day.

## VI. ACKNOWLEDGMENTS

R̲E̲F̲E̲R̲E̲N̲C̲E̲S̲

[1] R. B. Stull, *An Introduction to Boundary Layer Meteorology*, ch. 12, Stable Boundary Layer, pp. 499–543. Springer Netherlands, 1988.

[2] B. Fisher and D. Thomson, "Harmonization in the preprocessing of meteorological data for atmospheric dispersion models," Final Report COST710, COST Action 710, 1998.

[3] S. Liu and X. Liang, "Observed diurnal cycle climatology of planetary boundary layer height," *J. Climate*, vol. 23, pp. 5790–5809, 2010.

[4] G. J. Steeneveld, B. J. H. Van de Wiel, and A. A. M. Holtslag, "Modelling the evolution of the nocturnal boundary layer for three different nights in cases-99," in *In Proceedings of 16th Symposium on Boundary Layers and Turbulence*, August 2004.

[5] Z. Qiang and W. Sheng, "A study of the atmospheric boundary layer structure during a clear day in the arid region of Northwest China," *Acta Meteor. Sinica*, vol. 23, no. 3, pp. 327–337, 2009.

[6] Y.-K. Hyun, K.-E. Kim, and K.-J. Ha, "A comparison of methods to estimate the height of stable boundary layer over a temperate grassland," *Agricultural and Forest Meteorology*, vol. 132, pp. 132–142, Mar. 2005.

[7] W. M. Angevine, C. W. Hare, J. E.and Fairall, D. E. Wolfe, R. J. Hill, W. A. Brewer, and A. B. White, "Structure and formation of the highly stable marine boundary layer over the Gulf of Maine," *Journal Of Geophysical Research*, vol. 111, no. D23S22, pp. 1–13, 2006.

[8] D. Seidel, Y. Zhang, A. Beljaars, J.-C. Golaz, A. Jacobson, and B. Medeiros, "Climatology of the planetary boundary layer over the continental United States and Europe," *Journal Of Geophysical Research*, vol. 117, no. D17106, pp. 1–15, 2012.

[9] J. Cuxart, A. Holtslag, R. Beare, E. Bazile, A. Beljaars, A. Cheng, L. Conangla, M. Ek, F. Freedman, R. Hamdi, A. Kerstein, H. Kitagawa, G. Lenderink, D. Lewellen, J. Mailhot, T. Mauritsen, V. Perov, G. Schayes, G.-J. Steeneveld, G. Svensson, P. Taylor, W. Weng, S. Wunsch, and K.-M. Xu, "Single-column model intercomparison for a stably stratified atmospheric boundary layer," *Boundary-Layer Meteorology*, vol. 118, no. 2, pp. 273–303, 2006.

[10] N. Eresmaa, A. Karppinen, S. M. Joffre, J. Räsänen, and H. Talvitie, "Mixing height determination by ceilometer," *Atmos. Chem. Phys*, vol. 6, pp. 1485–1493, 2006.

[11] M. Haeffelin, F. Angelini, Y. Morille, G. Martucci, S. Frey, G. Gobbi, S. Lolli, C. ODowd, L. Sauvage, I. Xueref-Rémy, B. Wastine, and D. Feist, "Evaluation of mixing-height retrievals from automatic profiling lidars and ceilometers in view of future integrated networks in Europe," *Boundary-Layer Meteorology*, vol. 143, pp. 49–75, 2012.

[12] M. d. Haij, W. Wauben, and H. Baltink, "Continuous mixing layer height determination using the LD-40 ceilometer: a feasibility study," technical report, Royal Netherlands Meteorological Institute (KNMI), 2007.

[13] F. Giuseppe, A. Riccio, L. Caporaso, G. Bonafé, G. Gobbi, and F. Angelini, "Automatic detection of atmospheric boundary layer height using ceilometer backscatter data assisted by a boundary layer model," *Quarterly Journal of the Royal Meteorological Society*, vol. 138, pp. 649–663, 2012.

[14] S. Crewell and U. Löhnert, "Accuracy of boundary layer temperature profiles retrieved with multifrequency multiangle microwave radiometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 2195–2201, 2007.

[15] M. Collaud Coen, C. Praz, A. Haefele, D. Ruffieux, P. Kaufmann, and B. Calpini, "Determination and climatology of the planetary boundary layer height above the Swiss plateau by in situ and remote sensing measurements as well as by the COSMO-2 model," *Atmos. Chem. Phys.*, vol. 14, pp. 13205–13221, Dec. 2014.

[16] C. Münkel, N. Eresmaa, J. Räsänen, and A. Karppinen, "Retrieval of mixing height and dust concentration with lidar ceilometer," *Boundary-Layer Meteorology*, vol. 124, no. 1, pp. 117–128, 2007.

[17] B. Hennemuth and A. Lammert, "Determination of the atmospheric boundary layer height from radiosonde and lidar backscatter," *Boundary-Layer Meteorology*, vol. 120, pp. 181–200, 2006.

[18] W. P. Hooper and E. W. Eloranta, "Lidar measurements of wind in the planetary boundaryndary layer: The method, accuracy and results from joint measurements with radiosonde and Kytoon," *J. Climate Appl. Meteor.*, vol. 25, pp. 990–100, 1986.

[19] U. Löhnert, J. H. Schween, C. Acquistapace, K. Ebell, M. Maahn, M. Barrera-Verdejo, A. Hirsikko, B. Bohn, A. Knaps, E. O'Connor, C. Simmer, A. Wahner, and S. Crewell, "JOYCE: Jülich Observatory for Cloud Evolution," *Bull. Amer. Meteor. Soc.*, vol. 96, no. 7, pp. 1157–1174, 2015.

[20] J. H. Schween, A. Hirsikko, U. Löhnert, and S. Crewell, "Mixing-layer height retrieval with ceilometer and doppler lidar: from case studies to long-term assessment," *Atmos. Meas. Tech.*, vol. 7, pp. 3685–3704, Nov. 2014.

[21] T. Rose, S. Crewell, U. Löhnert, and C. Simmer, "A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere," *Atmospheric Research*, vol. 75, pp. 183–200, 2005.

[22] U. Löhnert and O. Maier, "Operational profiling of temperature using ground-based microwave radiometry at Payerne: Prospects and challenges," *Atmos. Meas. Tech.*, vol. 5, pp. 1121–1134, May 2012.

[23] S. Steinke, S. Eikenberg, U. Löhnert, G. Dick, D. Klocke, P. Di Girolamo, and S. Crewell, "Assessment of small-scale integrated water vapour variability during HOPE," *Atmos. Chem. Phys.*, vol. 15, pp. 2675–2692, Mar. 2015.

[24] U. Saeed, F. Rocadenbosch, and S. Crewell, "Synergetic use of lidar and microwave radiometer observations for boundary-layer height detection," in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, pp. 3945–3948, 26-31 July 2015.

[25] R. M. Measures, *Laser-Remote Sensor Equations, in Laser Remote Sensing: Fundamentals and Applications*. No. chap. 7, pp. 237-280, Krieger, Malabar, FL USA, 1992.

[26] D. Lange, J. Tiana-Alsina, U. Saeed, S. Tomás, and F. Rocadenbosch, "Atmospheric-boundary-layer height monitoring using a Kalman filter and backscatter lidar returns," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4717–4728, 2014.

[27] R. Barlow, *Estimation, in Statistics: A Guide to the Use of Statistical Methods in The Physical Sciences*. Wiley, Chichester, England, 1989. Chap. 5, pp. 68–95.

[28] F. Rocadenbosch, G. Vázquez, and A. Comerón, "Adaptive filter solution for processing lidar returns: Optical parameter estimation," *Applied Optics*, vol. 37, pp. 7019–7034, 1998.

[29] F. Rocadenbosch, C. Soriano, A. Comerón, and J. Baldasano, "Lidar inversion of atmospheric backscatter and extinction-to-backscatter ratios by use of a Kalman filter.," *Applied Optics*, vol. 38, pp. 3175–3189, 1999.

[30] D. Lange, F. Rocadenbosch, J. Tiana-Alsina, and S. Frasier, "Atmospheric boundary-layer-height estimation using a Kalman filter and a

frequency-modulated continuous-wave radar returns," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 3338–3349, June 2015.

[31] J. J. Moré, "The levenberg-marquardt algorithm: Implementation and theory," *Numerical Analysis*, pp. 105–116, 1977. Lecture Notes in Mathematics. Berlin, Germany: Springer-Verlag.

[32] U. Saeed, F. Rocadenbosch, and S. Crewell, "Performance test of the synergetic use of simulated lidar and microwave radiometer observations for mixing-layer height detection," in *Proc. SPIE*, vol. 9640, pp. 1–10, 2015.

[33] R. S. Bradley, F. T. Keimig, and H. F. Diaz, "Recent changes in the North American Arctic boundary layer in winter," *J. Geophys. Res.*, vol. 98, no. D5, pp. 8851–8858, 1993.

## Backscatter -vaisala @24.04.2013



(a)

## Potential temperature – MWR @24.04.2013



(b)
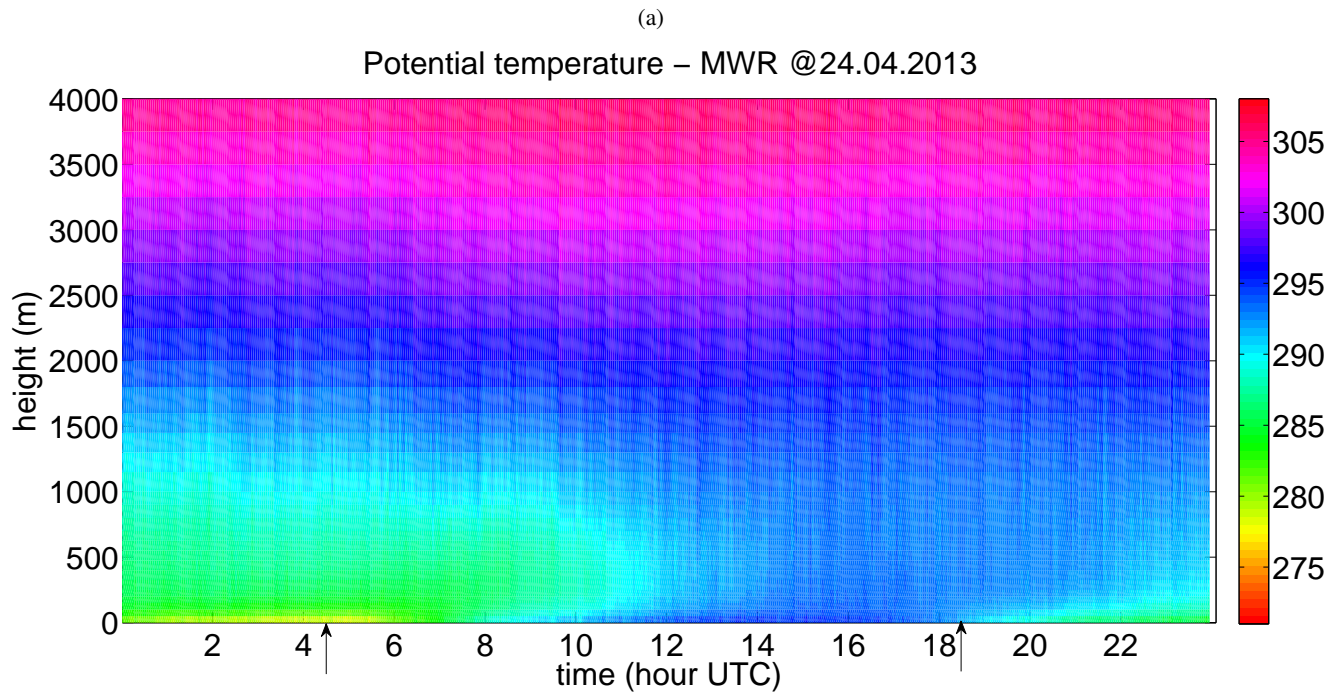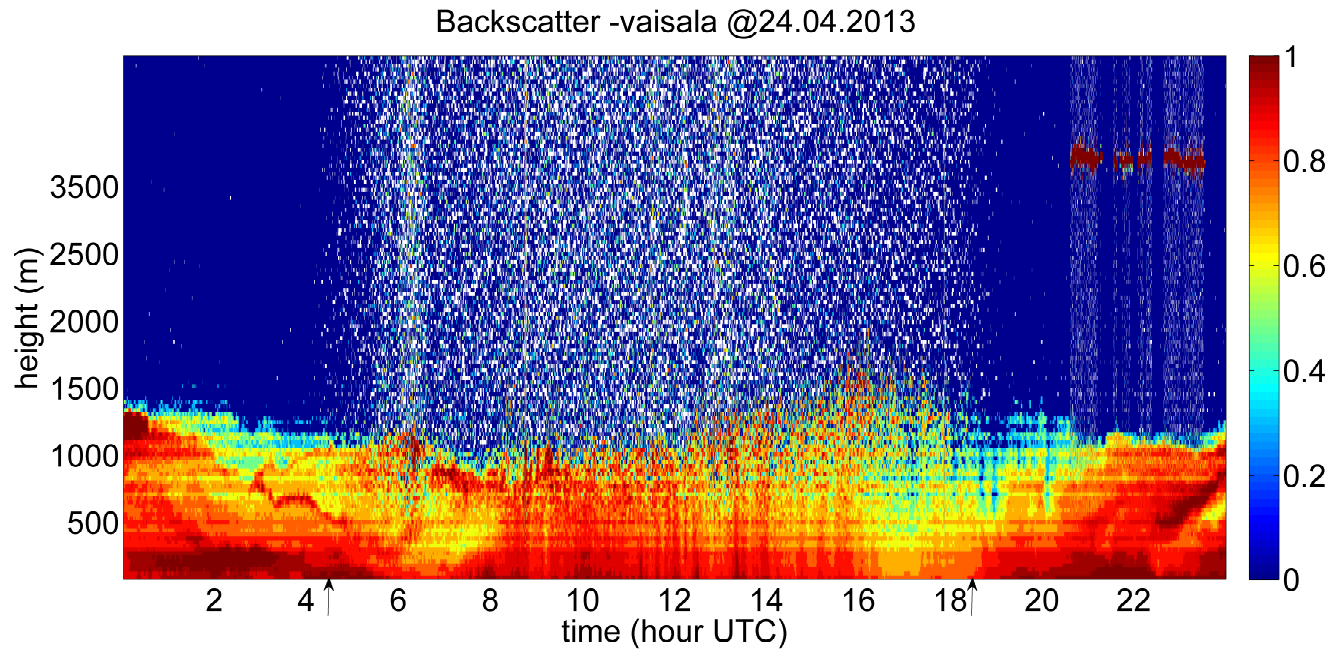
Fig. 1: Measured aerosol backscatter and potential temperature data from 24.04.2013, Jülich, Germany. Vertical arrows mark the sunrise and sunset times. (a) Colorplot of the backscatter profiles measured by the Vaisala CT25k ceilometer. (b) Colorplot from the MWR-retrieved potential temperature data.
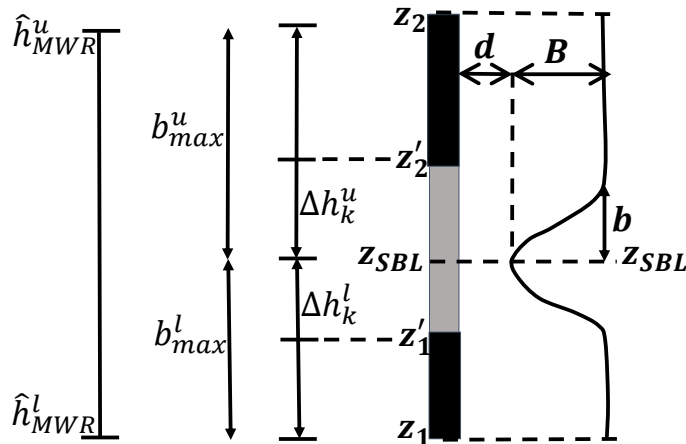
Fig. 2: Backscatter model variance profile: Conceptual sketch of a minimum variance region (MVR) modeled by an inverted Gaussian function along with fitting boundaries, $z_1$, $z_1'$, $z_2'$, and $z_2$. The inverted Gaussian bell on the right shows a MVR within the height interval, $\left[z_1', z_2'\right]$. The center of the MVR corresponds to the SBLH given by $z_{SBL}$ and its width is represented by parameter $b$. Parameter "$d$" stands for the background variance and parameter "$B$" indicates the variance amplitude above the background. Interval $\left[z_1, z_1'\right] U \left[z_2, z_2'\right]$ corresponds to the tail of the Gaussian outside the MVR. On the right, the fitting boundaries are labelled in terms of the parameters $h_{MWR}^{u,l}$, $\Delta h_k^{u,l}$, and $b_{max}^{u,l}$ (see Sect. III-D).
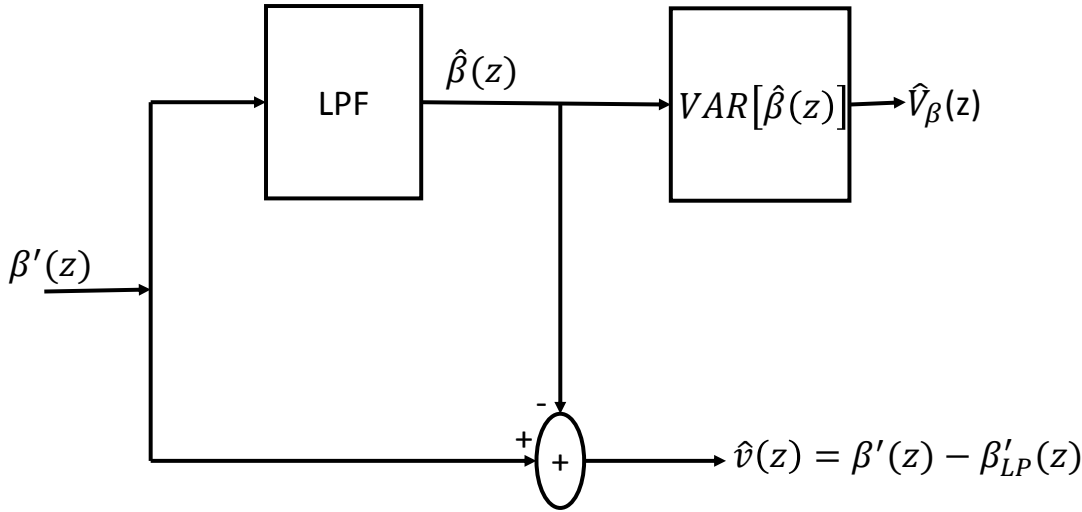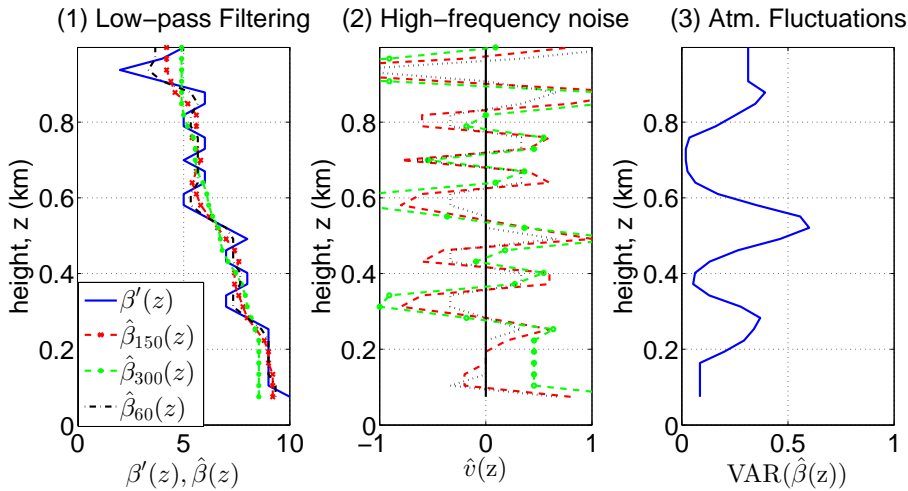
Fig. 3: Ceilometer signal-processing. (a) Block diagram to estimate the variance profile associated to the atmospheric backscatter, $V_{\hat{\beta}}(z)$. (b) Selection of the vertical smoothing window length and calculation of the backscatter variance, 24.04.2013, 21:00 UTC. (1) Vertical profiles of the de-noised backscatter coefficient, $\hat{\beta}_w(z)$, for different window lengths, $w = 60$ m (black line with circular markers), 150 m (red dashed line), and 300 m (green trace with dot markers) (rectangular-window smoothing as low-pass filter (LPF)). (2) High-pass residual noise, $\hat{v}_w(z)$, for the different window lengths chosen. (3) Estimated variance, $\hat{V}_{\beta}(z)$, for window length chosen (150 m) with two MVRs between the height ranges of approximately $300 - 500$ m and $550 - 900$ m.
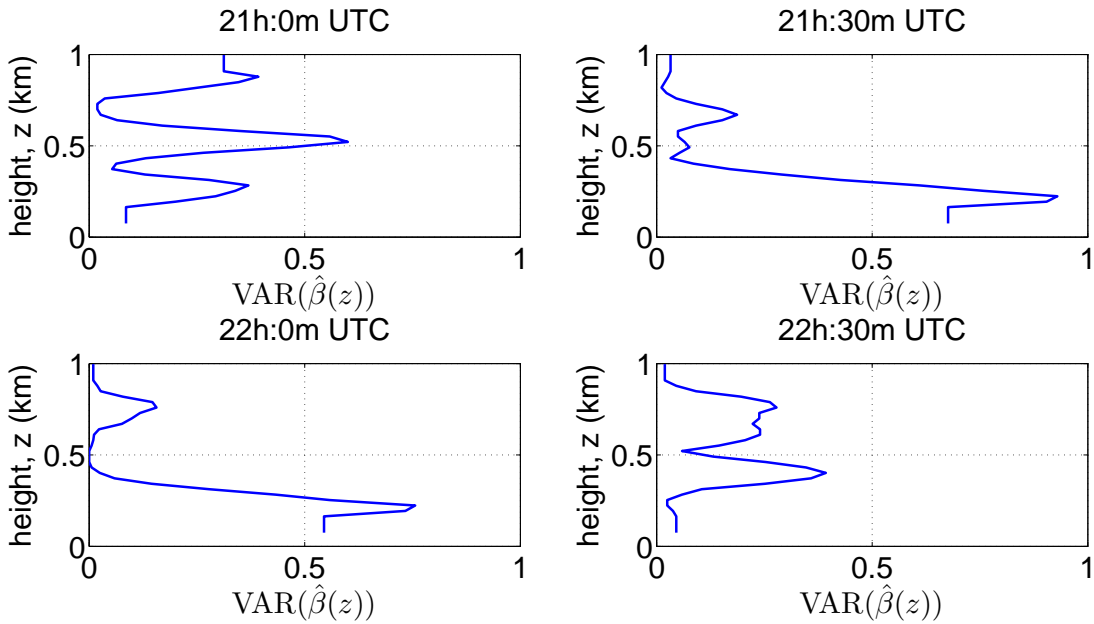
(a)
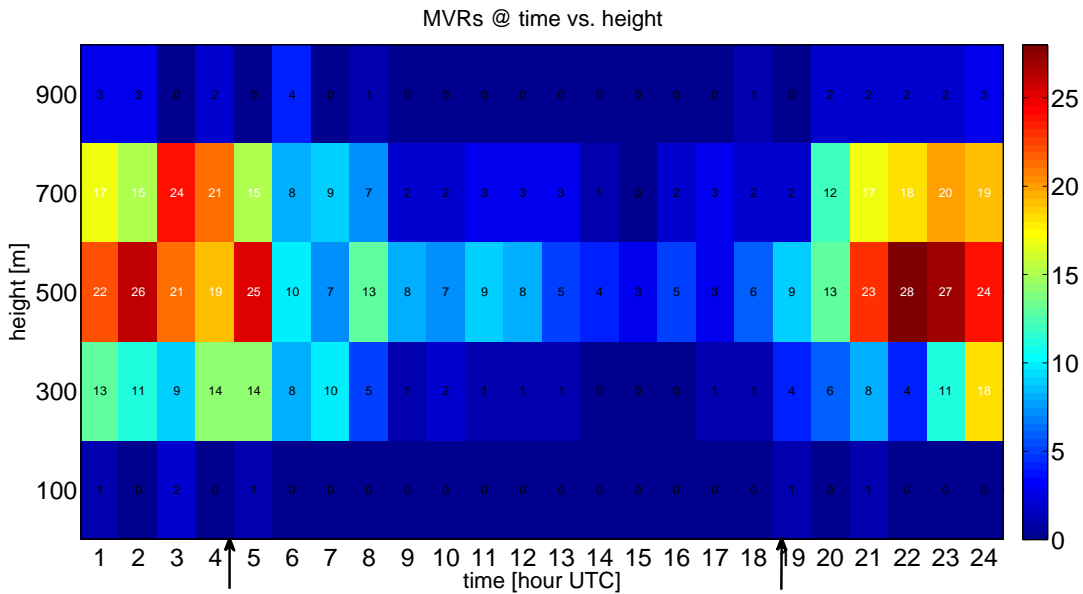


(b)

Fig. 4: (a) Vertical variance profiles based on the ceilometer measured aerosol backscatter data showing MVRs in the height range of $300 - 600$ m, 24.04.2013, 21:00-22:30 UTC. (b) Diurnal cycle of MVR occurrence based on one month of observations in April/May 2013. Number of MVR occurrence within 1 hr and 200 m is color coded. Vertical arrows mark the mean sunrise and sunset times.
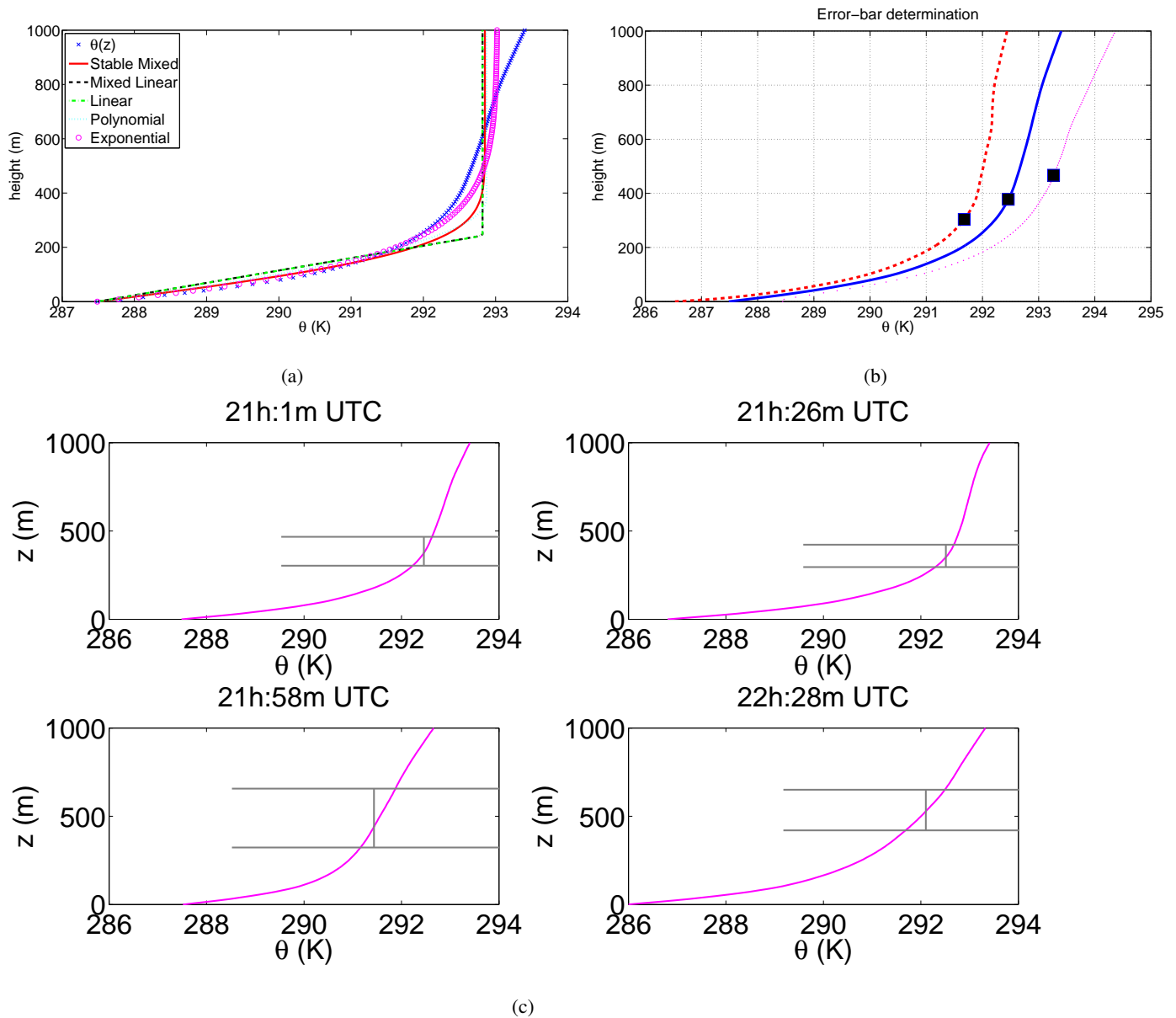
Fig. 5: MWR signal-processing. (a) Estimation of the SBLH from MWR data and potential temperature model profiles (Tab. I). (Blue crosses) Measured MWR-retrieved potential temperature, $\theta_{MWR}(z)$. (Solid color traces, see legends) Model profiles according to Tab. I. (b) Potential-temperature perturbed-error profiles, $\theta_{MWR}(z) + \epsilon(z)$ (upper error bound, dashed red), $\theta_{MWR}(z)$ (nominal profile, solid blue here and blue crosses in Fig. 5a), and $\theta_{MWR}(z) - \epsilon(z)$ (lower error bound, dotted red), used to estimate the errorbars associated to the MWR-retrieved potential temperature profile, $\theta_{MWR}(z)$. (c) Time-snapshots of MWR-retrieved potential temperature profiles with uncertainty due to measurements errors, (Eq. 28) at four time-instants from 24.04.2013, 21:00-22:30:00 UTC.
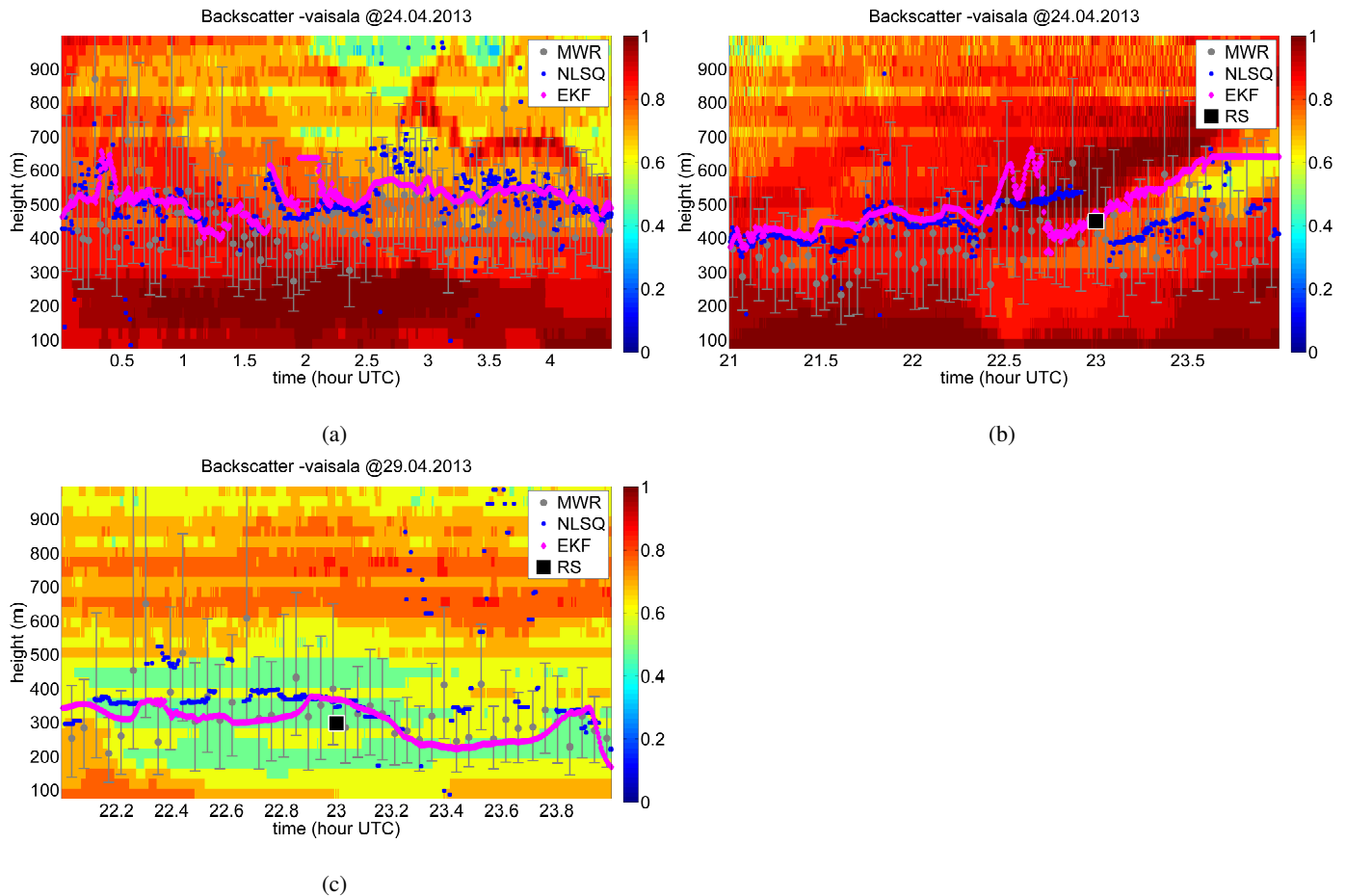
Fig. 6: Test cases based on data from HOPE campaign at Jülich, Germany. (a) Colorplot of the ceilometer backscattered signal for 24.04.2013, 00:00-04:30 UTC measurements along with SBLH estimates from the EKF and the NLSQ approaches. (b) Colorplot of the ceilometer backscattered signal for 24.04.2013, 21:00-24:00 UTC measurements along with SBLH estimates from the EKF and the NLSQ. SBLH estimate from the radiosonde (black square) launched at 23:00 is also shown. (c) Colorplot of the ceilometer backscattered signal for 29.04.2013, 22:00-24:00 UTC measurements along with SBLH estimates from the EKF and the NLSQ. (All panels) MWR-EKF search ranges $[z_{1,k}, z_{2,k}]$ (Eq. 29 and Fig. 2) are plotted as gray vertical bars.