Motivating a synergistic mixing-layer height retrieval method using backscatter lidar returns and microwave-radiometer temperature observations

M. P. Araújo da Silva, *Student Member, IEEE*, F. Rocadenbosch, *Senior Member, IEEE*, Robin L. Tanamachi, and Umar Saeed

Corresponding author: Francesc Rocadenbosch, roca@tsc.upc.edu

Abstract

Mixing-Layer-Height (MLH) retrieval methods using backscattered lidar signals from a ceilometer (Jenoptik CHM-15k Nimbus) and temperature profiles from a Microwave Radiometer (MWR, HATPRO RPG) are compared in terms of their complementary capabilities and associated uncertainties. The Extended Kalman Filter (EKF) is used for MLH retrieval from backscattered lidar signals and the parcel method is used for MLH retrieval from MWR-derived potential-temperature profiles.

The two principal sources of uncertainty in ceilometer-based MLH estimates are (i) incorrect layer attribution ($\sim$ hundreds of meters) and (ii) noise-induced errors (about 50 m at 3σ). MWR MLH uncertainties comprise (i) the total uncertainty in the retrieved potential temperature profile and (ii) ± 0.5 K uncertainty in the surface temperature. Ceilometer- and MWR-based MLH estimates are in turn compared with reference to MLH estimates from radiosoundings. Twenty one measurement days from the HD(CP)^2 Observational Prototype Experiment (HOPE) campaign at Jülich, Germany are considered.

It is shown that the MWR can track the full Mixed Layer (ML) diurnal cycle (i.e., including morning and evening transitions) with height-increasing error bars. The ceilometer-EKF MLH estimates are much smaller errorbars than those from the MWR under the well-developed clear-sky ML but the ceilometer-EKF is prone to ambiguous tracking some multilayer scenarios (e.g., the residual layer). We therefore introduce the synergistic MLH retrieval approach that combines both ceilometer and MWR estimates in order to optimize the benefits of both.

Marcos P. Araújo da Silva and Francesc Rocadenbosch are with CommSensLab-UPC, Department of Signal Theory and Communications, Universitat Politècnica de Catalunya (BarcelonaTech—UPC), E-08034 Barcelona, Spain. F.R. is also with the Institute of Space Studies of Catalonia (IEEC), E-08034 Barcelona, Spain (e-mails: marcos.silva@upc.edu and roca@tsc.upc.edu).

Robin L. Tanamachi is with the Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN 47907 USA (rtanamachi@purdue.edu).

U. Saeed is with the Department of Communications and Networking, Aalto University, Espoo, 00076, Finland.

 DOI 10.1109/TGRS.2022.3158401

Index Terms

Lidar, laser radar, ceilometers, microwave radiometry, remote sensing, error analysis, signal processing, atmospheric boundary layer height, mixed layer.

I. INTRODUCTION

Accurate monitoring of the atmospheric boundary layer (ABL) is a subject of wide interest. The ABL, by definition is directly affected by interactions with the surface of the Earth on a time scale of an hour or less [1], and it is also the layer of the atmosphere within which humans live. The mixed layer height (MLH) is a parameter of interest for many applications, including weather forecasting, air quality and chemical dispersion models, and aviation. In fair weather conditions, the interior of the ABL is well-mixed by convective turbulence (hence the term "mixed layer", or ML), and exhibits near-constant potential temperature (θ) and water vapor mixing ratio (r) throughout most of its depth. However, no remote or *in situ* instrument exists that can directly measure MLH. Instead, a proxy or tracer for the top of the ML must be used. Such tracers include gradients of aerosols, temperature, wind characteristics, or energy fluxes [2], [3]. Ground-based remote-sensing instruments that can detect these tracers include lidar, radar, and sodar, which are active, and microwave radiometer (MWR), which is passive.

Because of the varying properties of these MLH tracers and estimation methods, their accuracy varies widely. Different studies have proven the reliability of MWR retrievals by comparison with radiosoundings. Radiosondes, usually comprising an expendable, balloon-borne package of sensors, are a recognized reference instrument for boundary layer monitoring. Xu et al. [4] compared MWR-derived temperature profiles averaged over a 30-min (full-span) window centered around the radiosonde profiles. Analysing 403 (760) radiosonde launches from clear-sky (cloudy) days, they obtained correlation coefficients ≥ 0.85 up to ~ 7 km AGL. Good agreement between MWR- and radiosonde-derived temperatures is also described by Löhnert and Maier [5], who reported differences lower than 0.5 K between the profiles up to 4 km AGL. In relation to MLH assessment, Collaud Coen et al. [6] applied the parcel method [7] to MWR- and radiosonde-retrieved θ profiles, obtaining median bias of -25.5 m and coefficient of determination $\rho^2 = 0.75$ ($\rho = 0.87$) over 100 samples. In a related study, de Arruda Moreira et al. [8] found excellent agreement during convective conditions, with a root mean square error (RMSE) $= 190$ m and $\rho = 0.96$.

The MWR has higher temporal resolution (e.g., a few minutes) than the radiosonde, whose operational frequency may be as low as two launches per day. However, its vertical resolution decreases with height, hence increasing uncertainty in retrieved quantities. In contrast, the lidar is an active remote sensing instrument designed to measure vertical profiles of aerosol backscattering with high spatiotemporal resolution. Aerosol concentrations are usually relatively high and constant throughout the ML and much

lower aloft, thereby enabling lidar-based MLH estimation. In a fair-weather convective boundary layer (CBL), MLH-lidar and MLH-MWR tend to converge. Belegante et al. [9] compared MLH estimates retrieved from elastic-lidar range-corrected-signal (RCS) profiles averaged over 30-min intervals with those from MWR-derived virtual potential temperature, finding high correlation ($\rho \approx 0.98$) in CBL conditions. However, during evening transition times (ETT; i.e., at and after sunset, when thermally-driven turbulent mixing ceases), MLH-MWR is more accurate than MLH-lidar, which generally tended to track the RL (i.e., a remnant layer left over from earlier turbulent mixing) height [9]. Cimini et al. [10], using training data from multi-frequency, multi-angle MWR and lidar observations, designed a multivariate linear regression method to assess the mixing-layer height (MLH) directly from MWR brightness-temperature observations instead of retrieved profiles.

Lange et al. [11] (CommSensLab-UPC) estimated MLH by applying an extended Kalman filter (EKF) [12] to backscattered lidar returns (hereafter, MLH-LC-EKF). They showed that, for a CBL with moderate-to-low signal-to-noise ratio (SNR > 5), MLH-LC-EKF was more accurate than classical approaches such as the threshold [13], gradient [14], logarithmic gradient [15], inflection point [16], and variance methods [17]. Additionally, Banks et al. [18] showed the MLH-LC-EKF reliability against radiosonde-derived Bulk Richardson Number profiles and against Weather Research and Forecasting (WRF) model predictions for clear-air and pre- convective storm cases. Based upon these works, de Arruda Moreira et al. [8] found correlated results among MLH estimates in the CBL derived from LC-EKF, MWR and Doppler Wind Lidar (DWL) combinations. The EKF has also been used as an MLH retrieval method when applied to S-band radar returns [19], [20]. In a different context, Barrera-Verdejo et al., [21] combined brightness-temperature information from a MWR and water-vapour mixing-ratio Raman lidar profiles in order to derive absolute humidity vertical profiles.

Most of the previous work done for ML retrieval using backscattered lidar signals or temperature profiles has been focused on "stand-alone" retrieval methods only. However, tentative derivation of a synergistic algorithm spanning the full diurnal cycle and the study of the uncertainty sources associated to MLH estimation and their statistical impact on the retrievals has not received much attention [22].

To fill this void, this paper focuses on synergistic retrieval of MLH estimates with low uncertainty by combining ceilometer and MWR MLH-based retrievals using the EKF and the parcel method as respective estimators. Towards this end, this paper aims to study: (i) the performance of these two commonly-used algorithms under different atmospheric scenarios, (ii) the impact of key error sources, namely, measurement and retrieval errors, on MLH estimates and (iii) evaluate the performance of the proposed synergistic ceilometer-MWR method. The simplified processing chain of the ceilometer and MWR data is summarised in Fig. 1.

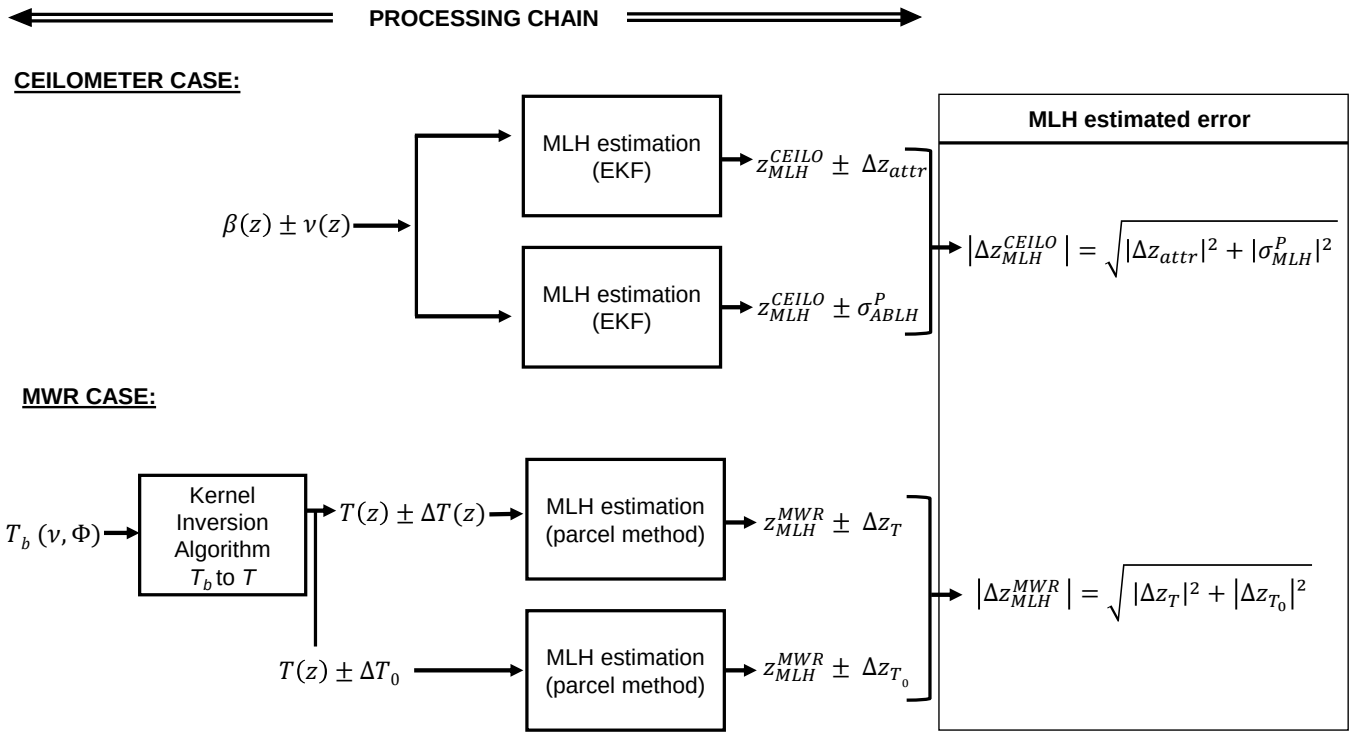

Figure 1: Block diagram illustrating the ceilometer (top) and MWR (bottom) processing chains used to estimate the mixing-layer height (MLH) and related error estimates. z is the vertical coordinate (height). $\beta(z)$ is the attenuated backscatter, $\nu(z)$ is the related corrupting noise, z_{MLH}^{CEILO} is MLH-LC-EKF, Δz_{attr} is the uncertainty of MLH-LC-EKF due to incorrect layer attribution, and σ_{MLH}^{P} is the MLH error component due to noise. $T_b(\nu, \phi)$ is the MWR brightness temperature measured at frequency ν and elevation angle ϕ, $T(z)$ is the retrieved temperature profile, Δz_T is the MLH-MWR error component due to the total uncertainty of $T(z)$, Δz_{T_0} is the MLH-MWR error component due to uncertainty in the auxiliary measured surface temperature T_0. Δz_{MLH}^{CEILO} and Δz_{MLH}^{MWR} stand for the total estimated error for MLH-LC-EKF and MLH-MWR, respectively.

This paper is organised as follows: Sect. II introduces instruments and data sets. Sect. III revisits the proposed MLH estimation methods and related error sources, assesses their uncertainties on the MLH estimates via error propagation from both ceilometer- and MWR-based methods, and introduces the synergistic method. Sect. IV presents statistical analysis and discussion of results based on 20 days selected from the HD(CP)^2 Observational Prototype Experiment (HOPE) [23] campaign. Finally, Sect. V gives concluding remarks.

Owing to the various combinations of boundary layer terminology, instruments, and methods used in this paper, we provide an acronym table for ease of reference (Table I).

Table I: List of acronyms and abbreviations used throughout this manuscript.

Acronym	Definition	Reference(s)
ABL	Atmospheric boundary layer	[1]
CBL	Convective boundary layer	[1]
ML	(Well-)mixed layer	[1]
FT	Free troposphere	[1]
EZ	Entrainment zone	[1]
RL	Residual layer	[1]
MLH	Mixed layer height	[1]
CBH	Cloud base height	[24]
MTT	Morning transition time	[1]
ETT	Evening transition time	[1]
JOYCE	Jülich ObservatorY for Cloud Evolution	[25]
HOPE	HD(CP)^2 Observational Prototype Experiment	[23]
MWR	Microwave radiometer	[25], [26]
LC	Lidar ceilometer	[24]
DWL	Doppler wind lidar	[26], [27]
RS	Radiosonde	[28]–[30]
EKF	Extended Kalman filter	[12]
VVSTD	Vertical Velocity Standard Deviation	[27], [31]
MLH-MWR	MLH estimated by application of the parcel method to MWR measurements	[3], [7]
MLH-LC-EKF	MLH estimated by application of the EKF method to LC measurements	[11], [19]
MLH-DWL	MLH estimated by application of the VVSTD method to DWL measurements	[27]
MLH-RS	MLH estimated by application of the parcel method to RS measurements	[3], [7]
SYN	Synergistic method optimally combining MLH-MWR and MLH-LC-EKF	This manuscript
MLH-SYN	MLH estimated by the SYN method	This manuscript

II. INSTRUMENTS AND DATA SET

Data used in this work was collected at the Jülich ObservatorY for Cloud Evolution (JOYCE), which is located in Forschungszentrum, Jülich, Germany (50°54′31″ N, 6°24′49″ E, 111 m MSL). The topography

in the area of JOYCE is generally flat apart from large lignite open pit mines. Farming, open-cast coal mining areas with major power plants and patchy settlements characterised the 50-km periphery. The climate is characterised by a temperate, humid climate with warm summers [25]. JOYCE contains a number of active and passive permanently installed remote sensing, and in-situ instruments aimed to the study of clouds and atmosphere. The HOPE campaign [23] was conducted at JOYCE from April 2013 to May 2013. One of the principal aims of HOPE was to characterize the evolution of the ABL over JOYCE for forecasting applications. We chose to examine the data from this campaign because they contain long-duration, simultaneous observation of MLH tracers by multiple independent instruments, a situation ideal for validation. For brevity, only those JOYCE instruments used in HOPE are described below. The reader is referred to [25], [32] for a complete listing.

A. Ceilometer

A lidar ceilometer is a single-wavelength elastic-backscatter lidar characterised by a low energy-aperture product. Under moderate-to-clear-air atmospheres (optical thickness, $\tau < 1$) and, particularly, towards the near-infrared, the profile of the attenuated-backscatter coefficient is essentially proportional to the aerosol concentration in the ABL [11], [24], [27].

Two LCs are used in this work. The first is a Jenoptik CHM-15k Nimbus, an $8-\mu$J, 1064-nm wavelength, 5-ns-pulse duration, 5-to-7-kHz repetition-rate ceilometer. Under clear-sky conditions, the maximum sounding range is about 15 km with a range resolution of 15 m. The temporal resolution of the instrument is 15 s. The Jenoptik ceilometer is a bi-axial system with separate optics for the transmitter and receiver so that the optical overlap is reached at about 350 m. The instrument provides range- and overlap-corrected profiles of the normalized backscattered power (i.e., the attenuated backscatter-coefficient profile).

The second LC used in the HOPE campaign was a Vaisala CT25K, used to monitor cloud-base height (CBH) and precipitation, and to complement Jenoptik's observations below 350 m (night-time and morning/afternoon transition periods). The Vaisala CT25K is a $1.6-\mu$J, 905-nm wavelength, 100-ns pulse duration, 5.6-kHz repetition-rate LC designed to retrieve profiles of the attenuated-backscatter coefficient in the sounding range from 60 m to 7 km, with a range resolution of approximately 30 m. The temporal resolution of the instrument is 15 s (including 3.3 s for processing and data transmission). Because of the lower pulse energy of the Vaisala LC as compared to the Jenoptik (and because both systems operate in similar, near-infrared wavelengths), Jenoptik's vertical profiles of the attenuated backscatter coefficient exhibit a comparatively high SNR.

B. MWR: Humidity And Temperature PROfiler (HATPRO)

A MWR measures the radiative emission of atmospheric gases. The emissions from molecular oxygen, in particular, are proportional to atmospheric temperature when the water vapor mixing ratio remains constant. Measurements at different frequency bands and elevation angles can be used to derive several physical quantities such as temperature, water vapor, integrated water vapor, and liquid water path. The Humidity And Temperature PROfiler (HATPRO) MWR manufactured by Radiometer Physics GmbH (RPG) [25], [26] measures the atmospheric brightness temperature at 14 frequencies in two bands and at six angles (depending on user's settings). Measurements in the K-band (7 channels), $22 - 31$ GHz, are used for water vapor and liquid water retrieval, and in the V-band, $51 - 58$ GHz, for temperature retrieval. The MWR instrument exhibits good temporal resolution (about 2.7 minutes). In principle, MWR can operate in all conditions except for rain, when the radiation measured is dominated by the emission and scattering from raindrops.

Brightness-temperature measurements are converted into a temperature profile by means of a statistical retrieval algorithm [33]–[36] together with auxiliary atmospheric temperature and pressure data. The latter are measured separately by surface-based *in situ* sensors. Ultimately, the retrieved potential temperature profile is used to estimate the MLH. In this formulation, two main error sources are delineated: (i) MLH estimation errors, Δz_T, originating as a total uncertainty in the retrieved temperature profile ($\Delta T(z)$ in Fig. 1) and (ii) MLH errors due to uncertainties in the auxiliary measurement of surface temperature, Δz_{T_0}. MLH estimates and related errors are computed according to the different spatial resolutions of each instrument. The MWR-derived temperature profile exhibits a coarse ($\geq$ 50 m) vertical resolution that increases with height [5], [35], and which is specific to the retrieval algorithm and to the number of measurement channels (sounding frequencies) being used. Because the number of independent pieces of information contained in the brightness temperature measurements at different frequencies and scanning angles is limited, the resulting Degree of Freedom (DoF) is low (≈ 4 for temperature boundary-layer profiling) [5]. As a result, the vertical resolution of the retrieved quantities substantially decreases (i.e., becomes coarser) with height. As a general rule, the vertical resolution is approximately equal to the height above the surface. For example, the vertical resolution at 1 km AGL is roughly 1 km (i.e., the retrieved temperature at 1 km AGL is valid for the 500 – 1500 m layer).

C. Doppler Wind Lidar

Because the Free Troposphere (FT) is generally not as turbulent as the ML, vertical gradients in the standard deviation of velocity can be used as a tracer of the MLH [31], [37]. A DWL indirectly measures the ABL mixing process via vertical wind velocity variance [26], [27]. Driven by the technological progress

in the wind energy industry in the last couple of decades, economical and useful DWL systems have been developed [38].

The DWL used in the HOPE campaign is the HALO Photonics Streamline Wind Lidar [27], [38], [39], a coherent system with an average pulse energy of 100 μJ and a frequency of 15 kHz. The vertical resolution of this system is 30 m, and the maximum range is about 8 km. However, in practice, its effective maximum range becomes limited by the reduced aerosol content above the ABL.

D. Radiosonde

The Radiosonde is a *de facto* standard for reference in the atmospheric sciences [28]–[30]. Radiosondes measure *in situ* profiles of the atmospheric thermodynamic state (temperature, pressure, water vapor and wind), and are usually launched attached to a large balloon. A radiosonde can rise up to 40 km in height over the course of several hours, though most water-vapour sensors usually cease to operate properly below the tropopause ($\sim$ 15 km in the midlatitudes).

The radiosonde used in this work is the Graw DFM-09 manufactured by Graw GmbH, which includes temperature, pressure, humidity and GPS sensors (20 channels). Wind speed and direction are determined from the changes in the GPS position and GPS velocity vector. The transmission rate is one full set of observations per second. The main drawbacks of using radiosondes for MLH determinations are the sparse temporal resolution (owing to the expense associated with each launch) and horizontal drift on ascent. During HOPE campaign, 226 soundings were launched up to a maximum of six per day (07:00, 09:00, 11:00, 13:00, 15:00, and 17:00 UTC; local time, LT=UTC+1h) from a site located 3.8 km southeast of JOYCE.

III. METHODS AND UNCERTAINTIES OF MLH ESTIMATION

When comparing ceilometers and MWRs, the instruments at the focus of this paper, ceilometers typically have higher vertical resolution (e.g., 15 m for the Jenoptik CHM-15k Nimbus ceilometer versus 50 m or greater for the MWR). Because the ceilometer relies on attenuated backscatter returns from atmospheric aerosols and molecules, which are proxies of the thermodynamic state of the atmosphere, to identify the MLH, layer-attribution problems are very common. On the other hand, MWR-derived temperature profiles have a much coarser vertical resolution than profiles from LCs (Sect. II). Additionally, parcel method-based MLH estimates suffer from uncertainties associated with surface temperature Δz_{T_0}. We now explore these sources of uncertainty in greater detail.

A. MLH estimation from ceilometer data

MLH estimation method: Several classical methods for LC-based MLH estimation, such as the gradient method [14], [40], the inflection-point method [16] and the variance method [17], among others [13], [41], are based on the detection of a meaningful ML-to-FT sharp transition in the vertical profile of the attenuated backscatter coefficient. However, thermal updrafts, intermittent turbulence, and measurement noise often lead to time-inconsistent MLH retrievals even in single aerosol layer scenarios. Apart from noise, these fluctuations represent real physical processes in the atmosphere, but they can complicate MLH tracking. To counteract detrimental effects of these fluctuations on MLH tracking, backscatter profiles are often time and/or height averaged [11]. Therefore, temporal resolution of LC-based MLH estimates is usually relatively low ($\sim$ 30 min). LC-based algorithms for MLH estimation (in convective and stable regimes) in recent literature [6], [42], [43] use time-continuity and morphological criteria based on a combination of empirically-tuned gradient and variance criteria, signal and SNR conditions, and climatological data.

In this study, we apply the EKF MLH estimation method [11], [19], which departs from previous works of [44], using a time-adaptive, optimal predictive model to delineate the shape of the sharp ML-to-FT transition (Fig. 2a). The Kalman filter is essentially the Wiener solution [45] of the optimal filter problem in which one wants to compute a statistical estimate of an unknown signal (the MLH as a function of time) using a related signal (i.e., the ceilometer attenuated backscatter height profile as a function of time) to produce the estimate as an output. The two main distinguishing features of the Kalman filter formulation are: (i) vector modelling of the random processes under study (i.e., the MLH and shape parameters defining the ML-to-FT transition as a function of time), and (ii) recursive processing of the input noisy measurements (the ceilometer's) at each successive discrete time. This time-adaptive behaviour of the filter is based on minimization of the mean squared error over time of the so-called *a posteriori* error covariance matrix (the term *a posteriori* meaning "once the filter is updated with the present-time measurement", i.e., from the ceilometer). The reader is referred to [11], [19] for a complete derivation of the EKF-based method for MLH tracking in LC data (hereafter, MLH-LC-EKF), particularly Appendix A of [11] for a summary of the filter's constitutive equations and proof of its optimality. In following Sect. III-A1 the MLH-LC-EKF is briefly summarised for self-contained purposes and notation definition.

As a result, MLH-LC-EKF [11] enables consistent time-tracking of the MLH without need of averaging techniques or training from ancillary climatological records. Because the filter estimates are generated at the same temporal and spatial resolution of the input measurement data, the filter can even be implemented as a real-time processor. Best performance of the EKF with reference to the classical methods above has been shown in [11], [18], [46] with reference to different parameterizations of the WRF model.

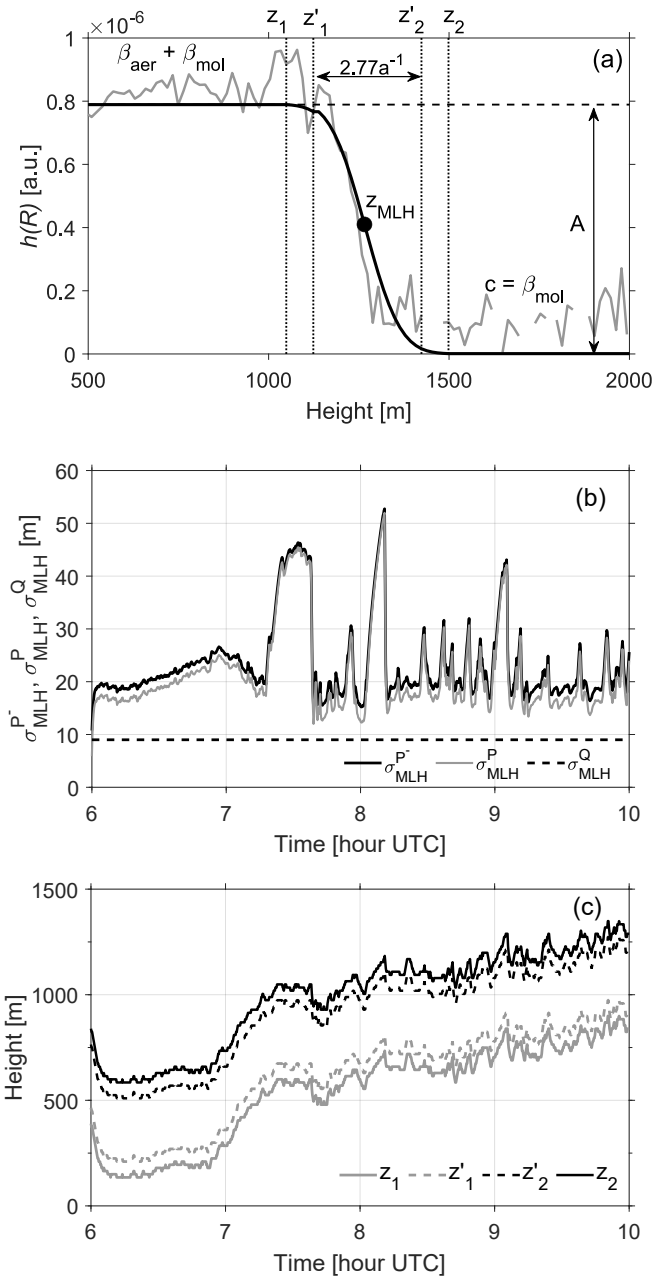

Figure 2: Example of the MLH-LC-EKF estimation technique (20 April 2013, 0600-1000 UTC). (a) The ML-to-FT transition model. (Grey trace) Example of background-subtracted attenuated backscatter-coefficient profile measured by the LC. (Black curve) Fitted erf-model profile described in [11]. (b) MLH-LC-EKF *a priori* error, $\sigma_{MLH}^{P^-}$, *a posteriori* error, σ_{MLH}^{P}, and state-noise standard deviation, σ_{MLH}^{Q} as a function of time. (c) Temporal evolution of search boundaries z_1, z_1', z_2' and z_2.

1) Review of the MLH-LC-EKF: Central to the MLH-LC-EKF method is the assumption of the erf-like ML-to-FT model,

$$h(z; z_{MLH}, a, A, c) = \frac{A}{2}\left\{1 - erf\left[\frac{a}{\sqrt{2}}(z - z_{MLH})\right]\right\} + c, \tag{1}$$

where z_{MLH} is the MLH, a is a scaling factor related to the entrainment zone (EZ) thickness ($2.77a^{-1}$) [31], [40], A is the total backscatter coefficient, and c is the free troposphere (FT) molecular backscatter background, which acts as an offset term to the filter.

The erf model depicted in Fig. 2a models the ML-to-FT transition of the attenuated backscatter coefficient measured by the ceilometer. This model is used as a proxy of the total atmospheric backscatter coefficient under the assumption of clear-to-moderately cloudy sky conditions (optical thickness, $\tau < 1$). The shape parameters of this model profile along with the MLH parameter itself give rise to the formulation of the state vector (to be estimated),

$$\boldsymbol{x}_k = [z_{MLH,k}, a_k, A_k, c_k]^T, \tag{2}$$

where subscript k is a discrete time. $z_{MLH,k}$ is the key parameter of interest, and a_k, A_k, and c_k are auxiliary parameters determining the change in shape of the ML-to-FT interface with time.

The *state-vector model* represents the transition of the state vector from time t_k to t_{k+1}. It is formulated as

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{w}_k, \tag{3}$$

where $\boldsymbol{w}_k$ is the state-noise vector with diagonal covariance matrix, $\boldsymbol{Q}_k = \boldsymbol{E}[\boldsymbol{w}_k \boldsymbol{w}_k^T]$, where $\boldsymbol{E}$ denotes the ensemble mean (or expectation) operator [47]. For enhanced filter stability [48], $\boldsymbol{Q}_k$ is modelled in stationary diagonal form as $\boldsymbol{Q} = diag[\boldsymbol{\sigma_Q^2}], \boldsymbol{\sigma_Q} = [\sigma_{z_{MLH}}, \sigma_a, \sigma_A, \sigma_c]$, where $\sigma_{z_{MLH}}, \sigma_a, \sigma_A$ and σ_c are the guessed standard deviations associated with the state-vector components, $z_{MLH},\ a,\ A$ and c, respectively. For example, $\sigma_{z_{MLH}}$ models the standard deviation of the MLH (a random variable) around its mean value. In practice, for simplicity, the input vector $\boldsymbol{\sigma_Q}$ is constructed as

$$\boldsymbol{\sigma_Q} = \mu_Q \hat{\boldsymbol{x}}_{\boldsymbol{0}}^{-}, \tag{4}$$

where μ_Q is the so-called Q-intensity factor (a scalar) and $\hat{\boldsymbol{x}}_{\boldsymbol{0}}^{-}$ denotes the initial guess of the state vector at filter start-up, to be specified by the user. In what follows, we have used $\mu_Q = 0.1$ (10%). This means that if we assume an initialization $z_{MLH,0} = 1500$ m then we expect MLH fluctuations of approximately ± 150 m (10%) at 1σ. Because the state vector is recursively recomputed at each filter step, an intensity factor $\mu_Q = 0.1$ is usually more than sufficient to search the full ML height span. Increasing this factor above 10% increases the search "nervousness" of the filter. This effect is usually not beneficial because it

may lead the filter to jump between different aerosol layers and, hence, be more prone to divergence. The initial guess vector $\hat{\boldsymbol{x}}_{\boldsymbol{0}}^{-}$ can easily be estimated by plotting the erf model against one measured ceilometer attenuated backscatter profile as in Fig. 2a.

The *measurement model* relates the ceilometer measurement vector, $\boldsymbol{y_k}$, to the state-vector as

$$\boldsymbol{y_k} = h(\boldsymbol{x_k}) + \boldsymbol{v_k}, \tag{5}$$

where h is the erf-like ML-to-FT function model given by Eq. 5 above, $\boldsymbol{v_k}$ is the observation noise at time t_k with noise covariance matrix, $\boldsymbol{R}_k = \boldsymbol{E}\left[\boldsymbol{v_k}\boldsymbol{v_k^T}\right]$, and z is the vertical range. The measurement vector, $\boldsymbol{y_k} = \left(y_{1,k}, y_{2,k}, \ldots, y_{N,k}\right)$ is the noise-corrupted ceilometer attenuated backscatter signal at discrete ranges, $z = z_i, i = 1 \ldots N$. Because there is only one single measurement realization available at each time $\boldsymbol{t_k}$, the instantaneous noise covariance matrix $\boldsymbol{R}_k$ is estimated in piece-wise form over range cells instead of time cells as described in [19], Eq. (19). The nonlinear model of Eq. 5 is linearised through its Jacobian, which is passed to the filter.

2) Error sources: There are two key sources of uncertainty concerning MLH estimation from ceilometer data: (i) layer-attribution errors and (ii) noise-induced errors.

(i) Impact of layer-attribution errors on the estimated MLH, Δz_{attr}: Layer-attribution errors arise from the existence of multiple layers or gradients in the attenuated backscatter profile. Depending upon the number of layers and their separation, the estimated MLH can be significantly different from the actual MLH. For example, during the evening transition time (ETT; also called the afternoon-to-evening transition or AET, see [49]) under quiescent conditions, the ML is replaced by the residual layer (RL) [1]. In this case, it often happens that the MLH-LC-EKF estimate follows the RL because RL-to-FT aerosol gradient characteristics are similar to those of the ML top [50] , particularly in the 1–2 hr period following local sunset. Provision of acceptable initial guesses for the state-vector, $\hat{\boldsymbol{x}}_0^{-}$, state-vector covariance, $\boldsymbol{Q}$, and a priori state-vector error covariance matrix, $\boldsymbol{P_0^{-}}$, are key to preventing layer-attribution errors. This is especially true during the morning transition time (MTT, i.e., at and after local sunrise), when the ML starts to develop and there are relatively steep backscatter aerosol gradients between the ML and the RL or FT above. An example of this phenomenon is illustrated in Fig. 3a. Previous work on assessing the uncertainty of the MLH estimate due to layer-attribution errors has been carried out by [50].

The *a priori* and *a posteriori* error-covariance matrices are defined as

$$\boldsymbol{P}_k^{-} = \boldsymbol{E}\left[\boldsymbol{e}_k^{-}\boldsymbol{e}_k^{-T}\right], \; \boldsymbol{P}_k = \boldsymbol{E}\left[\boldsymbol{e}_k\boldsymbol{e}_k^{T}\right], \tag{6}$$

respectively, where $\boldsymbol{e_k^{-}} = \boldsymbol{x_k} - \hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}$ is the *a priori* error and $\boldsymbol{e_k} = \boldsymbol{x_k} - \hat{\boldsymbol{x}}_{\boldsymbol{k}}$ is the *a posteriori* error, i.e., before and after assimilating the current measurement ($\boldsymbol{y_k}$). Here, $\boldsymbol{x_k}$ is the true atmospheric state (unknown) and $\hat{\boldsymbol{x}}_{\boldsymbol{k}}^{-}$ and $\hat{\boldsymbol{x}}_{\boldsymbol{k}}$ are the *a priori* and *a posteriori* state vectors estimated by the filter, respectively.

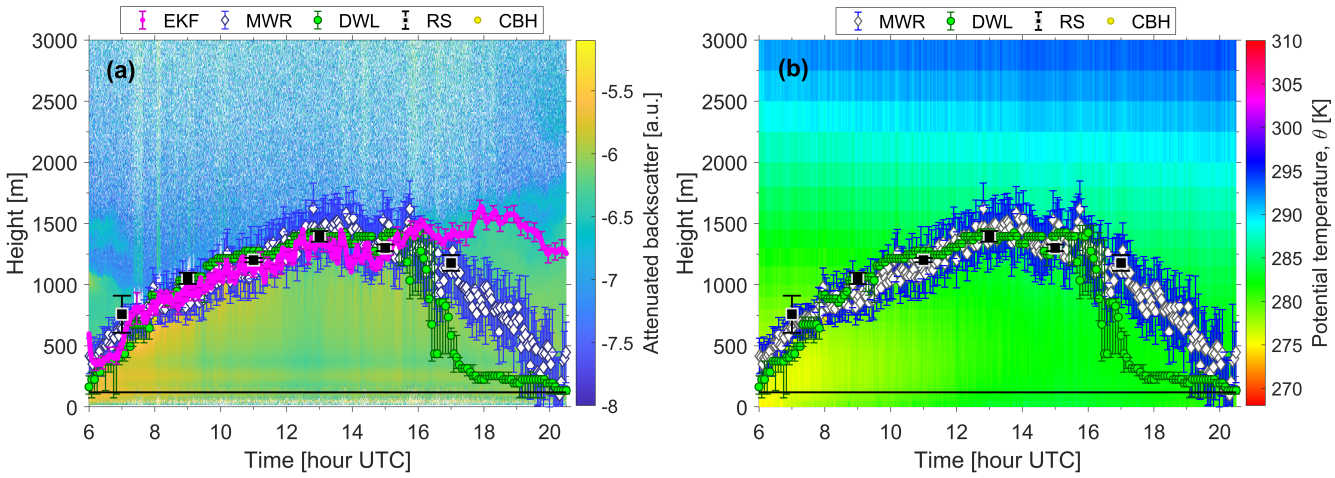

Figure 3: (a) Attenuated ceilometer backscatter (in a.u.) measured by the Jenoptik CHM-15 and (b) MWR-retrieved potential temperature (in K) for 20 April 2013 at Jülich, Germany. In panel (a), magenta dots (along with noise-related error bars) represent MLH-LC-EKF. In both panels, white diamonds represent MLH-MWR with associated uncertainty Δz_{MLH}^{MWR} (blue error bars); black squares are MLH-RS; green dots MLH-DWL; and yellow dots are CBH estimates from the Vaisala CT25. In both panels, the solid black line represents a lower bound (120 m) on MLH-DWL.

Covariance matrices $\boldsymbol{P_k^-}$ and $\boldsymbol{P_k}$ are updated with each successive step of the recursive loop of the filter (Fig. 2b) as a function of the current information available to the filter at time t_k (actual Kalman or projection gain, $\boldsymbol{K}_k$, linearised Jacobian, state vector, $\boldsymbol{x}_k$, state-vector covariance matrix, $\boldsymbol{P}_k$, and measurement-noise covariance matrix, $\boldsymbol{R}_k$) as well as initial settings at t_0. The initial guess of the a priori state-vector error covariance matrix, $\boldsymbol{P}_0^-$, is a rough estimate of the uncertainty associated with the initial guess of the state vector, $\hat{\boldsymbol{x}}_{\boldsymbol{0}}^-$, in the form of a diagonal matrix $\boldsymbol{P_0^-} = diag\left[\boldsymbol{\sigma_P^2}\right], \boldsymbol{\sigma_P} = \left(\sigma_{e,z_{MLH}}, \sigma_{e,a}, \sigma_{e,A}, \sigma_{e,c}\right)$, where $\sigma_{e,X}, X = [z_{MLH}, a, A, c]$ represents the assumed uncertainty of the initial guess, $\hat{\boldsymbol{x}}_{\boldsymbol{0}}^- = \left[z_{MLH,0}, a_0, A_0, c_0\right]^T$ at 1σ level. $\boldsymbol{\sigma_P}$ is shorthand notation for $\boldsymbol{\sigma}_{\boldsymbol{P_0^-}}$. We compute the input vector $\boldsymbol{\sigma_P}$ as

$$\boldsymbol{\sigma_P} = \mu_P \hat{\boldsymbol{x}}_{\boldsymbol{0}}^-, \tag{7}$$

where μ_P denotes the P-intensity factor, to be specified by the user. Factors in the range $\mu_P = 0.1 - 0.3 (10 - 30\%)$ have been used in the examples of Sect. IV, with $\mu_P = 0.3$ the default setting. High/low values $(0.3/0.1)$ tend to increase/decrease the search span of the filter during the first iterations. For example, assuming $z_{MLH,0} = 1500$ m, the setting $\mu_P = 0.3$ tells the filter that the user expects the MLH to be roughly at 1500 ± 450 m at filter start up. $P-$ and $Q-$intensity factors are partially coupled parameters because of the recursive nature of the filter and, therefore, the guidelines above are just orientative. Layer-attribution errors are common because, irrespective of the user's initializations for $\hat{\boldsymbol{x}}_{\boldsymbol{0}}^-$ and $\boldsymbol{\sigma_P}$,

successful filter operation is always conditioned to the existence of aerosol gradients. Therefore, Δz_{attr} can only be known by comparison to a reference instrument (e.g., the radiosonde).

(ii) Impact of observation noise and resulting a posteriori error on the estimated MLH, σ^{P}_{MLH}: Noise-induced errors are due to the presence of noise in $\beta'(z)$ and propagate an error to the MLH-LC-EKF estimate. The recursive loop of the EKF provides by itself convenient error estimates ($\boldsymbol{P}_k^-$, $\boldsymbol{P}_k$ and $\boldsymbol{Q}_k$) of the estimated state vector and, therefore, of the estimated MLH at each discrete time t_k.

Fig. 2b shows the estimated *a priori* and *a posteriori* errors for the case shown in Fig. 3 computed as the time-dependent standard deviations, $\sigma^{P^-}_{MLH,k}$, $\sigma^{P}_{MLH,k}$, respectively. These are compared to the time-static state noise standard deviation, σ^{Q}_{MLH}. Standard deviations $\sigma^{P^-}_{MLH,k}$, $\sigma^{P}_{MLH,k}$ and $\sigma^{Q}_{MLH,k}$ are computed as the square root of the first diagonal element of these matrices during the recursive loop of the filter. The error of the instantaneous MLH-LC-EKF estimate at time t_k is given by the *a posteriori* error as

$$\sigma^{EKF}_{MLH,k} = \sigma^{P}_{MLH,k} \ , \tag{8}$$

which is the key error indicator of interest. In Fig. 2b, it can be seen that the a posteriori error magnitude, $\sigma^{P}_{MLH,k}$, is always smaller than the *a priori* error, $\sigma^{P^-}_{MLH,k}$. This variance reduction, $\left(\sigma^{P}_{MLH,k}\right)^2 < \left(\sigma^{P^-}_{MLH,k}\right)^2$, means that the assimilation of the present measurement $\boldsymbol{y_k}$ counteracts the detrimental effects of observation noise. The latter term merges into a single body both measurement noise $v(z)$ (with $\boldsymbol{R_k}$ being estimated run time) and modelling noise (i.e., approximation of the ML-to-FT transition by the erf model used in MLH-LC-EKF). $\sigma^{Q}_{MLH,k}$ is just a reference baseline of the user's assumed MLH variability. Fig. 2c depicts time evolution of search boundaries z_1, z_1', z_2' and z_2 (Fig. 2a) during the first four hours of data processing [11].

The error estimates above are obviously subordinated to filter convergence and tracking of the ML-to-FT interface (i.e., no attribution errors). Otherwise, the total error from sources (i) and (ii) above can be calculated by using the error superposition principle as (see Fig. 1)

$$\left|\Delta z^{CEILO}_{MLH}\right| = \sqrt{\left|\Delta z_{attr}\right|^2 + \left|\sigma^{P}_{MLH}\right|^2}. \tag{9}$$

B. MLH estimation from DWL data

To estimate the MLH from DWL measurements, the standard deviation of time-height profiles of vertical velocity are calculated every 5 min within a ± 15 min window. Second, the standard deviation is corrected for instrument noise [51]. Finally, the MLH is estimated as the first height at which the Vertical Velocity STandard Deviation (VVSTD) falls below a predetermined threshold [27], [52], [53] (hereafter, MLH-DWL).

Useable thresholds for VVSTD range from 0.2 m s^{-1} to 0.4 m s^{-1} [52], [53]. Based on the work of [27], who studied the sensitivity of the mixing layer height derived from different thresholds, here we use

a 0.4 m s^{-1} threshold. Because 0.4 m s^{-1} is at the high end of the accepted range, it represents a major source of uncertainty.

DWL-based estimates of ML height (MLH-DWL) are generally limited to daytime hours, when the boundary layer is turbulent due to solar heating and thermally-driven ML overturn. MLH-DWLs at night are less reliable because aerosol content is reduced in stable situations at night. Additionally, mixing layer height values below 120 m AGL are rejected as the DWL is not sufficiently sensitive below this height.

The uncertainty associated with the DWL-derived MLH is estimated by applying a 25% variation to the nominal threshold ($th_{\pm} = 0.4$ m·s^{-1} ± 0.1 m·s^{-1}) [27]. Upper and lower uncertainty bounds are derived from the MLH estimations using 0.5- and 0.3-m·s^{-1} thresholds, respectively, minus the MLH estimation using the nominal 0.4 m·s^{-1} threshold. Formally,

$$\sigma_{MLH,k}^{DWL,\pm} = \left| z_{MLH,k}^{DWL\ (th=0.4\pm0.1)} - z_{MLH,k}^{DWL\ (th=0.4)} \right|, \tag{10}$$

where the plus *(minus)* superscript in $\sigma_{MLH,k}^{DWL,\pm}$ denotes the upper *(lower)* uncertainty bound, which is computed by inserting threshold $th = 0.4 + 0.1$ m·s^{-1} *($th = 0.4 - 0.1$ m·s^{-1})* in the first term of the right hand side of Eq. (10) above. A similar approach was previously employed by Villalonga et al. [54].

C. MLH estimation from MWR data

MLH estimation method: The parcel method is commonly used for MLH estimation using potential temperature data [3], [7]. For a given profile of physical temperature, $T(z)$, retrieved from brightness measurements, the first step is to convert it to potential temperature profile, $\theta(z)$, by using

$$\theta(z) = T(z) \left(\frac{p_0}{p(z)} \right)^{\frac{R}{C_p}}, \tag{11}$$

where p_0 is the surface atmospheric pressure, $p(z)$ is the atmospheric pressure profile, $R = 287$ J·K^{-1}·kg^{-1} is the universal gas constant, and $C_p = 1004$ J · K^{-1} · kg^{-1} is the specific heat capacity for dry air at a constant pressure [55]. Physically, the potential temperature represents the temperature an air parcel at an altitude z would have if it were lowered, dry adiabatically, to the surface. In the parcel method, the MLH is defined as the lowest point in a given potential temperature profile for which $\theta(z) > \theta(0)$, where $\theta(0)$ is the surface value of the potential temperature. Small scale effects (e.g., surface properties and shielding of the sensor) can bias the estimate of surface temperature $T(0) = \theta(0)$, to which the parcel method is very sensitive [6], [56].

Changes in $\theta(z)$ with respect to height, $\frac{d\theta}{dz}$, are indicative of the stability of the atmosphere with respect to displacement of unsaturated air parcels. The atmosphere is stable when $\frac{d\theta}{dz} > 0$, neutral when $\frac{d\theta}{dz} = 0$, and unstable when $\frac{d\theta}{dz} < 0$. Under quiescent conditions in daytime, the mixed layer (ML) is characterized by continuous convective mixing, driven from below by buoyant thermal plumes from the relatively warm

surface and, sometimes, from above, by evaporatively driven downdrafts initiated within clouds in the EZ. As a result of this continuous mixing, the interior of the ML exhibits nearly uniform temperature and moisture throughout most of its depth. Thus, $\frac{d\theta}{dz} = 0$ in the ML, and negative in the surface layer (unstable), $\frac{d\theta}{dz} < 0$. At the top of the ML, an increase in temperature and reduction in moisture delineates the EZ, i.e., a transition layer between the ML and the FA. The MLH is typically computed as falling halfway between the top of the ML and the bottom of the FA, near where the magnitudes of the temperature and/or moisture gradients are maximized.

While the parcel method was designed for use with sounding data, Stull [1] (p. 474) cautions against estimating MLH using only a single radiosounding. This is because a single radiosounding may not be representative of average conditions in a horizontally heterogeneous ML, as would be the case in a convectively active BL characterized by thermal updrafts and downdrafts. MWR offers a partial solution to this issue in that it provides a time series of potential temperature profiles. Temporal averaging can ameliorate to some extent the perturbations caused by individual updrafts and downdrafts, allowing longer-term (e.g., $\sim$O(30 min–1 hr)) trends in MLH to be discerned.

To estimate the MLH from MWR data, first, MWR potential-temperature profiles are interpolated to a uniform vertical resolution of 10 m. Second, a five-point (50 m) moving average is applied to smooth the profiles [27]. Then, the MLH-MWR is derived by using as surface temperature reference, $\theta(0)$, that from the JOYCE meteorological tower at 2 m, thus assuring reliable retrievals [6]. The tower-derived temperature, which has a raw temporal resolution of 6 min, is interpolated to the MWR temporal resolution (2.7 min).

In Fig. 3, we present an example comparison of MLH-LC-EKF (Fig. 3a) with MLH-MWR (Fig. 3b). In spite of the relatively coarse spatial resolution (which decreases with height) of the MWR potential temperature profiles, it can be seen that MLH-MWR follows the overall trend of the radiosonde-derived MLH, and particularly during morning (0600-1000 UTC) and evening (1600-1900 UTC) transition times MTT and ETT, respectively. The MLH-MWR captures the collapse of the convective BL during the latter period, after sunset, while the MLH-LC-EKF estimate continues to follow the elevated RL.

Error estimation: The uncertainty associated with MLH-MWR has been approximated by two main error sources (Fig. 1): (a) the total uncertainty in the retrieved temperature profile $T(z)$ and consequent error propagated to the MLH calculation, which gives rise to a MLH error, Δz_T, and (b) the uncertainty in the estimated surface temperature, T_0 (or, equivalently, $\theta(0)$), which gives rise to a MLH error, Δz_{T_0}.

(a) Assessment of the MWR-retrieved temperature-profile error in the estimated MLH, Δz_T: A study on the performance of the retrieved temperature profile [35] using a long-term data set of representative atmospheric profiles and noise levels found uncertainty in the 0.1–0.5 K range. A synthetic brightness

temperature data set generated from over 10 000 radiosoundings (of which 5334 were used for training and 4954 were used as a validation data set) was used to test the performance of the statistical retrieval algorithm (analogous to multi-regression algorithm [57]). Altitude-dependent temperature uncertainties, $\Delta T(z)$, were generated, varying from 0.44 K on the ground to 1.60 K at 4 km. An example of the height-dependent temperature-retrieval kernels is given in Fig. 7 of [5].

In order to assess the uncertainty of MLH-MWR due to temperature-retrieval errors, Δz_T, we adopt the following approach: At each time instant, the retrieved height-dependent temperature errors, $\Delta T(z)$ are converted into potential-temperature errors, $\Delta\theta(z)$. These are added to and subtracted from the retrieved potential temperature profile, resulting in delineation of its "upper" and "lower" error bounds. In this approach, the uncertainty in the temperature profile acts like a bias, i.e., a consistent under- or over-estimation throughout the profile (see also [58]). The parcel method is then applied to all three of these profiles (Fig. 4a): (i) $\theta_{MWR}(z) + \Delta\theta(z)$, or the upper error-bound profile for the potential temperature, (ii) $\theta_{MWR}(z)$, or the nominal profile, and (iii) $\theta_{MWR}(z) - \Delta\theta(z)$, or the lower error-bound profile, and consequently, a MLH error bar is obtained. Fig. 4a shows an example error bar calculation for the temperature profiles retrieved from HATPRO MWR measurements at Jülich, Germany at 1401 UTC on 24 April 2013. It can be observed that retrieval errors on the order of less than 2 K throughout the vertical profile introduce an uncertainty of about $\Delta z_T \approx 500$ m in the MLH estimates.

(b) Assessment of surface-temperature errors on the estimated MLH, Δz_{T_0}: Following a similar perturbation approach, the uncertainty of the estimated MLH due to surface-temperature errors Δz_{T_0} is calculated by adding and subtracting the approximate uncertainty in the surface temperature (± 0.5 K) [35] to T_0. Fig. 4b shows three resulting MLH estimations. As a result of this perturbation in the surface temperature T_0, the uncertainty in the MLH is about 150–300 m.

Finally, the total error from error sources (a) and (b) above is computed by error superposition (Fig. 1) as

$$\left|\Delta z_{MLH}^{MWR}\right| = \sqrt{|\Delta z_T|^2 + |\Delta z_{T_0}|^2} \tag{12}$$

D. MLH estimation from radiosonde data

The procedure to estimate the nominal MLH from RS data is the parcel method (refer to Sect. III-C). RS-MLH errorbars are derived in similar fashion to Sect. III-C.b assuming $\pm$ 0.5 K surface temperature uncertainty. Unlike the uncertainty for temperature profiles $T(z)$ retrieved from MWR, however, the uncertainty associated with RS-measured $T(z)$ is not altitude dependent, but constant ($\pm$ 0.5 K) with height. Thus, $\left|\Delta z_{MLH}^{RS}\right| = \pm$ 150–300 m.

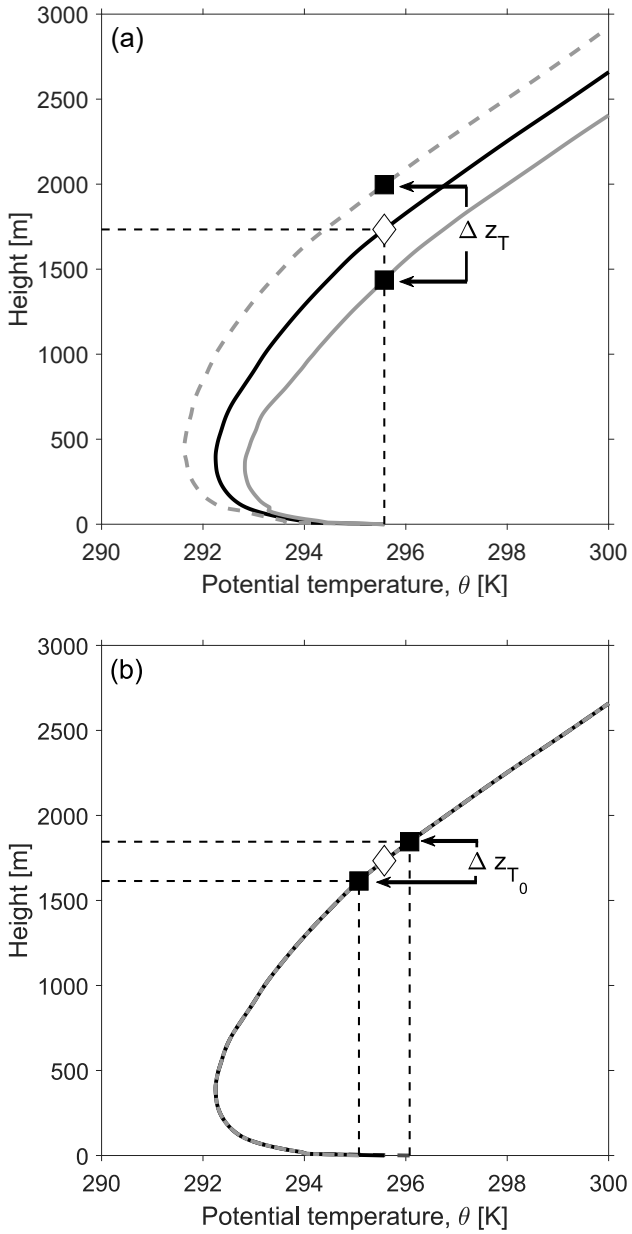

Figure 4: Overview of MLH-MWR (parcel method) estimation errors (24 April 2013, 1401 UTC, Jülich, Germany). (a) MLH-MWR error arising from the uncertainty inherent in the MWR retrieval of the temperature profile, Δz_T. Upper (dashed grey)- and lower (solid grey)-bound profiles are obtained by adding and subtracting the height-dependent temperature error-perturbation profile, $\Delta\theta(z)$, to the nominal potential-temperature profile, $\theta(z)$ (solid black line). The white diamond represents MLH-MWR obtained from the nominal profile, and the black squares are corresponding MLH-MWR obtained from the perturbed profiles. (b) MLH-MWR error due to the uncertainty in the measured surface temperature, Δz_{T_0}, obtained by adding and subtracting 0.5 K from the reference temperature.

E. 30-min averaged MLH retrievals and error assessment

In order to inter-compare MLH retrievals in a meaningful statistical sense we standardize MLH-LC-EKF, MLH-MWR, and MLH-DWL to a common temporal resolution of 30 min via Maximum Likelihood (MaxL) as [59]

$$MLH_X(t_h) = \frac{\sum_k z_k/\sigma_k^2}{\sum_k 1/\sigma_k^2}, \tag{13}$$

where X stands for the instrument or method used ($X = EKF$, MWR, DWL), t_h is the center time of the 30-min time window ($t_h = 0630, 0700, \ldots, 2000$ UTC), z_k (formally, $z_k = z^X_{MLH,k}$) is the instantaneous MLH estimated at time $t_k \in [t_h - 15$ min$, \ldots, t_h + 15$ min$]$, i.e., with the raw temporal resolution of the instrument/method X, and σ_k (formally, $\sigma_k = \sigma^X_{MLH,k}$) is the associated uncertainty of MLH estimate z_k. σ_k is computed via Eqs. (8), (12) and (10) for the MLH-LC-EKF, MLH-MWR, and MLH-DWL, respectively.

The associated 30-min MLH uncertainty is computed as

$$\sigma_X(t_h) = \sqrt{\sigma^2_{X_1}(t_h) + \sigma^2_{X_2}(t_h)}, \tag{14}$$

where σ_{X_1} is the standard deviation of the estimated MLH by instrument or method X (a proxy of the instantaneous MLH variability),

$$\sigma_{X_1} = std(z_k) \tag{15}$$

and σ_{X_2} is the uncertainty associated with the MLH estimation by instrument or method X,

$$\sigma_{X_2}(t_h) = \sqrt{\frac{1}{\sum_k 1/\sigma_k^2}}. \tag{16}$$

In Sect. IV, intercomparisons of 30-min MLH will be performed on pairs of instruments. Bias between instruments X and Y is computed as

$$bias^{X,Y}(t_h) = MLH_X(t_h) - MLH_Y(t_h), \tag{17}$$

and bias variability as

$$\sigma^{X,Y}_{bias}(t_h) = \sqrt{\sigma_X(t_h)^2 + \sigma_Y(t_h)^2}. \tag{18}$$

F. The ceilometer-MWR synergistic (SYN) method

The MLH-LC-EKF and MLH-MWR methods feature contrasting behaviours as exemplified in Fig. 3a. On one hand, MLH-MWR uncertainties span several hundred meters (blue error bars), and expand with height as a consequence of the MWR coarser spatial resolution at higher altitudes. On the other hand, MLH-LC-EKF have much smaller uncertainties (magenta error bars), on the order of tens of meters, which

lie within those of MLH-MWR when the ABL is well developed (e.g., 1000–1400 UTC). However, during the ETT (1500–1800 UTC), MLH-LC-EKF detaches from the ABL, following the RL, instead.

Accordingly, we are motivated to introduce a synergistic MLH estimation method (SYN), which yields an optimal MLH estimate (MLH-SYN) that improves upon MLH-MWR and MLH-LC-EKF estimates considered in isolation. The SYN method combines the 30-min MLH retrievals and associated uncertainties of these two methods as follows: the MLH estimate provided by the synergistic method, MLH_{SYN}, is computed as

(i) the maximum likelihood between MLH-LC-EKF and MLH-MWR as

$$MLH_{SYN}(t_h) = \frac{\frac{MLH_{EKF}(t_h)}{\sigma^2_{EKF}(t_h)} + \frac{MLH_{MWR}(t_h)}{\sigma^2_{MWR}(t_h)}}{1/\sigma^2_{EKF}(t_h) + 1/\sigma^2_{MWR}(t_h)} \quad if\ I_{EKF}(t_h) \cap I_{MWR}(t_h) \neq \emptyset \tag{19}$$

in two situations:

a) when their respective MLH uncertainty intervals (Eq. (14)), $I_{EKF}(t_h) = [\![MLH_{EKF}(t_h) - \sigma_{EKF}(t_h), MLH_{EKF}(t_h) + \sigma_{EKF}(t_h)]\!]$ and $I_{MWR}(t_h) = [\![MLH_{MWR}(t_h) - \sigma_{MWR}(t_h), MLH_{MWR}(t_h) + \sigma_{MWR}(t_h)]\!]$ overlap at least partially (i.e., Eq. (19)), or

b) when t_h falls in a strongly convective period, I = [1000 – 1400] UTC. I corresponds to mid-afternoon at the JOYCE site, when, assuming quiescent atmospheric conditions, the CBL reaches is maximum depth and maturity. In a more generalized formulation, I would be location-dependent.

The associated uncertainty for the synergistic MaxL estimate of Eq. (19) above is given by

$$\sigma_{SYN}(t_h) = \sqrt{\frac{1}{1/\sigma^2_{EKF}(t_h) + 1/\sigma^2_{MWR}(t_h)}}. \tag{20}$$

(ii) The MLH-MWR estimate elsewhere (i.e., out of statements (i.a) and (i.b) above). Formally,

$$\begin{cases} MLH_{SYN}(t_h) = MLH_{MWR}(t_h) \\ \sigma_{SYN}(t_h) = \sigma_{MWR}(t_h) \end{cases}. \tag{21}$$

In this latter case, the uncertainty of the MLH-SYN is equal to that of MLH-MWR (Eq. (14)).

Eq. (19) is essentially the MaxL definition given by Eq. (13) but applied to each pair of MLH estimates, EKF and MWR, at each 30-min time step, t_h. The SYN method is discussed next in Sect. IV-A.

In case (i), this formulation balances MLH-LC-EKF and MLH-MWR estimates by attributing higher weight to the estimates with lower uncertainty. Typically, MLH-LC-EKF is favored during the mid-afternoon peak in convective boundary layer growth. Case (ii) typifies MLH development or decay during MTT and ETT, respectively. During these periods, MLH-LC-EKF and MLH-MWR tend to diverge, and the SYN method retains MLH-MWR as the most reliable estimate. This constraint ensures that the SYN method avoids MLH-LC-EKF tracking the RL (i.e., layer-attribution error).

IV. DISCUSSION

30-min MLH-SYN estimates are compared to MLH-LC-EKF, MLH-MWR and MLH-DWL (Sect. III) considered in isolation with reference to MLH-RS. The statistical analysis is limited to the 0600-2030 UTC time interval, which delineates the CBL diurnal cycle including MTT and ETT over the JOYCE site. Sect. IV-A discusses the synergistic method in the context of one "textbook" clear-day example, Sect. IV-B gives an overview of the campaign dataset, and Sect IV-C evaluates performance statistics of the different MLH retrieval methods for the whole campaign.

A. Synergistic method example

Case day 20 April 2013 (Fig. 3a) is used to exemplify the synergistic method retrieval in Fig. 5a. MLH estimates derived from the different methods are plotted with 30-min resolution (Eq. 13). This day was characterized by a cloud-capped mixing layer from 0700 to 1100 UTC and by clear sky otherwise. For visual reference in the plots next, MTT is defined as [0600–1000) UTC, ETT as [1400–2030) UTC, and peak convective boundary layer growth time as [1000–1400) UTC. Solar noon at JOYCE is 1130 UTC.

From 0600 to 0800 UTC, MLH-LC-EKF, MLH-MWR and MLH-DWL track closely with one another but fall below MLH-RS (Fig. 5a). The MLH-LC-EKF agreed well with MLH-MWR, despite the presence of aerosols in the 500–1000 m layer. In contrast, MLH-DWL fell slightly below MLH-MWR because of the relatively weak turbulence in the early morning, and the the use of a constant VVSTD threshold (Sect. III-B). MLH estimates from all methods coincided from 0800 to 1500 UTC. In other words, during most of this interval (case (i.a) in Sect. III-F), the MLH-LC-EKF and MLH-MWR errorbars (computed as the $\pm 3\sigma$ value from Eq. (14)) partially or totally overlapped and hence, the MLH-SYN) was the MaxL estimate between the MLH-LC-EKF and MLH-MWR retrievals in isolation (Eq. 19) with an associated uncertainty given by Eq. (20).

During the strongly convective interval, 1000 to 1400 UTC (case (i.b) in Sect. III-F), the ML-to-FT gradient was sharply defined, and reliability of MLH-LC-EKF increased, with values closer to MLH-RS. In this interval, the ML was free from layer attribution errors, and the MLH-SYN was constrained to the MaxL value between MLH-LC-EKF and MLH-MWR, irrespective of whether their individual error bars overlapped (e.g., at 1400 UTC). The latter shows the case of thermal updrafts causing MLH-MWR and MLH-DWL to jump slightly above the smoother MLH-LC-EKF time series. As a result of the smaller MLH-LC-EKF error bars during this interval, the MLH-SYN followed MLH-LC-EKF and inherited its smaller uncertainty (Eq. 20).

After 1500 UTC (i.e., the start of the ETT), each MLH estimate started to develop a distinct behaviour: MLH-DWL fell quickly as thermal turbulence decayed, whereas MLH-MWR decreased smoothly, thus

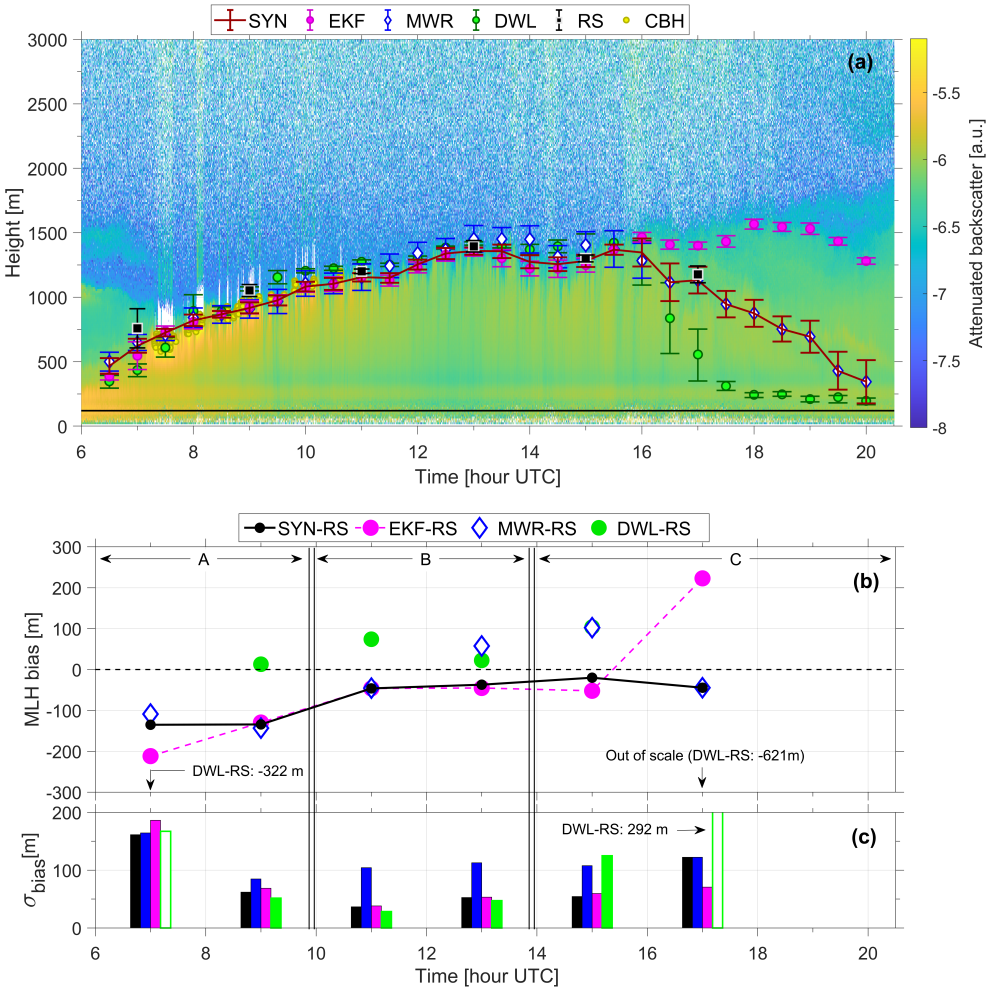

Figure 5: Performance of the SYN algorithm and MLH-LC-EKF, MLH-MWR, and MLH-DWL methods in isolation with reference to MLH-RS estimates as function of hour of day (case 20 April 2013, Fig. 3). (a) 30-min MLH estimates (Eq. (13)). (b) MLH bias (Eq. (17)). (c) MLH bias variability (Eq. (18)). Labels A, B, and C indicate typical MTT, peak convective, and ETT intervals, respectively (Sect. IV-A).

coinciding with the radiosonde at 1700 UTC. In contrast, MLH-LC-EKF kept on tracking the elevated aerosol layer, which became the RL. In this situation, MLH-SYN followed MLH-MWR (case (ii) in Sect. III-F, with MLH-SYN error bars overlapping MLH-MWR error bars (plotted in dark red and blue, respectively, in Fig. 5a), therefore avoiding layer-attribution errors typical of MLH-LC-EKF in the ETT.

Bias performance of the different MLH estimation methods with reference to MLH-RS are shown in Fig 5b. MLH-SYN and MLH-MWR (labeled SYN-RS and MWR-RS, respectively) yielded bias (Eq. 17) of less than $\pm$150 m during the whole period (0600–1400 UTC). Additionally, MLH-SYN bias variability (Eq. 18) was always lower than that of MLH-MWR (Fig 5c, SYN-RS and MWR-RS bars), which increases our confidence in the SYN algorithm. In the highly convective time interval *B*, it is evident that the MLH-LC-EKF bias variability is much lower than that of MLH-MWR (Fig 5c, EKF-RS and MWR-RS

bars). MLH-DWL bias in the 0900–1500 UTC interval (corresponding to the well-developed ML) was $\pm$ $\sim$100 m, comparable to MLH-SYN bias. However, outside of this interval (e.g., at 0700 and 1700 UTC, empty green circles) MLH-DWL bias dramatically increased to 300 and 600 m, respectively. As mentioned in Sect. III-B, this is a consequence of using a constant VVSTD threshold.

B. Dataset overview

Having examined a single case day in detail, we now expand our analysis to consider the entire data set from the 31-day HOPE campaign (01 April to 06 June 2013), which included 80 radiosondes.

Twenty-one days were selected from this set with a total of 55 radiosondes available. Selected days were either clear-sky days (cloud cover below 3 km lower than 10%, 8 days) or cloudy days (cloud cover below 3 km not greater than 70%, 13 days): Clear-sky days included days with single (Fig. 5a) or multiple aerosol layers (Fig. 6a) in the transition times. Cloudy days were characterized by a cloud-capped BL (Fig. 6b), sometimes with additional midlevel clouds well above the BL (Fig. 6c). Eight of the selected days included light drizzle events ($<$ 0.5 mm/h, $<$ 30 min/event, accumulated rain (0600-2000 UTC) $<$ 0.1 mm), which usually occurred during MTT or ETT. Excluded days (10) were days with cloud cover below 3 km greater than 70% (7 days) and rainy days (rain intensity $>=$ 0.5 mm/h, 2 days).

During the period of peak CBL growth (1000–1430 UTC), CBHs are usually at the same height as the MLH (Fig. 6b), as is common in the spring and summer time [27]. MLH-RS using the parcel method overestimated the MLH at 11, 13 and 15 UTC by some 300 m, which motivated us to compare MLH-RS with the MLH retrieved using another thermodynamically-based alternative, the Bulk Richardson Number method [27]. Like the parcel method, the Bulk Richardson Number method is also based on the temperature profile, but less sensitive to perturbations in the surface temperature, T_0. For consistency when computing statistics for the whole campaign, we retain MLH-RS using the parcel method.

Virga (precipitation streamers attached to the base of the clouds, e.g., in Fig. 6c from 1730-2000 UTC) were also problematic for MLH-LC-EKF, which tended to track the sharp gradient at the cloud tops from 1800 UTC onwards. Because cloud cover strongly changes the incoming solar radiation and, consequently, ML growth, cloudy days are particularly challenging for MLH-DWL, which uses VVSTD as a proxy of the turbulent mixing. Qualitatively, the SYN algorithm delineated fairly well the typical ML diurnal cycle (Fig. 6a-c).

C. Performance statistics

In this subsection, statistical measures of central tendency and variability for the 21-day sample considered are used to demonstrate superiority of MLH-SYN over MLH-MWR and MLH-LC-EKF estimates considered in isolation. Definitions for the statistical indicators are given in Appendix A.

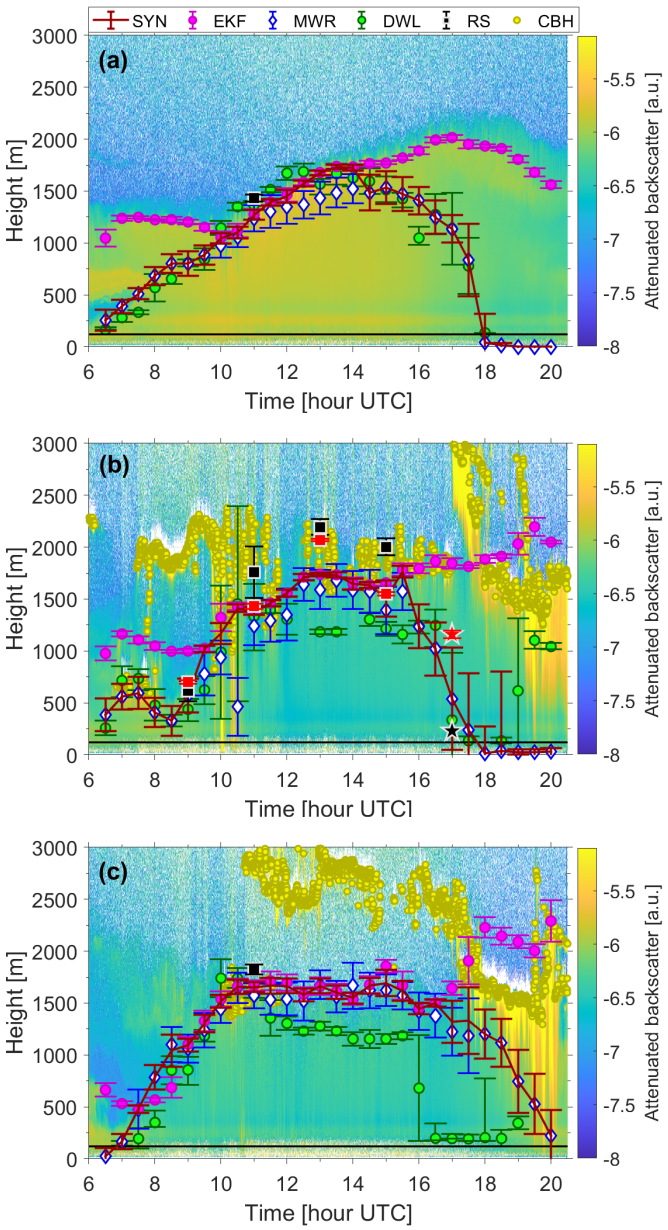

Figure 6: As in Fig. 5a, but representing the three typical atmospheric modes observed during the HOPE campaign: (a) Clear-sky day (22 April 2013) showing multiple aerosol layers from 0600 to 1000 UTC; (b) partially cloud-capped boundary layer day (13 April 2013); and (c) midlevel cloud day (16 May 2013) with virga from 1730 to 2000 UTC. In panel (b), red squares are MLH estimates made using Bulk Richardson Number (BRN) derived from RS $\theta(z)$ profiles. Black and red stars at 1700 UTC indicate the stable BL height estimates retrieved by the gradient method and BRN, respectively.

During the course of this research, it was found that MLH-DWL performance statistics for the whole campaign are inferior to those of the other MLH estimation methods, mainly because MLH-DWL is hampered by the use of a fixed VVSTD threshold. In particular, during MTT and ETT, intermittent

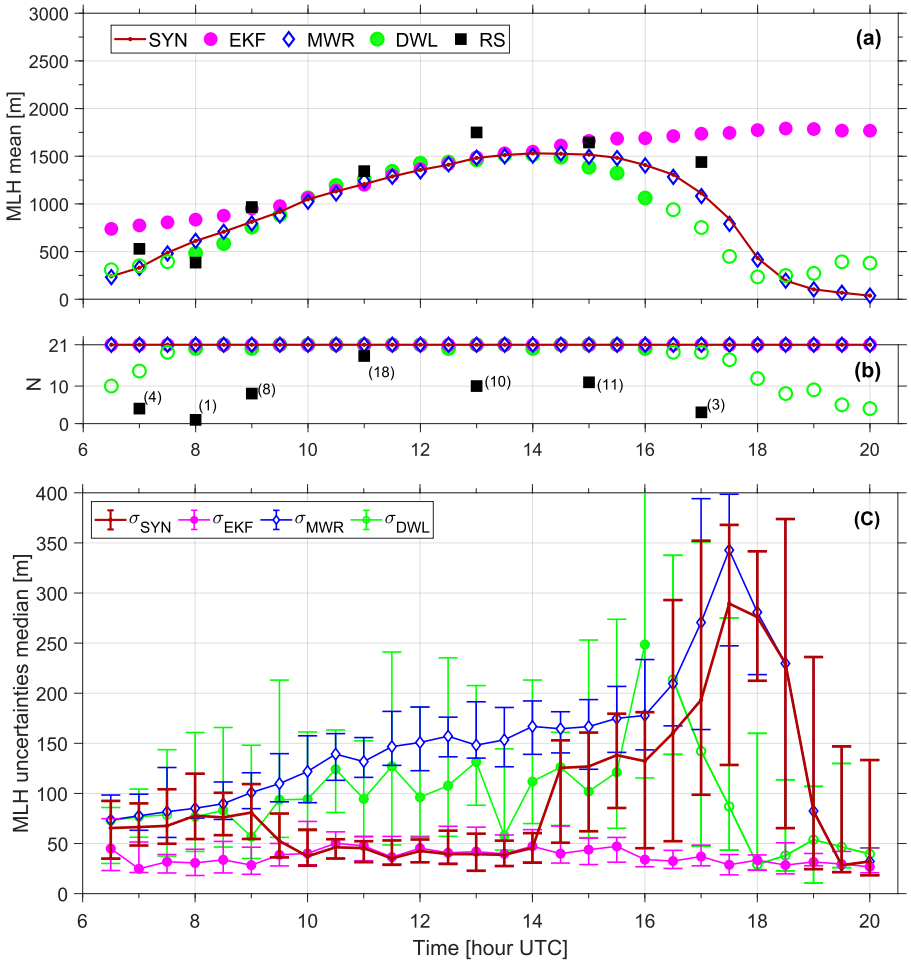

Figure 7: Means and uncertainties of 30-min MLH estimates over the selected 21 clear-sky days as a function of hour of day in UTC (LT=UTC+1h). (a) (Top panel) means (Eq. (A.1)) of MLH-SYN (red solid trace), MLH-LC-EKF (magenta dots), MLH-MWR (white diamonds), MLH-DWL (green dots) and MLH-RS (black squares). (b) Number of case days (out of 21 total days) used to compute mean values. (c) Median values (markers) and interquartile ranges (bars) of the 30-min MLH uncertainties (Eq. 14).

turbulence tends to cause MLH-DWL instability if the threshold is not adjusted. Dynamic adjustment of the VVSTD threshold in MLH-DWL falls beyond the scope of the present work. MLH-DWL performance characteristics will nonetheless be reported in this study in order to motivate future research in this area.

I. MLH estimates by the different methods: MLH estimates averaged over the selected 21 days as a function of hour of day (local time) are shown in Fig. 7a. While both MLH-MWR and MLH-EKF methods were based on 21 samples (one for each hour, each day; Fig. 7b), MLH-DWL furnished fewer samples during MTT (0600-0800 UTC) and ETT (1600-2030 UTC) , open green circles). In these time intervals, MLH-DWL was often < 120 m, and such estimates were rejected as outliers by the reasoning discussed in Sect. III-B.

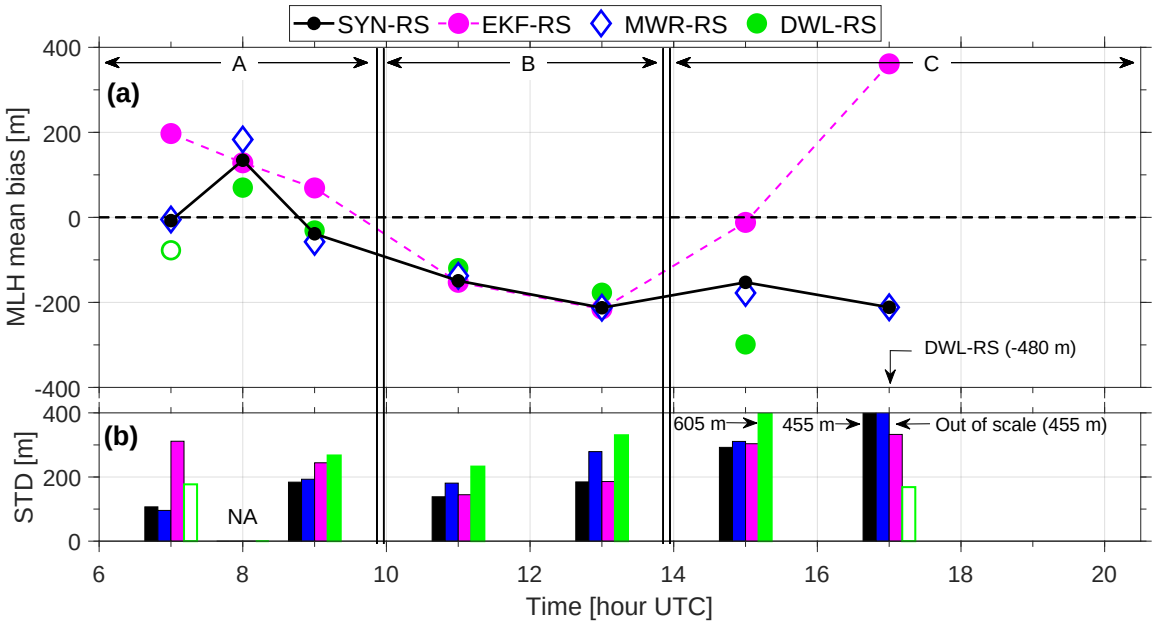

Figure 8: Biases relative to MLH-RS of the 30-min MLH estimations over the selected 21 clear-sky days as a function of hour of day in UTC. (a) Mean biases. Label "X-RS" (X=SYN, EKF, MWR, DWL) stands for the mean bias $\mu_{bias}^{X,RS}(t_h)$ (Eq. (A.3)) between an instrument/method combination X and MLH-RS. (b) Mean biases standard deviation $\sigma_{bias}^{X,RS}(t_h)$ (Eq. (A.4))

The best agreement among all MLH estimation methods occurred from 1000 to 1400 UTC. As previously mentioned, this is the interval in which, under quiescent conditions, intensive, thermally driven convection deepens the CBL and, consequently, MLH reaches its highest altitudes of the daily cycle [1]. The MLH-LC-EKF attribution-layer error clearly evidenced during MTTs and ETTs, during which the SYN algorithm successfully took over. From 0600 to 0900 UTC, the MLH-LC-EKF exceeded MLH from other estimators, revealing that MLH-LC-EKF retrievals were frequently affected by the persistent aerosol layers from the previous night. During early the morning, the MLH-SYN followed MLH-MWR. Both rose in concert with MLH-RS 0700 and at 0900 UTC. MLH-SYN, MLH-MWR and MLH-DWL rose together until 1400 UTC. As exemplified by previously discussed case of 20 April 2013 (Fig. 5), the disagreement among all MLH estimation methods grew sharply from 1500 UTC onwards (i.e., the start of the ETT), when the turbulence decayed, the CBL separated from the RL, and the stable boundary layer began to form in response to the reversal of surface radiative flux. MLH-LC-EKF tracked this RL. While MLH-DWL properly tracked the turbulence decay, it exhibited worse agreement than MLH-MWR with MLH-RS at 1700 UTC (Fig. 7a).

30-min MLH estimation uncertainty (Eq. (14)), as a function of hour of day is shown in Fig. 7c by computing medians and interquartile ranges (25-to-75-th percentiles) for the total sample of 21 days. MLH-LC-EKF exhibited much lower medians (by $\approx$ 40 m) and spreads than MLH-MWR in all hours

of the strongly convective period (e.g., $\approx$ 140 m, 1000–1400 UTC). This result further favours the SYN algorithm during this interval.

II. Performance of the SYN method and of MWR and EKF methods in isolation with reference to RS retrievals:

MLH bias vs. RS - MLH mean bias (Fig. 8a, Eq. (A.3)) is computed as the difference between the 21-day means of MLH-SYN, MLH-MWR, MLH-DWL and MLH-RS estimates in a 30-min time window centred at the RS launch time.

MLH-SYN and MLH-MWR yielded mean biases less than $\pm$ 150 m overall (SYN-RS and MWR-RS labels in Fig. 8a, respectively), and -150 m during the convective interval (*B*; 1000–1400 UTC). The latter is consistent with the findings of [18] (Fig. 4a therein) who estimated bias of $\approx$ 200 m for MLH-LC-EKF. As expected, MLH-LC-EKF (EKF-RS label in Fig. 8a) performed poorly outside the 1000–1400 UTC convective interval (*B*), during the MTT and ETT. Its V-shaped mean bias curve reaches +200 and +400 m at 0700 and 1700 UTC, respectively.

MLH bias-to-RS variability.- MLH bias-to-RS variability (Fig. 8b) is computed as the standard deviation (STD) given by Eq. (A.4). Because all the MLH estimates are subject to the same atmospheric, day-to-day variability, comparative differences in bias STD among them are attributable to their individual performance. Therefore, each hourly set of vertical bars ranks the different MLH estimation methods by bias performance with reference to RS. Thus, in the convective interval (1000-1400 UTC) MLH-LC-EKF and MLH-SYN had the lowest bias STD (EKF-RS and SYN-RS labels, respectively). In contrast, during TTs, MLH-MWR and MLH-SYN exhibited the lowest bias STD. In summary, the the SYN algorithm performed best over the whole daily interval because it inherited the superior performance characteristics of its constituent methods. MLH-LC-EKF performance in TTs was – as expected – worst (Fig. 8a). MLH-DWL only gave usable estimates between 900-1500 UTC, and even then its performance (green bars) was always worse than that of the MWR (blue bars). Outside of this interval, there were retrieval issues (empty green dots) related either to the selection of a fixed VVTSD threshold or MLHs below the instrument minimum measurable height of 120 m.

Correlation and regression analysis.- Fig. 9 compares the different MLH retrievals to MLH-RS (e.g., $X = MLH_{RS}$ and $Y = MLH_{MWR}$ in panel (a)). Two time intervals were investigated: *(i) One capturing the full daytime diurnal cycle* [0600-2030) UTC, which included 55 radiosondes, and *(ii) the shorter, strongly convective period in the local afternoon* [0930-1430) UTC, which included 28 radiosondes. MLH mean bias (MB) values in each of these two time intervals were comparable for both SYN and MWR methods. Additionally, the MB obtained in time intervals *(i)* and *(ii)* (red and blue text labels, respectively) for each method approximately coincided with the average of the hourly biases plotted in Fig. 8a in these

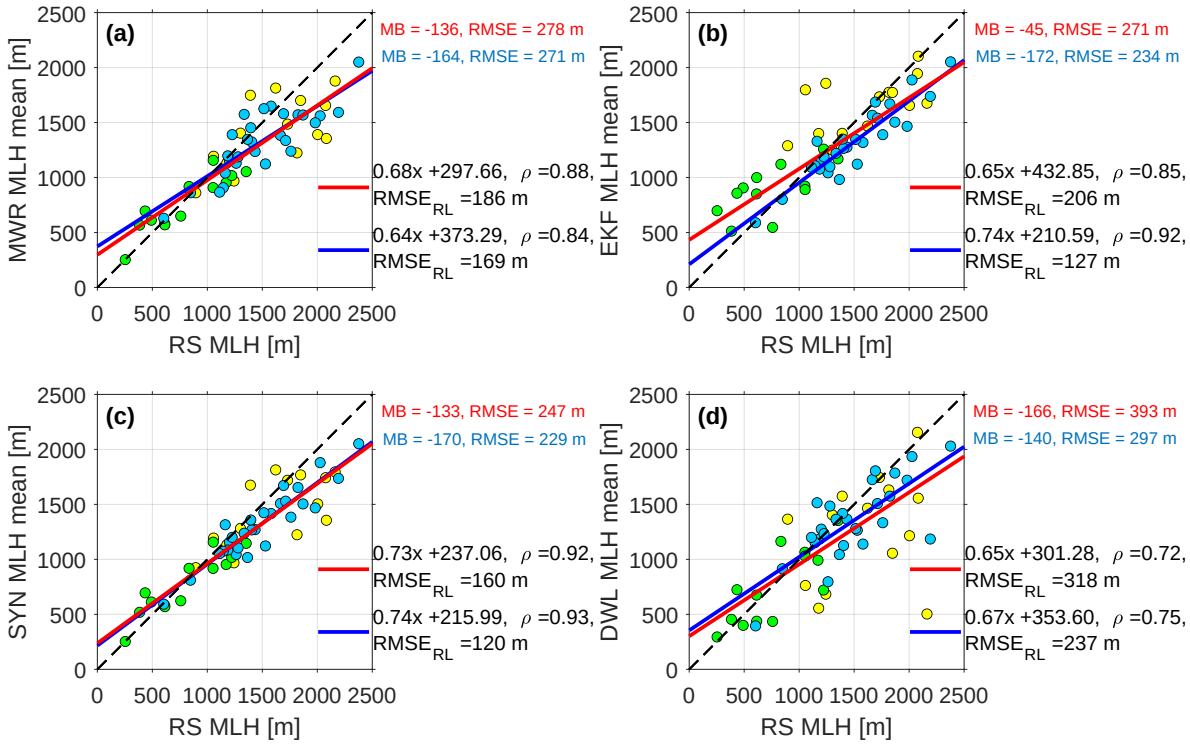

Figure 9: Scatterplot comparison of 21 clear-sky day, 30-min (a) MLH-MWR, (b) MLH-LC-EKF, (c) MLH-SYN and (d) MLH-DWL with MLH-RS. Green, blue and yellow dots depict MLH estimates in the [0600-0930), [0930-1430) and [1430-2030) UTC time intervals, respectively. Red and blue lines are regression lines over diurnal [0600-2030) UTC and convective [0930-1430) UTC MLH estimates, respectively. RMSE_{RL} denotes regression-line root-mean-square error. For reference in all panels, the 1:1 line is drawn as a black dashed line. Mean bias (MB) and RMSE are annotated in colors corresponding to their respective time intervals.

intervals. The root-mean-square error (RMSE) was slightly lower for the SYN method ($RMSE^{(i)}_{SYN} = 247$ m, $RMSE^{(i)}_{MWR} = 278$ m, see labels) and relatively high for the DWL $RMSE^{(i)}_{SYN} = 393$ m. The latter value is in accordance with the RMSE of 359 m (0800-1600 UTC) reported by Schween et al. [27]. All RMSE indicators improved in convective time interval (ii) because of the lower variability of the MLH.

The significance of the indicators above warrant some comments. Gross outliers were particularly abundant during TTs. Gross outliers are defined as biased estimates (MLH_X-MLH_{RS}, X=SYN, EKF, MWR, DWL) above $\pm 1\sigma$ of the mean of the associated MLH bias histogram computed hourly. To further improve the significance of results, gross outliers were removed prior to evaluating correlation statistics via a similar procedure as that described in [18] and [60]. When Fig. 9 was regenerated after gross outliers were removed (not shown), the linear regression lines ("RL" subscript in Fig. 9) became virtually coincident with the 1:1

line, indicating virtually no bias. Furthermore, MLH-SYN remained superior to MLH-MWR and MLH-DWL. In the diurnal time interval *(i)* above, the SYN exhibited $\rho^{(i)}_{SYN,1\sigma} = 0.98$, $RMSE^{(i)}_{SYN,1\sigma} = 76\ m$, followed by the MWR, $\rho^{(i)}_{MWR,1\sigma} = 0.96$ and $RMSE^{(i)}_{MWR,1\sigma} = 103\ m$ and the DWL, $\rho^{(i)}_{DWL,1\sigma} = 0.95$ and $RMSE^{(i)}_{DWL,1\sigma} = 157\ m$. Over the convective time interval *(ii)*, the SYN achieved $\rho^{(ii)}_{SYN,1\sigma} = 0.99$ and $RMSE^{(ii)}_{SYN,1\sigma} = 41\ m$ (as compared to $\rho^{(ii)}_{MWR,1\sigma} = 0.94$ and $RMSE^{(ii)}_{MWR,1\sigma} = 86\ m$ prior to gross outlier removal).

V. CONCLUSIONS

A synergistic MLH retrieval algorithm combining MWR and ceilometer-based estimates was presented along with performance statistics covering 21 days of the HOPE campaign. The SYN method used a maximum-likelihood algorithm (Sect. III-F) that combined MLH-LC-EKF and related error uncertainties in the strongly convective BL time interval (local afternoon, which at JOYCE is 1000–1400 UTC) with MLH-MWR and associated uncertainties outside of this interval. MLH-LC-EKF were derived from the ML-to-FT gradient in the attenuated backscatter profile by using a Kalman filter, yielding time-adaptive MLH estimates with a temporal resolution equal to that of the LC. MLH-MWR were estimated by the parcel method applied to MWR-derived temperature profiles.

The motivation for creating MLH-SYN, which combined MLH estimates by different methods, was the inherent weaknesses each method exhibited in isolation. Specifically, MLH-LC-EKF tends to track aerosol gradients, which led the filter to follow RLs during MTT and ETT. The latter was the most common layer-attribution error. In experiments, the EKF was able to detect gradients as low as 1.5 to 1 times the mean FT level (Fig. 2a). Assuming no attribution errors, the MLH uncertainty was given by the *a posteriori* error of the filter (Sect. III-A2). We therefore formulated MLH-SYN to equal MLH-LC-EKF in the strongly convective afternoon interval (1000-1400 UTC). Outside this interval, MLH-MWR proved more reliable, with an uncertainty given by the inherent error in the MWR-retrieved temperature profile and parcel-method surface-temperature error (Sect. III-C). MLH-DWL was excluded from the formulation of MLH-SYN; intermittent turbulence, particularly during the ETT, along with the assumption of a fixed VVSTD threshold, were major issues.

The variability in MLH estimation uncertainty matched that of its component algorithms (MLH-MWR and MLH-LC-EKF) as a function of time of day (Fig. 7b). Specifically, during the strongly convective interval (1000–1400 UTC), the MLH-SYN exhibited $\approx$ 40 m (median) as compared to $\approx$ 140 m by MLH-MWR, which indicates that SYN inherited the steady tracking performance of the MLH-LC-EKF. Outside this interval, the variability of MLH-SYN approached that of MLH-MWR.

Comparative performance of the different MLH estimation methods with reference to MLH-RS was also addressed. Statistical analysis over the 21-day sample showed that MLH mean bias was inflated

by gross outliers associated with TTs and day-to-day atmospheric variability. Additionally, the parcel method (used in both MLH-MWR and MLH-RS) was particularly sensitive to the accuracy of surface temperature T_0. Non-representativeness of T_0 on specific days involving complex micrometeorological effects was a challenging difficulty. Nonetheless, with reference to radiosoundings (with and without outlier removal from the data collection), the SYN algorithm outperformed all other MLH estimation methods in isolation. In the *diurnal*$^{(i)}$⟦*convective*$^{(ii)}$⟧ time interval, MLH-SYN achieved a correlation coefficient $\rho^{(i)(ii)}_{SYN,1\sigma} \geq 0.98$ and RMSE, $RMSE^{(i)}_{SYN,1\sigma} = \ 76m$ ($RMSE^{(ii)}_{SYN,1\sigma} = 41\ m$). These results compare favorably to MLH-MWR alone: $\rho^{(i)(ii)}_{MWR,1\sigma} \geq 0.94$ and $RMSE^{(i)}_{MWR,1\sigma} = 103\ m$ ($RMSE^{(ii)}_{MWR,1\sigma} = 86\ m$).

To sum up, although SYN used the simplistic assumption of a fixed strongly convective time interval (1000–1400 UTC), the time-adaptive combination of two largely independent methods for MLH tracking, one based on aerosol gradient-based observations (MLH-LC-EKF) and the other temperature-based (MLH-MWR) has shown superior MLH tracking skill. Further research is planned to extend this methodology over the whole diurnal cycle, as well as to explore further synergy with DWL sensors. Besides, a comparison of all the sources of uncertainty in terms of how they balance and propagate would be an interesting study for the remote sensing community.

ACKNOWLEDGMENTS

This work was supported by the Spanish Government – European Regional Development Funds under PGC2018-094132-B-I00 project, and via European Union (EU) Funds under projects ACTRIS-IMP (GA-871115) and GA 101008004 (ATMO-ACCESS). CommSensLab-UPC is a María-de-Maeztu Excellence Unit (MDM-2016-0600) funded by the Agencia Estatal de Investigación, Spain. Data were provided by Jülich Observatory for Cloud Evolution (JOYCE-CF), a core facility funded by Deutsche Forschungsgemeinschaft via grant DFG LO 901/7-1. The work of Marcos P. Araújo da Silva was supported by the Spanish Ministry of Science, Innovation and Universities under Grant PRE2018-086054. Dr. J. H. Schween and Prof. S. Crewell are thanked for the stay of U. Saeed at the Institute for Geophysics and Meteorology, University of Cologne, and data provision. Three unknown reviewers are also thanked for their insightful comments to this paper.

APPENDIX A
MEASURES OF CENTRAL TENDENCY AND VARIABILITY

The mean of the 30-min MLH for the instrument/method combination, *X*, denoted $\mu_{MLH,X}(t_h)$, is computed at time t_h as

$$\mu_{MLH,X}(t_h) = \frac{1}{N}\sum_{i=1}^{N} MLH_X(t_h, d_i), \tag{A.1}$$

where $N = 21$ is the total number of selected days (statistical sample) and d_i denotes the i-th day, $i = 1, \ldots, N$. The variability of the estimated MLH in Eq. (A.1) above is computed as the standard deviation over the sample population,

$$\sigma_{MLH,X}(t_h) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(MLH_X(t_h, d_i) - \mu_{MLH,X}(t_h)\right)^2}. \tag{A.2}$$

The mean of the MLH bias (Eq. 17) between MLH estimates from two different instruments/datasets denoted X and Y is computed at each time t_h as

$$\mu_{bias}^{X,Y}(t_h) = \frac{1}{N}\sum_{i=1}^{N} bias^{X,Y}(t_h, d_i). \tag{A.3}$$

The variability of the MLH bias given by Eq. (A.3) above is computed as the standard deviation,

$$\sigma_{bias}^{X,Y}(t_h) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(bias^{X,Y}(t_h, d_i) - \mu_{bias}^{X,Y}(t_h)\right)^2}. \tag{A.4}$$

REFERENCES

[1] R. B. Stull, *An Introduction to Boundary Layer Meteorology*. Netherlands: Springer, 1988.

[2] S. Emeis, *Surface-based Remote Sensing of the Atmospheric Boundary Layer*. Springer, 2010.

[3] P. Seibert, F. Beyrich, S.-E. Gryning, S. Joffre, A. Rasmussen, and P. Tercier, "Review and intercomparison of operational methods for the determination of the mixing height," *Atmospheric Environment*, vol. 34, pp. 1352–2310, 2000.

[4] G. Xu, B. Xi, W. Zhang, C. Cui, X. Dong, Y. Liu, and G. Yan, "Comparison of atmospheric profiles between microwave radiometer retrievals and radiosonde soundings," *Journal of Geophysical Research: Atmospheres*, vol. 120, no. 19, pp. 10,313–10,323, 2015.

[5] 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.

[6] M. C. 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." *Atmospheric Chemistry & Physics*, vol. 14, no. 23, 2014.

[7] G. Holzworth, "Estimates of mean maximum mixing depths in the contiguous United States," *Monthly Weather Review*, vol. 92, pp. 235–242, 1964.

[8] G. de Arruda Moreira, J. L. Guerrero-Rascado, J. A. Bravo-Aranda, J. A. Benavent-Oltra, P. Ortiz-Amezcua, R. Róman, A. E. Bedoya-Velásquez, E. Landulfo, and L. Alados-Arboledas, "Study of the planetary boundary layer by microwave radiometer, elastic lidar and doppler lidar estimations in southern iberian peninsula," *Atmospheric Research*, vol. 213, p. 185–195, 2018.

[9] L. Belegante, D. Nicolae, A. Nemuc, C. Talianu, and C. Derognat, "Retrieval of the boundary layer height from active and passive remote sensors. comparison with a nwp model," *Acta Geophysica*, vol. 62, no. 2, pp. 276–289, 2014.

[10] D. Cimini, F. D. Angelis, J.-C. Dupont, S. Pal, and M. Haeffelin, "Mixing layer height retrievals by multichannel microwave radiometer observations," *Atmospheric Measurement Techniques*, vol. 6, pp. 4971–4998, 2013.

[11] 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.

[12] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 03 1960. [Online]. Available: https://doi.org/10.1115/1.3662552

[13] S. H. Melfi, J. D. Spinhirne, S.-H. Chou, and S. P. Palm, "Lidar observations of vertically organized convection in the planetary boundary layer over the ocean," *J. Climate Appl. Meteor*, vol. 24, pp. 806–821, 1985.

[14] R. Endlich, F. Ludwig, and E. Uthe, "An automatic method for determining the mixing depth from lidar observations," *Atmospheric Environment*, vol. 13, pp. 1051–1056, 1967.

[15] C. Senff, J. Bösenberg, G. Peters, and T. S. Chaberl, "Remote sensing of turbulent ozone fluxes and the ozone budget in the convective boundary layer with dial and radar-rass: a case study," *Contributions to Atmospheric Physics/Beitraege zur Physik der Atmosphaere*, vol. 61, no. 9, pp. 161–176, 1996.

[16] L. Menut, C. Flamant, J. Pelon, and P. H. Flamant, "Urban boundary-layer height determination from lidar measurements over the paris area," *Appl. Opt*, vol. 38, pp. 945–954, 1999.

[17] 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.

[18] R. F. Banks, J. Tiana-Alsina, F. Rocadenbosch, and J. M. Baldasano, "Performance evaluation of the boundary-layer height from lidar and the weather research and forecasting model at an urban coastal site in the north-east iberian peninsula," *Boundary-Layer Meteorology*, vol. 157, no. 2, pp. 265–292, 2015.

[19] 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.

[20] R. Tanamachi, S. Frasier, J. Waldinger, A. LaFleur, D. Turner, and F. Rocadenbosch, "Progress toward characterization of the atmospheric boundary layer over northern alabama using observations by a vertically pointing, s-band profiling radar during vortex-southeast," *Journal of Atmospheric and Oceanic Technology*, vol. 36, no. 11, pp. 2221–2246, 2019.

[21] M. Barrera-Verdejo, S. Crewell, U. Löhnert, E. Orlandi, and P. D. Girolamo, "Ground-based lidar and microwave radiometry synergy for high vertical resolution absolute humidity profiling," *Atmospheric Measurement Techniques*, vol. 9, no. 8, p. 4013–4028, 2016.

[22] M. A. LeMone, W. M. Angevine, C. S. Bretherton, F. Chen, J. Dudhia, E. Fedorovich, K. B. Katsaros, D. H. Lenschow, L. Mahrt, E. G. Patton, J. Sun, M. Tjernström, and J. Weil, "100 years of progress in boundary layer meteorology," *Meteorological Monographs*, vol. 59, pp. 9.1–9.85, 2018.

[23] A. Macke, P. Seifert, H. Baars, C. Barthlott, C. Beekmans, A. Behrendt, B. Bohn, M. Brueck, J. Bühl, S. Crewell, T. Damian, H. Deneke, S. Düsing, A. Foth, D. Girolamo, H. P., H. E., H. R., K. A., K. J., K. N., K. S., L. M., M. U., B. L., V. Maurer, S. K. Muppa, J. Schween, I. Serikov, H. Siebert, C. Simmer, F. Späth, S. Steinke, K. Träumner, S. Trömel, B. Wehner, A. Wieser, V. Wulfmeyer, and X. Xie, "The hd(cp)^2 observational prototype experiment (hope) – an overview," *Atmos. Chem. Phys*, vol. 17, pp. 4887–4914, 2017.

[24] R. M. Measures, *Laser Remote Sensing: Fundamentals and Applications*, ser. Laser-Remote Sensor Equations. Krieger, Malabar, FL USA: John Wiley & Sons, 1992, ch. 7, pp. 237-280.

[25] 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.

[26] 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.

[27] 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, November 2014.

[28] 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.

[29] 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.

[30] E. J. O'Connor, A. J. Illingworth, I. M. Brooks, C. D. Westbrook, R. J. Hogan, F. Davies, and B. J. Brooks, "A method for estimating the turbulent kinetic energy dissipation rate from a vertically pointing doppler lidar, and independent evaluation from balloon-borne in situ measurements," *Journal of Atmospheric and Oceanic Technology*, vol. 27, no. 10, pp. 1652–1664, 2010.

[31] S. A. Cohn and W. M. Angevine, "Boundary layer height and entrainment zone thickness measured by lidars and wind-profiling radars," *Journal of Applied Meteorology*, vol. 39, no. 8, pp. 1233–1247, 2000.

[32] AG Crewell IGMK, "Instruments operated by RG integrated remote sensing," http://gop.meteo.uni-koeln.de/ag_crewell/doku.php?id=instruments:instruments, 2020. (accessed, Dec. 2020).

[33] J. Güldner and D. Spänkuch, "Remote sensing of the thermodynamic state of the atmospheric boundary layer by ground-based microwave radiometry," *J. Atmos. Ocean. Tech*, vol. 18, pp. 925–933, 2001.

[34] D. Cimini, T. J. Hewison, L. Martin, J. Güldner, C. Gaffard, and F. S. Marzano, "Temperature and humidity profile retrievals from ground based microwave radiometers during tuc," *Meteorol. Z*, vol. 15, pp. 45–56, 2006.

[35] 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.

[36] U. Löhnert, S. Crewell, and C. Simmer, "An integrated approach toward retrieving physically consistent profiles of temperature, humidity, and cloud liquid water," *Journal of Applied Meteorology*, vol. 43, no. 9, pp. 1295 – 1307, 01 Sep. 2004. [Online]. Available: https://journals.ametsoc.org/view/journals/apme/43/9/1520-0450_2004_043_1295_aiatrp_2.0.co_2.xml

[37] R. J. Hogan, A. L. M. Grant, A. J. Illingworth, G. N. Pearson, and E. J. O'Connor, "Vertical velocity variance and skewness in clear and cloud-topped boundary layers as revealed by doppler lidar," *Quarterly Journal of the Royal Meteorological Society*, vol. 135, no. 640, pp. 635–643, 2009.

[38] G. Pearson, F. Davies, and C. Collier, "Remote sensing of the tropical rain forest boundary layer using pulsed doppler lidar," *Atmos. Chem. Phys*, vol. 10, pp. 5891–5901, July 2010.

[39] G. N. Pearson and C. G. Collier, "A pulsed coherent CO2 lidar for boundary-layer meteorology," *Quarterly Journal of the Royal Meteorological Society*, vol. 125, no. 559, pp. 2703–2721, 1999.

[40] C. Flamant, J. Pelon, P. H. Flamant, and P. Durand, "Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary-layer," *Bound.-Layer Meteor*, vol. 83, pp. 247–284, 1997.

[41] R. Boers and E. W. Eloranta, "Lidar measurements of the atmospheric entrainment zone and the potential temperature jump across the top of the mixed layer," *Bound.-Layer Meteor*, vol. 34, pp. 357–375, 1986.

[42] M. de Bruine, A. Apituley, D. P. Donovan, H. Klein Baltink, and M. J. de Haij, "Pathfinder: applying graph theory to consistent tracking of daytime mixed layer height with backscatter lidar," *Atmospheric Measurement Techniques*, vol. 10, no. 5, pp. 1893–1909, 2017. [Online]. Available: https://amt.copernicus.org/articles/10/1893/2017/

[43] Y. Poltera, G. Martucci, M. Collaud Coen, M. Hervo, L. Emmenegger, S. Henne, D. Brunner, and A. Haefele, "Pathfinderturb: an automatic boundary layer algorithm. development, validation and application to study the impact on in situ measurements at the jungfraujoch," *Atmospheric Chemistry and Physics*, vol. 17, no. 16, pp. 10 051–10 070, 2017. [Online]. Available: https://acp.copernicus.org/articles/17/10051/2017/

[44] 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.

[45] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press, 1964.

[46] R. F. Banks, J. Tiana-Alsina, J. M. Baldasano, F. Rocadenbosch, A. Papayannis, S. Solomos, and C. G. Tzanis, "Sensitivity of boundary-layer variables to pbl schemes in the wrf model based on surface meteorological observations, lidar, and radiosondes during the hygra-cd campaign," *Atmospheric Research*, vol. 176-177, pp. 185–201, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169809516300412

[47] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in The Physical Sciences*, R. E. F. Mandl and D. Sandiford, Eds. Chichester, England: Wiley, 1989, ch. 4, Errors, pp. 48–67.

[48] Reif, K. and Gunther, S. and Yaz, E. and Unbehauen, R., "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714–728, 1999.

[49] S. Wingo and K. Knupp, "Multi-platform observations characterizing the afternoon-to-evening transition of the planetary boundary layer in northern alabama, usa," *Boundary-Layer Meteor*, vol. 155, pp. 29–53, 2015.

[50] M. Haeffelin, F. Angelini, Y. Morille, G. Martucci, S. Frey, G. Gobbi, S. Lolli, C. O'Dowd, 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.

[51] D. Lenschow, J. C. Wyngaard, and W. T. Pennell, "Mean-field and second-moment budgets in a baroclinic, convective boundary layer," *Journal of the Atmospheric Sciences*, vol. 37, no. 6, pp. 1313–1326, 1980.

[52] S. C. Tucker, C. J. Senff, A. M. Weickmann, W. A. Brewer, R. M. Banta, S. P. Sandberg, D. C. Law, and R. M. Hardesty, "Doppler lidar estimation of mixing height using turbulence, shear, and aerosol profiles," *Journal of Atmospheric and Oceanic Technology*, vol. 26, no. 4, pp. 673–688, 2009.

[53] K. Träumner, C. Kottmeier, U. Corsmeier, and A. Wieser, "Convective boundary-layer entrainment: Short review and progress using doppler lidar," *Boundary-Layer Meteorology*, vol. 141, no. 3, pp. 369–391, 2011.

[54] J. Villalonga, S. L. Beveridge, M. P. Araujo da Silva, R. L. Tanamachi, F. Rocadenbosch, D. D. Turner, and S. J. Frasier, "Convective boundary-layer height estimation from combined radar and Doppler lidar observations in VORTEX-SE," in *Remote Sensing of Clouds and the Atmosphere XXV*, A. Comerón, E. I. Kassianov, K. Schäfer, R. H. Picard, K. Weber, and U. N. Singh, Eds., vol. 11531, International Society for Optics and Photonics. SPIE, 2020, pp. 192 – 201. [Online]. Available: https://doi.org/10.1117/12.2576046

[55] J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*. Elsevier, 2006.

[56] D. J. Seidel, C. O. Ao, and K. Li, "Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis," *Journal of Geophysical Research: Atmospheres*, vol. 115, no. D16, 2010. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD013680

[57] U. Löhnert and S. Crewell, "Accuracy of cloud liquid water path from ground-based microwave radiometry - Part I: Dependency on cloud model statistics and precipitation," *Radio Science*, vol. 38, no. 3, 2003.

[58] U. Löhnert, D. D. Turner, and S. Crewell, "Ground-based temperature and humidity profiling using spectral infrared and microwave observations. part i: Simulated retrieval performance in clear-sky conditions," *Journal of Applied Meteorology and Climatology*, vol. 48, no. 5, pp. 1017 – 1032, 01 May. 2009. [Online]. Available: https://journals.ametsoc.org/view/journals/apme/48/5/2008jamc2060.1.xml

[59] F. Rocadenbosch, R. L. Tanamachi, M. P. Araujo da Silva, J. Villalonga, S. J. Frasier, and D. D. Turner, "Atmospheric boundary layer height disambiguation using synergistic remote sensing observations: case examples from VORTEX-SE," in *Remote Sensing of Clouds and the Atmosphere XXV*, A. Comerón, E. I. Kassianov, K. Schäfer, R. H. Picard, K. Weber, and U. N. Singh, Eds., vol. 11531, International Society for Optics and Photonics. SPIE, 2020, pp. 109 – 120. [Online]. Available: https://doi.org/10.1117/12.2576093

[60] F. Rocadenbosch, R. Barragán, S. J. Frasier, J. Waldinger, D. D. Turner, R. L. Tanamachi, and D. T. Dawson, "Ceilometer-based rain-rate estimation: A case-study comparison with s-band radar and disdrometer retrievals in the context of vortex-se," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8268–8284, 2020.

Marcos P. Araújo da Silva received the B.S. degree in Environmental Engineering from the Federal University of Rio Grande do Norte, Natal, Brazil, in 2015, the M.S. degree in Climate Sciences from the same University, in 2018. He is currently pursuing his Ph.D. degree in Telecom Engineering, Remote Sensing (advisor, Prof. F. Rocadenbosch) at María-de-Maeztu Excellence Unit CommSensLab - UPC, granted by the Ministerio de Ciencia e Innovación (Spain). His research interests include atmospheric remote sensing and boundary layer, related multi-sensor processing and wind energy.

Francesc Rocadenbosch (Senior Member, IEEE) received the B.S. and Ph.D. degrees in telecommunications engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1991 and 1996, respectively, and the M.B.A. degree from the University of Barcelona, Barcelona, in 2001.

In 1993, he joined the Department of Signal Theory and Communications, UPC, where he is a Full Professor. Since 1996, he has steered the development of CommSensLab-UPC Excellence-Unit activities on lidar, including the UPC unmanned lidar station (ESFRI 2016 roadmap), and he has served on over ten collaborative remote sensing projects as PI. His research interests include remote sensing (lidar-radar-radiometry) for atmospheric observation, related signal processing, and off-shore wind-lidar.

Dr. Rocadenbosch was a recipient of the “Salvà i Campillo” Best Research Project award in 1997, the National Telecom Award in 2003 (group), and the EU KIC recognition for EOLOS spin-off in 2015. He is an Associate Editor of the IEEE Transactions on Geoscience and Remote Sensing.

Robin Tanamachi received the B. S. degree in Atmospheric and Oceanic Sciences from the University of Wisconsin - Madison in 2001, and M. S. and Ph. D. degrees in Meteorology from the University of Oklahoma in 2004 and 2011, respectively. She is currently an Assistant Professor in the Department of Earth, Atmospheric, and Planetary Sciences at Purdue University. Her interests include radar meteorology, severe thunderstorm dynamics and kinematics, and methods for assimilating radar observations into numerical weather prediction models. Dr. Tanamachi is a member of the American Meteorological Society (AMS) and served on the AMS Radar Meteorology committee from 2013 to 2019.

Umar Saeed received the B.E. degree in electronics engineering from the National University of Sciences and Technology (NUST), Karachi, Pakistan, in 2007, the M.Sc. degree in communications engineering, with specialization in digital signal processing, from Aalto University, Helsinki, Finland, in 2012, and the Ph.D. degree in atmospheric remote sensing from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 2016. From 2013 to 2016, he was a Marie Curie Early Stage Researcher in the Initial Training for Atmospheric Remote Sensing network funded by the Seventh Framework Programme (FP7) of the European Union. His research interests include active optical and passive microwave atmospheric remote sensing, atmospheric boundary layer, sensor synergy, and statistical and adaptive signal processing.