ICOS improved sensors, network and interoperability for GMES

D3.4

Report on quality and performance of interpolated BLH retrieval product and cross-validation method

D. Feist, Juan Antonio Bravo Aranda, Christophe Pietras, G.Biavati and J.Tarniewicz

ICOS INTEGRATED CARBON OBSERVATION SYSTEM

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n°313169.

ICOS-INWIRE (GA n° 313169)
Document for programme participants
Version 1.0, December 2015
www.icos-inwire.lsce.ipsl.fr

Deliverable:	**D3.4, Report on quality and performance of interpolated BLH retrieval product and cross-validation method**
Author(s):	**D. Feist, J. Bravo Aranda, Ch. Pietras, G. Biavati and J. Tarniewicz**
Date:	**15/12/2015**
Activity:	**WP3**
Lead Partner:	**MPG/LMD/CEA**
Document Issue:	**1.0**
Dissemination Level:	**Programme participants(PP) - internal**
Contact:	**Jerome Tarniewicz, CEA/LSCE** **Jerome.tarniewicz@lsce.ipsl.fr**

	Name	Partner	Date
Written by	D. Feist, G. Biavati J. Bravo Aranda, Ch. Pietras, J. Tarniewicz	MPG CNRS-LMD CEA/LSCE	15/12/2015
Reviewed by	Jean-Daniel Paris	CEA	16/12/2015
Approved by	Jean-Daniel Paris	CEA	17/12/2015

Table of content

Introduction

One of WP3 objectives is to demonstrate the contribution of lidar equipping ICOS to diagnose boundary layer height (BLH). BLH is a critical variable that defines the dilution of CO_2 and CH_4, emitted or taken up at the surface, in the atmosphere.

Two atmospheric boundary layer retrieval algorithms have been selected in WP3 (PyBL & STRAT), and the choice was made to have them improved by their respective developers (resp. MPG & CNRS-LMD) to meet the specific needs of ICOS-INWIRE in terms of BLH restitution and implementation on site in the ICOS network (see MS7, Develop reference implementation for unified pre-processing and BLH retrieval algorithm). It results in two adapted versions (PyBL_ICOS & STRAT+) of original algorithms that are compared to each other in a first part of this deliverable, over a common dataset of measurement, thanks to standard data format for backscatter signals from lidar/ceilometers defined in MS6.

Since optical systems will not necessarily be exactly collocated with atmospheric stations in the ICOS network, spatial interpolation of mixing heights is investigated in a second part of this deliverable. Two interpolation schemes (distance-based and geo-statistical) are tested. In particular results of kriging methods with mixing heights from atmospheric model to physically constrain the interpolation are presented.

In a last part, quality and performance of BLH retrieval are studied. In particular lidar/ceilometer sensitivity for BLH retrieval is addressed, and cross validation of BLH retrievals by ceilometers versus micro wave radiometer and radiosondes mixing heights is shown.

Algorithms description and comparison

STRAT+ algorithm

STRAT+ is an algorithm for the PBL height attribution based on lidar and sonic anemometer observations considering the boundary layer processes for the purpose of attribution (Pal et al., 2013). STRAT+ is based on three main modules: STRAT, variance method and final attribution summarized in Figure 1.

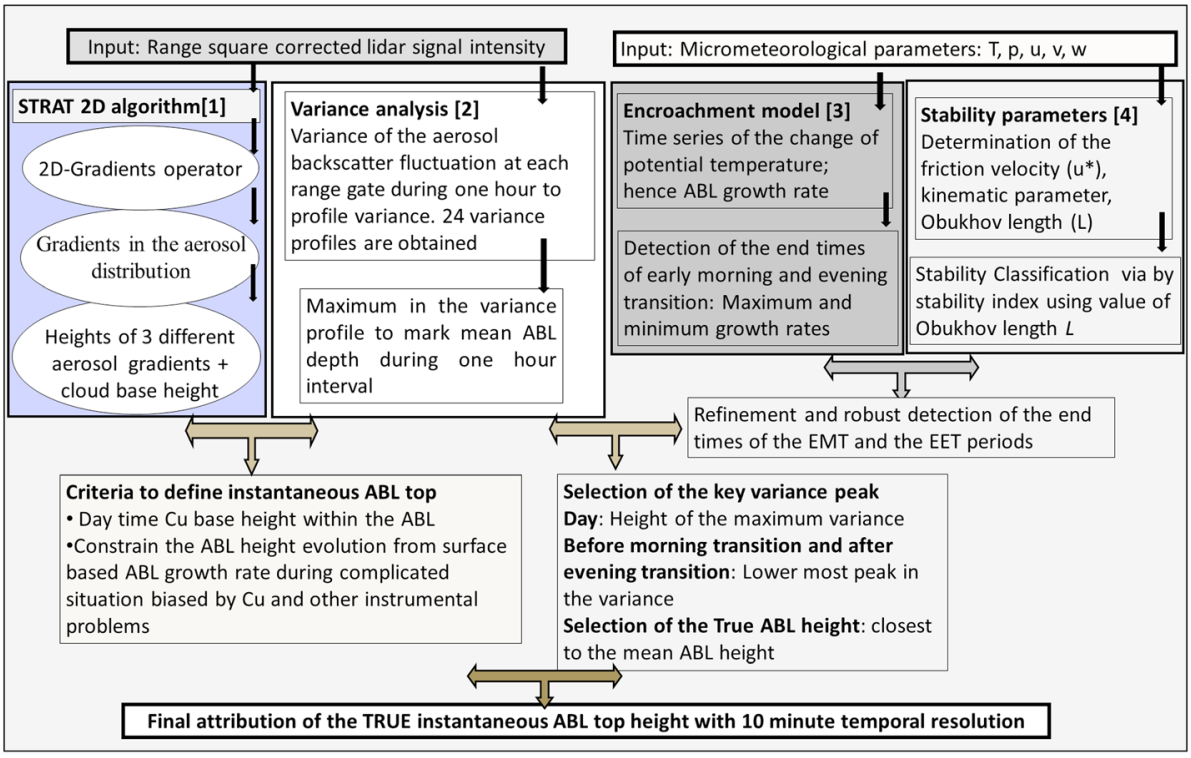

Figure 1: STRAT+ flow chart to derive high-resolution BLH.

STRAT+ algorithm is available upon request by directly getting into contact with SIRTA observatory.

PyBL_ICOS algorithm

For more details, PyBL_ICOS retrieval algorithm is described in depth in ICOS-INWIRE deliverable D3.3 – "Unified Lidar preprocessing and mixing height retrieval algorithm" deliverable. A preliminary version of the software PyBL_ICOS [MS7] has been developed in Python, which is freely available for all platforms, and can be downloaded through Mercurial repository [https://projects.bgc-jena.mpg.de/hg/PyBL_ICOS/]. Please get into contact with MPG to get credentials to access the repository.

Algorithm comparison

ALS 400 (from Leosphere, France) is a lidar system that was deployed at SIRTA observatory (Palaiseau, near Paris, France) between 2008 and 2013. The ALS lidar performed measurements for climatological studies (e.g Pal et al, 2013 and 2015). One month of this dataset (June 2011) is independently processed by algorithms from MPG and LMD. BLH time series for the whole month of measurement are then plotted and compared on Figure 2.

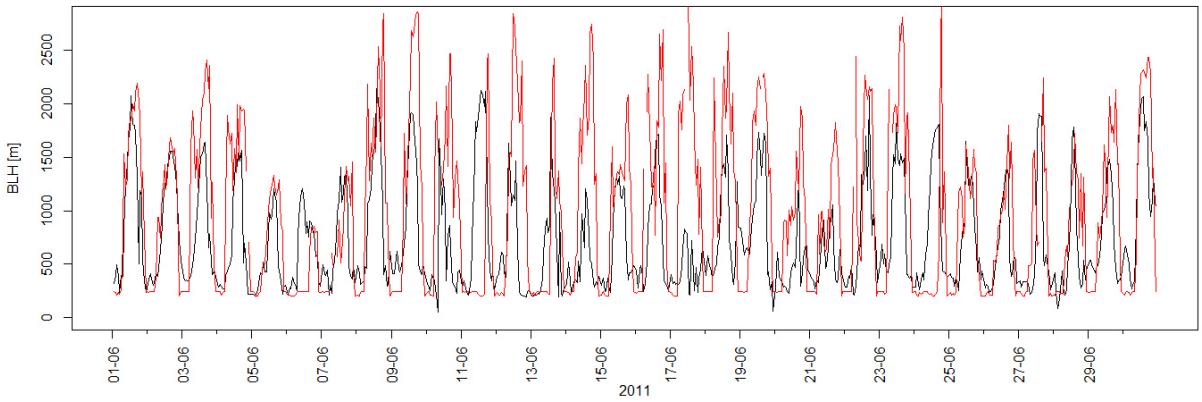

Figure 2: Time series of estimated BLH with one common dataset acquired at SIRTA [Palaiseau, France] for STRAT+ (red line) and PyBL_ICOS (black line) for one month of measurement in June 2011. BLH estimates are given on an hourly basis.

The daily cycle of BLH is well reproduced with both algorithms. Differences in restitution are nevertheless observed PyBL_ICOS providing smaller values than STRAT+. Day to day variation of BLH estimates is represented on scatter plots on Figure 3. Since lidar is an optical method, first thing is to have a look to meteorological conditions for June 2011 in order to try to explain difficulties in BLH retrievals for PyBL_ICOS and STRAT+ (see Table 1).

Meteorological conditions	Rainy days	Foggy days or with low clouds	Clear sky
Days #	4,5, 9-14, 18-20	6-7, 16-17, 21-22	1-3, 8, 15, 23-27, 29-30.

Table 1 – Summary of meteorological conditions observed for June 2011.

PyBL an STRAT+ algorithm generally provide good daily BLH evolution with smaller values provided by PyBL. and STRAT+ provide unrealistic results on 12, 17, 20 and 6, 11, 13 respectively due to complex meteorological situation (rainy and low clouds). A weak interaction between the surface and the low part of the troposphere (e.g., strong air mass advection) can be the largest difficulty for

the BLH retrieval and thus, the comparison of BLH estimated under different meteorological conditions clearly needs to be taken to a step further.

High BLH are retrieved by STRAT+ for some days (9 and 16 for example) whereas PyBL_ICOS BLH remains below 2 km almost the whole period. Although high BLH are not common in mid latitude for the season, recent paper on BLH retrievals (S. Pal et al, JGR 2015) shows that BLH could be as high as 3000 meters and could last several hours a day in spring and summer.

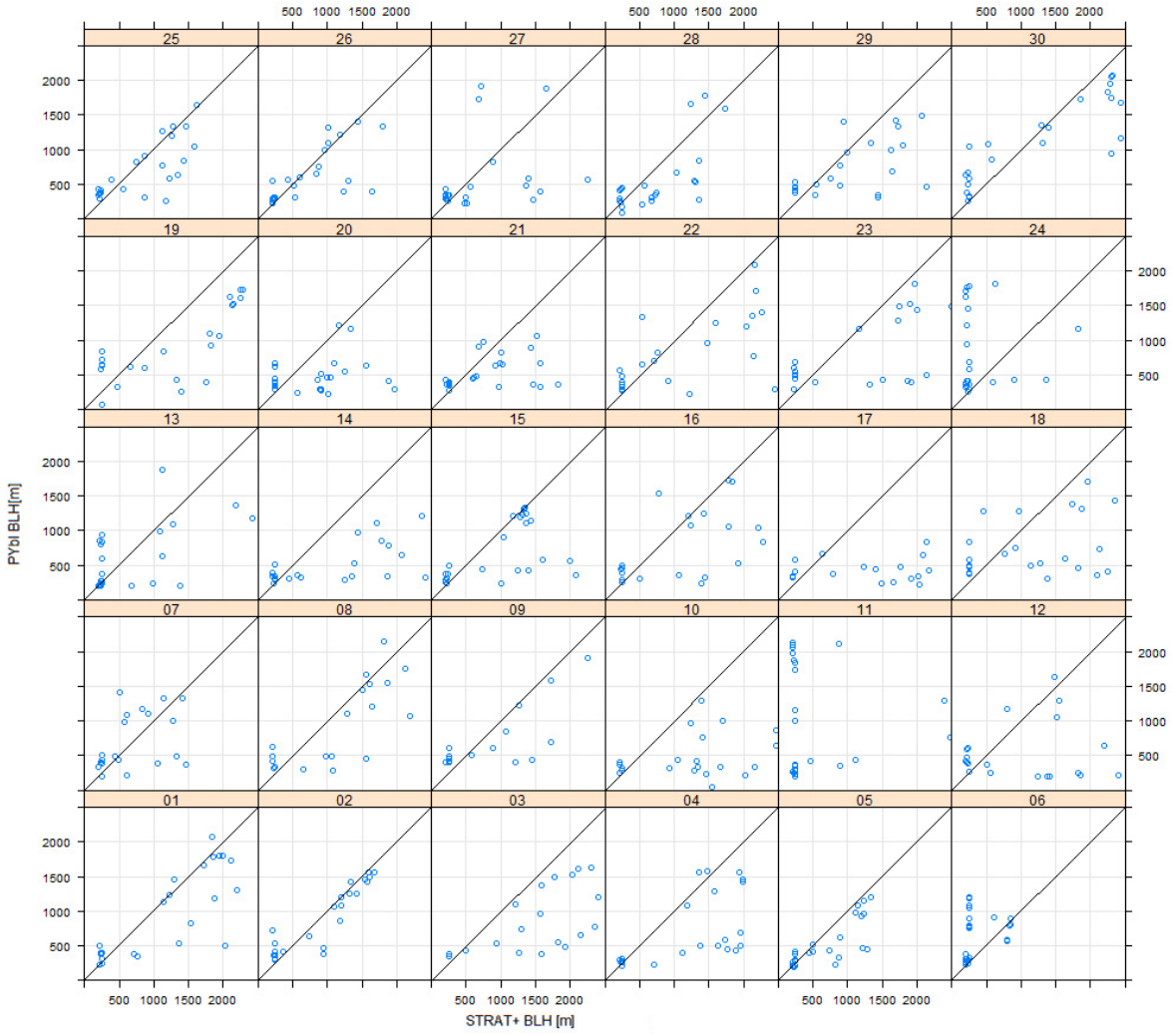

Figure 3 – Day to day scatter plots of STRAT+ versus PyBL_ICOS estimates for June 2011 [blue points].

BLH derived from STRAT show larger temporal variability than PyBL_ICOS because: on the first hand, STRAT+ is fully based on experimental data whereas PyBL_ICOS is much more constrained by its underlying atmospheric transport model. On the second hand, the low STRAT+ BLH resolution (10-min for STRAT+ and 1hour for PyBL_ICOS) used to properly track the BL development between the sunrise and midday. On the third hand, BL top horizontally varies according to the updraft and

downdraft of the convective cells. Although the ceilometers stations are fixed, convective cells can be advected and thus, the BLH derived from ceilometers varies depending on the atmospheric region explored (either updraft or downdraft).

Some unrealistic STRAT+ BLH developments are observed. This is probably caused either by a wrong variance profile (probably due to low signal-to-noise ratio), by the discontinuity of the maximum BLH thresholds between night- and day-time or by wrong stability conditions derived from sonic measurements. Although BLH attribution could still be improved using a "pathfinder" method, STRAT+ relies on the BL physical foundation whereas PyBL_ICOS relies on an underlying atmospheric transport model.

Mixing height interpolation and verification

Already Holzworth (1967) concluded that a good knowledge of boundary layer height (BLH) or mixing height (MH) over the spatial domain of interest is a requirement for studies on pollution and in general fluxes from the surface. However, models often fail in producing estimates of this key parameter. Therefore, the idea of interpolating BLH to a larger domain from individual observation sites was explored as part of this project.

Mixing height estimates from ceilometer networks

In recent times, networks of ceilometers have been set up in several countries (see Figure 4). The main goal of such a network is estimation of cloud base height. The increasing quality of the instruments also permits retrieval of MH (as well as BLH) under certain conditions.

This network of instruments is a good starting point for testing how the obtained information can be used to constrain transport models in order to increase the accuracy of greenhouse gas and pollutant estimates. This study tried to assess the quality of the data within the network and to compare the retrieved MH with independent datasets.

The data provided by the German Weather Service/Deutscher Wetterdienst (DWD) contained a short time series for the stations in Figure 4. The time period ran from 25 September to 14 October 2009. The instruments were operated within the nominal parameters but not all instruments provided data during the whole time period.

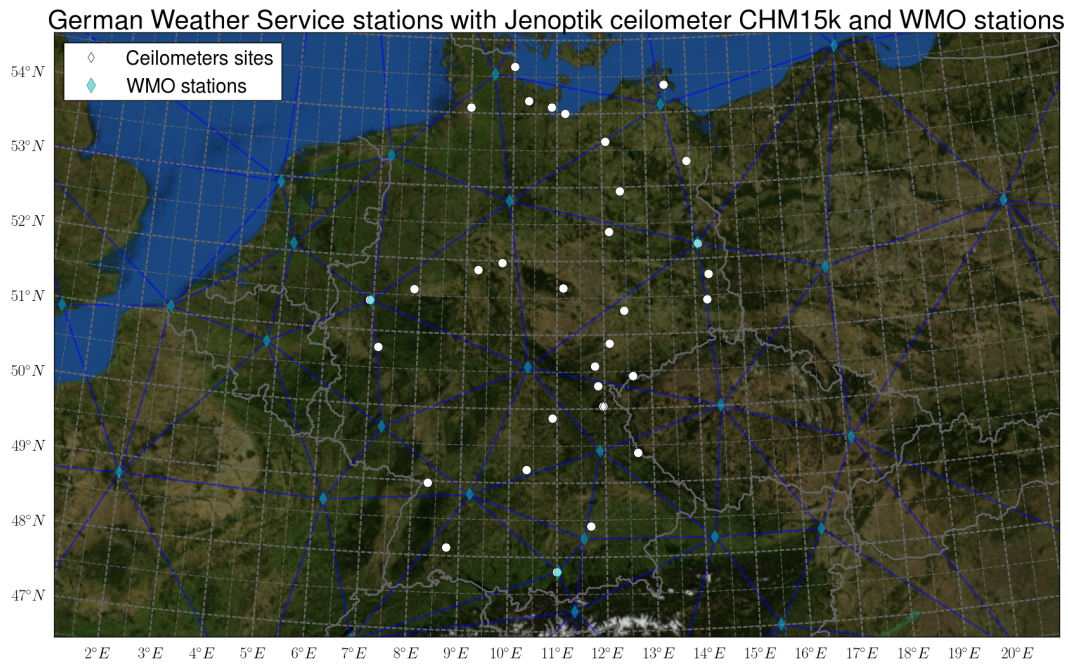

Figure 4: Stations of the German Meteorological Service equipped with a CHM15k ceilometer (white dots) as well as WMO radiosonde network stations (blue diamond).

Comparisons with MH derived from radiosondes

To evaluate the ability of the ceilometer network to reproduce the evolution of the boundary layer and, in particular the Convective Boundary Layer, we need an objective reference to test our estimates of MH. As done for the Jenoptik ceilometer installed at Lindenberg, the results can be compared with MH obtained from radiosondes. Unfortunately, only a few of the ceilometer stations in Figure 4 also have radiosondes available for comparison. However, Figure 4 also shows WMO stations where radiosondes are launched at 00:00 and 12:00 UTC. At a smaller subset of stations, radiosondes are also launched at 06:00 and 18:00 UTC.

For the comparison, two different strategies were considered for obtaining an estimate of MH at the ceilometer locations: a linear interpolation of the estimated MH at the WMO stations and a geo-statistical interpolation driven by model data.

Comparison with distance-based interpolation

For this comparison, only data at 00:00 and 12:00 UTC were considered. The use of other times would imply a time-space interpolation that can cause additional errors. This is because only a few stations produce data at 06:00 and 18:00 UTC. However, it was already known from previous studies on the Lindenberg dataset that the nocturnal boundary layer is mostly outside the capability of the installed CHM15k ceilometers.

The radiosonde stations are scattered over the European domain. A preliminary consideration must be taken to perform a spatial interpolation of MH: the MH can be defined as height above ground or above sea level. The first choice is the most reasonable. Using values referred above sea level could potentially generate negative interpolated values.

The interpolation of scattered data can be done by a Delaunay triangulation. In mathematics and computational geometry, a Delaunay triangulation for a set P of points in a plane is a triangulation DT(P) such that no point in P is inside the circumcircle of any triangle in DT(P). This method is rather common and based on distance weighting. It defines a particular tessellation of the domain for each point where we want to interpolate on a set of three WMO stations for which the influence is considered optimal. Then, the quantities are calculated for the target point as a weighted average.

Delaunay triangulation cannot be used directly to perform temporal interpolations because time and space cannot be mixed in weighting. Therefore, space interpolation is done first, followed by time interpolation afterwards.

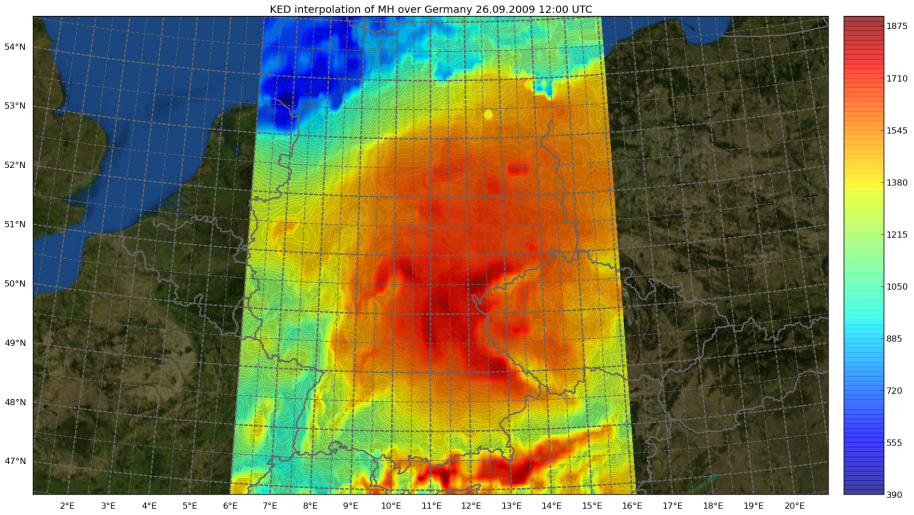

Figure 5: Example of KED-interpolated MH from high-resolution model data by Kretschmer et al. (2014).

Comparison with geo-statistical interpolation

To compare the MH retrieved on the WMO network of stations with a random station of the DWD, we already introduced distance-based interpolation. Other methods are reported in the literature. In particular, we also tested geostatistical interpolation. Without going into detail, geostatistical interpolation (or *Kriging*), is an advanced methodology to define weighting factors for the interpolation. The main idea is to use time series and spatial maps of MH in order to produce optimal weights to be more consistent with the observed phenomena.

In the same department, the use of transport models for estimating greenhouse gas sources and sinks were also explored (Kretschmer et al. (2012), Kretschmer et al. (2014)). Figure 5 shows an example of the high-resolution interpolated MH maps that were produced by applying "Kriging with External Drift" (KED) to meteorological fields from the high-resolution WRF model.

Global	*MH0*	*MH*	*KED*	*rs*
MH0	1.00	0.15	0.13	0.15
MH	0.15	1.00	0.47	0.53
KED	0.13	0.47	1.00	0.76
rs	0.15	0.53	0.76	1.00
Day	*MH0*	*MH*	*KED*	*rs*
MH0	1.00	0.19	0.11	0.14
MH	0.19	1.00	0.25	0.34
KED	0.11	0.25	1.00	0.67
rs	0.14	0.34	0.67	1.00
Night	*MH0*	*MH*	*KED*	*Rs*
MH0	1.00	0.18	0.17	0.18
MH	0.18	1.00	0.49	0.54
KED	0.17	0.49	1.00	0.76
rs	0.18	0.54	0.76	1.00

Table 2: Linear correlation of different estimates of mixed layer height within the DWD network of meteorological stations. The results are grouped into three different classes. Global: all available data; Day: data from 06:00 to 18:00 UTC; Night: data from 18:00 to 06:00 UTC.

Results for the DWD network

Statistics for MH estimated using the ceilometer network and the different interpolation methodologies are presented in the following paragraphs. Results are focused on four different estimates: the most likely edge *MH0*; the estimate obtained using the selection method *MH*; the linear interpolation of the values estimated from the WMO stations *rs*; the geospatial interpolation *KED*. The linear correlation calculated on the data using different classifications of the data is presented in Table 2. We also performed the classification per hour of the day as shown in Figure 6.

The results of Table 2 reveal several things.. The couple *rs-KED* has the best correlation of the table. This result was expected, because the process that produces *KED* uses the same MH obtained from the WMO stations.

An unexpected result is the difference of correlations of the couples *MH-rs* and *MH-KED*. The selected MH follow *rs* better than *KED*. This remains also valid when checking Figure 6. The worst hour of the day is 05:00 UTC. This reflects the fact that at 06:00 UTC at most of the WMO stations, there is no radiosonde launch. The negative values for the couples including *MH0* show that this choice is the worst.

Another relevant result is the value of correlation at 12:00 UTC. The selection method *MH* seems to be weaker than the *MH0*. Around this time, the solar forcing is generally stronger, the convection reaches the highest altitudes and the MH is located mostly below the free atmosphere. The presence of weaker edges within the MH makes it easier to follow a trend of a reference model. This also works when it reproduces the physical phenomena wrongly.

The non-promising results can also be related to the meteorological conditions. During the period of available data, a low pressure system covered central Europe. There were only short periods of clear sky with pure convection. Due to strong clouds and frontal structures passing over the stations, the ergodicity of the measurements was heavily affected by the meteorological conditions. Also the height of the MH was often low. That is a problem when the MH is in the altitude range where the field of view of the ceilometer's telescope and the cone of backscattered light only overlaps partially.

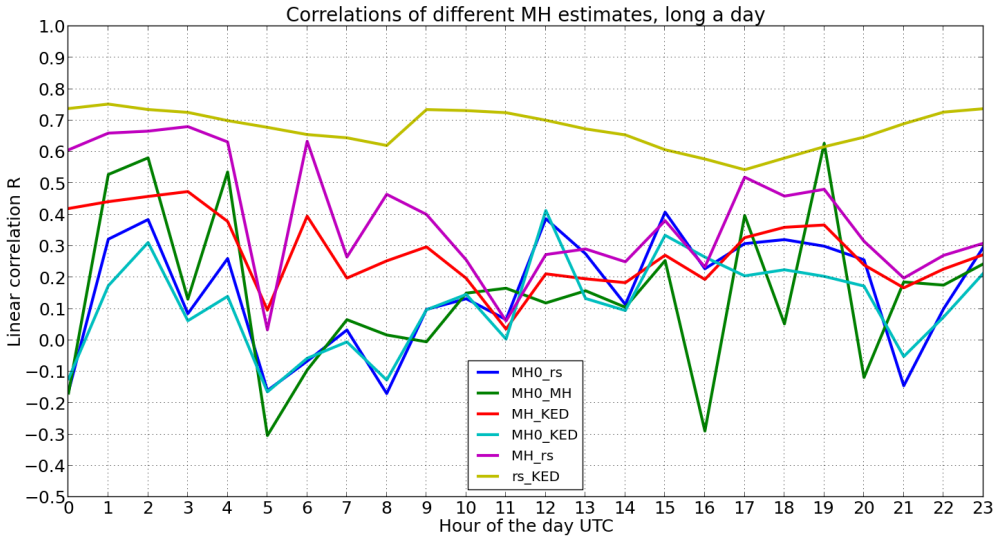

Figure 6: Correlation of different MH estimates over the DWD stations on an hourly basis.

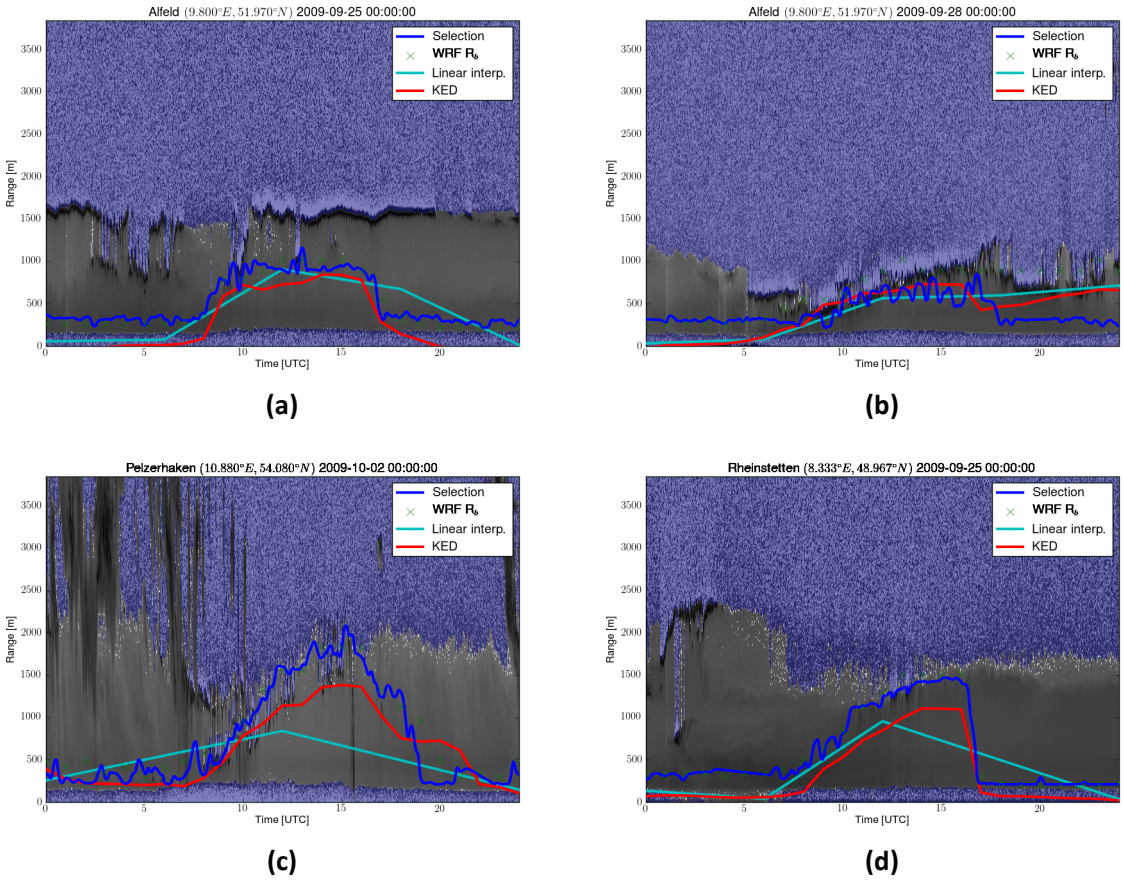

Figure 7: Each of the plots presents the attenuated volume backscatter coefficient signal in gray. The red line represents the KED. The distance-based interpolation of the RBN selections are plotted in cyan. The green crosses represent the result of RBN applied to the WRF model output and is considered here as the reference for the selection method. The blue line is the result after the selection of candidates MH.

The best option would be to perform the analysis on a longer period. Preferably, this should be done during summer, when MH is higher, far from the partial overlap region of the ceilometers used in this study. Some examples of daily evolutions of MH are presented in Figure 7 as well as Figure 8.

A relevant example can be seen when comparing Figure 8 (a) and (b). The two instruments co-located at Lindenberg have different performances. This can be seen by comparing the signal-to-noise mask. The results are similar, but not identical. The exact location of the two instruments was not reported. From the time delay of signals from overpassing clouds, we estimated the distance to be at least several hundred meters. This difference can have strong implications for the validity of the retrieval.

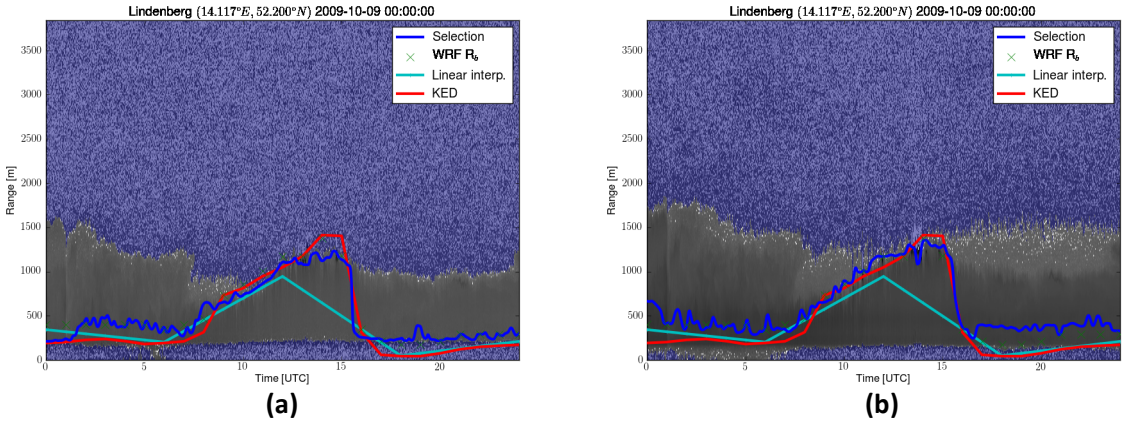

Figure 8: Analogous to Figure 7. However, the plots were produced using two different CHM15k ceilometers installed at Lindenberg: (a) CHM060012, (b) CHM080066.

The method of selection on the 15-second-profiles also has a strong effect on the retrieval. High frequencies are still visible even after time-averaging on 30-minute-wide windows. These high frequency fluctuations can be one of the reasons for failure of the comparison of the selected data with both *KED* and *rs*.

The better correlations obtained at night time have to be considered with care. Fluctuations of MH during night are much smaller than during daytime. The edges within the residual layer offer many candidates for selection so that it is easier to get a good match with the model which more or less reflects the interpolations.

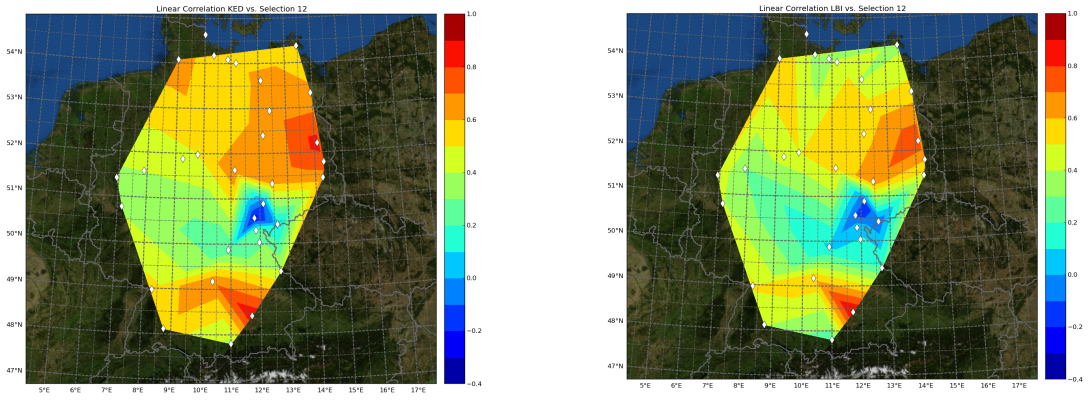

Figure 9: Comparison of KED-interpolated MH (a) and linearly interpolated MH (b) at 12:00 UTC.

The reason why the comparison with *KED* gave the worst results must be explored further. Conditions were not optimal: independent estimates of MH and more ceilometer data would be needed to produce better statistics with classes discriminating for different degrees of stability.

There is the need for capturing the evolution of MH for a longer time and at a higher temporal resolution.

In Figure 6, correlations between the parameters retrieved from the ceilometer network and the interpolated MH from the WMO radiosonde network are provided over time. Figure 9 provides a complementary comparison over space. The maps show the correlation between the *MH* value (from the selction method) retrieved at the ceilometers sites at 12:00 UTC with interpolated MH from the WMO network.

Two different methods were used to interpolate the WMO network data: geostatistical (KED) and linear interpolation. On both maps, the correlation is only high at the ceilometer stations that also have radiosonde launches. There is an interesting anticorrelation in east-central Germany which is probably due to topographic effects and the fact that none of these sites have radiosonde launches. In general, the overall correlation is slightly better for KED than for the linear interpolation.

Discussion

The measurements done by the LIDAR are only representative for a comparatively small volume. This means that they are influenced by local conditions which make it hard to extend them to a larger scale. The MHs retrieved as diagnostic variables from models are hardly comparable to the point measurements. The 10-km-grid resolution used in the WRF model, which is used to make the geostatistical interpolation and the reference for the selection method, makes it impossible to obtain a linear correlation of 1.0 with the estimated MH. This is because they are affected by high frequencies. In other words: because of sub-grid variability that the model cannot reproduce.

The need of an accurate map of mixing height, to constrain transport models, as introduced by Holzworth (1967), was the idea at the beginning of this study. The idea of installing instruments for estimating MH on a network still remains an optimal choice. However, nothing can be said about the minimal distance required between instruments in order to produce correct maps of MH. Together with the network, a proper interpolation must be performed in order to cover the full domain. The most promising way remains the geostatistical interpolation as introduced by Kretschmer et al. (2012) and Kretschmer et al. (2014).

The obtained results are still not mature enough to be used for the interpolations. The number of stations of the DWD is probably too small, and their spatial distribution has relevant holes in the domain of Germany.

More studies should be performed in order to combine radiosonde profiles, optical measurements and model data. In particular, studies on the sub grid variability and on the interpolation methods should provide the directions for the development of such networks together with the way to use the estimated data.

Quality and performance of BLH retrieval

Aerosol characterization (e.g., aerosol vertical structure and aerosol optical properties) derived from ceilometers measurements were possible the last years due to hardware & firmware improvements although they were originally developed for cloud base height detection. Particularly, BLH has become one of the most important parameters that can be provided together with the attenuated aerosol-particle backscatter coefficient. However, ceilometer types show firmware and hardware differences that may affect the final product. These differences have to be considered by the BLH retrieval algorithms, such as STRAT+ that was initially developed to be applied to lidar systems but was later optimized to be applied to ceilometers. STRAT+ algorithm performance (number of BLH detection, instrument BLH retrieval sensitivity and the agreement with the BLH determined by means of radiosonde) has been assessed using measurements from 12 co-located ceilometers (six different types) and a lidar system (Table 3) during the CEILINEX experimental campaign. This campaign was carried out in the framework of TOPROF (Cost Action ES1303) from June to September 2015. Additionally, STRAT+ algorithm has been applied to measurements acquired by the CHM15k model at 3 stations over Europe to investigate its performance for different situations and atmospheric conditions (Granada, Spain; Payern, Switzerland and Palaiseau, France).

#	Instrument type	Instrument name	STRAT+ detected-to-potential BLH percentage (%)
1	CHM 15k Nimbus	CHM100110	97
2		CHM140101	96
3	CHX 15k Nimbus	CHX080082	96
4		CHXLMU	93
5	CL31	CL31RUB	91
6		CL31RAO	94
7	CL51	CL51RAO	96
8		CL51CG	94
9	CS135	CS1	94
10		CS2	92
11	LD40	LD40_002	94
12		LD40_003	95
13	PollyXT	RALPH	92

Table 3 - List of ceilometers and lidar systems participating in CEILINEX2015 at Lindenberg (Germany).

Table 3 shows the detected-to-potential BLH percentage (%) for all systems. STRAT+ provided between 91% and 97% of the potential 10-min BLH during the campaign. The models CHM, CL51 and LD40 showed the largest percentages (between 94% and 97%). Although the model differences are almost negligible for all of them (less than 3%), CHM and LD40 are the most stable performance with a difference of 1% between the CHM100110 and CHM140101 and LD40_002 and LD40_003.

STRAT+ and radiosonde comparison

During the CEILINEX campaign, 4 radiosondes per day were launched. Richardson number method was used to derive the BLH. These results were used to test STRAT+ performance on different ceilometer types. Since ceilometer and radiosonde are based on different tracers, either aerosols or temperature, discrepancies during day- and night-time can be expected.

Figure 10 shows the number of BLH determined simultaneously by means of radiosonde and STRAT+ during the CEILINEX campaign. Up to 300 BLH coincidences were found, between 150 and 175 during day- and night-time. The CHM (CHM100101 and CHM140110) and LD40 (LD40_002 and LD40_003) models provided larger number of BLH retrievals. The large difference within the same model occurred between the CL51RAO and the CL51CG. This could be related to the firmware changes carried out throughout the campaign. Radiosondes and STRAT+ perform well to retrieve BLH throughout day and night since same number of coincidences is obtained. BLH coincidences were lower for CS135 models. Since no relevant technical problems were reported for them, this behavior could be related either to the low signal-to-noise ratio or to the 'Vaisala-adapted' firmware version implemented in the instrument by the company, which may affect the BLH detection (pers. com., Mike Brettle).

Figure 11 shows the cases where BLH differences between radiosonde and STRAT+ are lower than 250 ($P_{<250m}$), 500 ($P_{<500m}$), 1000 meters and larger than 2000 meters during the CEILINEX campaign. $P_{<500m}$ is around 70% for all ceilometers and even 80% for CL31RAO. Figure 12 shows $P_{<250m}$ during night time larger than 80% for all ceilometers. This is due to the lower BLH dynamic range during night-time. During day-time, $P_{<500m}$ is around 60% for the CHM, CHX and CL51, and the lidar system whereas $P_{<250m}$ is below 30% for CS135 and LD40 indicating larger instability in the BLH retrieval for these two models. Therefore, differences up to 500 meters during day-time and rather 250 meters during night-time may be expected for most ceilometers.

We considered outliers as differences larger than 2km. Outliers range spans between 0 and 10%, however most of the cases are below 5%. Same model of ceilometers showed similar outliers percentages except for CHX model (10% for CHXLMU and 1% for CHX80082). Deeper analyses are

required to clarify this discrepancy. Preliminary investigation points to technical problems in the CHX80082 (Frank Wagner, pers. com.).

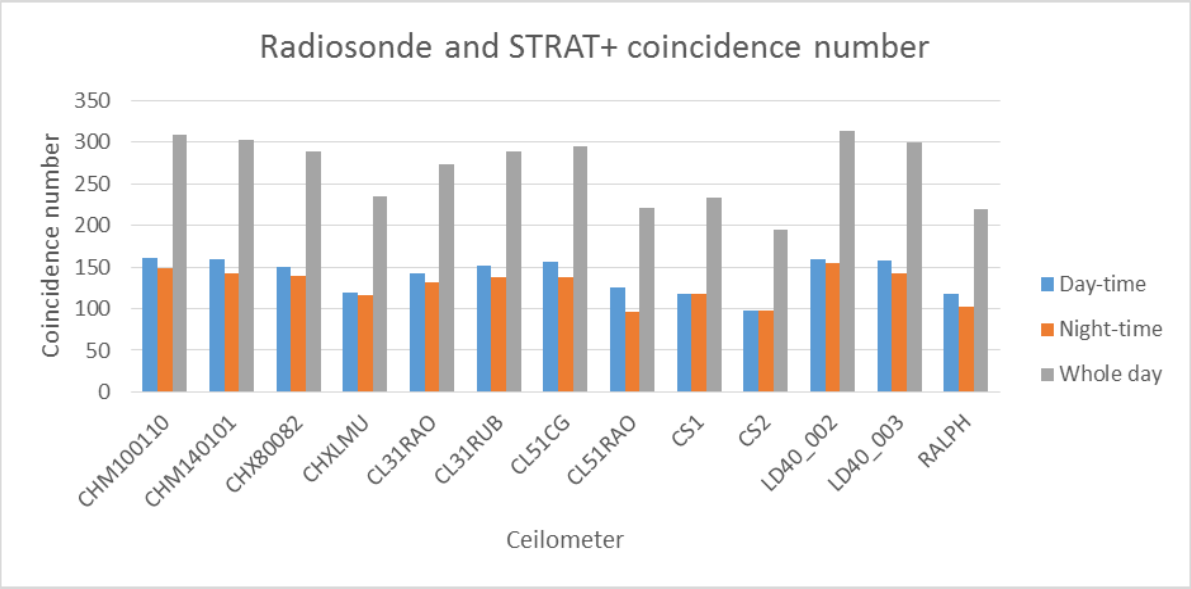

Figure 10 - Number of radiosonde and STRAT+ BLH coincidences during the CEILINEX campaign.

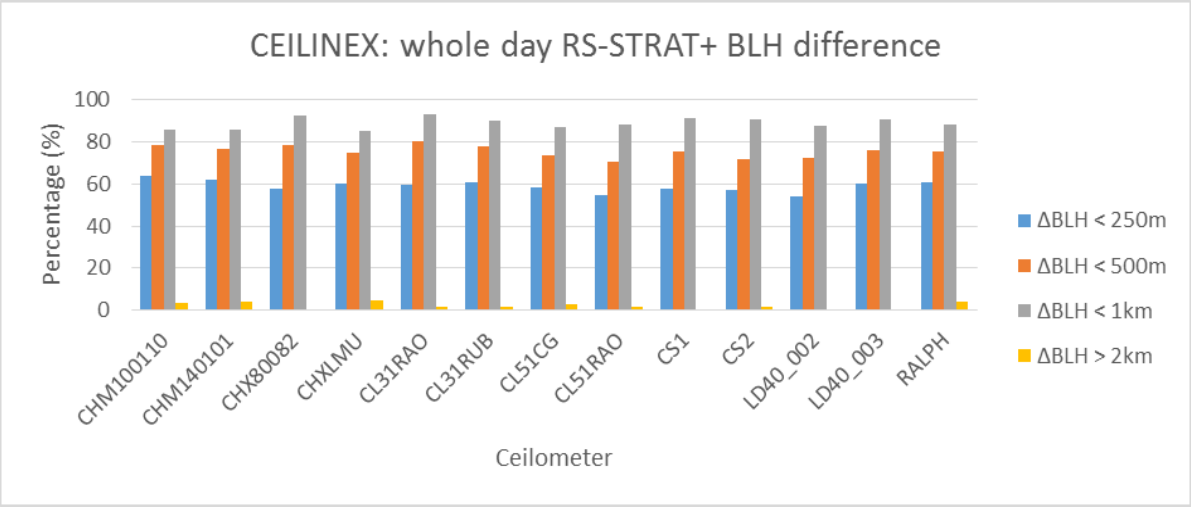

Figure 11 - Percentage of radiosonde and STRAT+ BLH differences (ΔBLH) during CEILINEX.

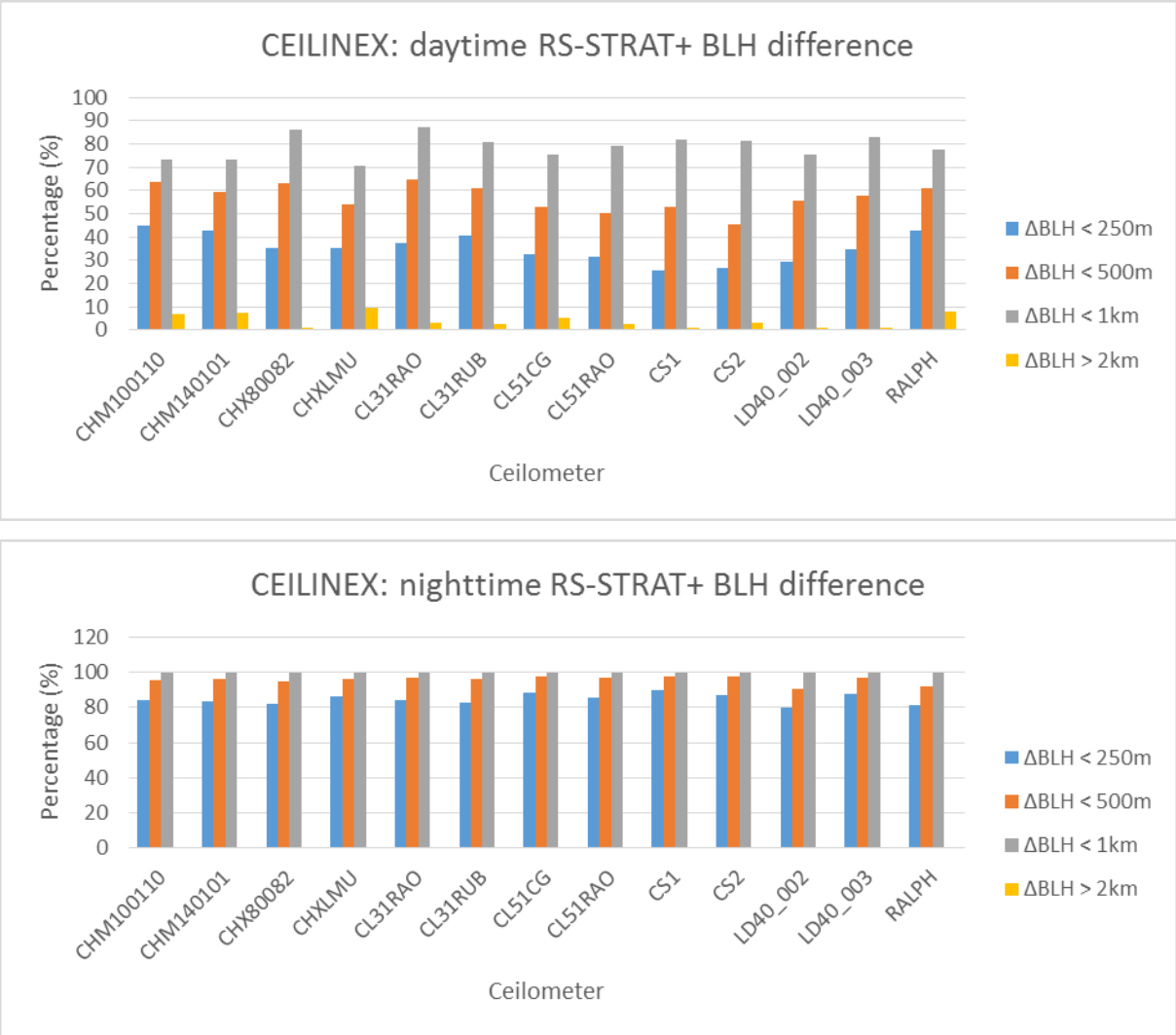

Figure 12 - Day (a) and night-time (b) percentage of radiosonde and STRAT+ BLH differences (ΔBLH) during CEILINEX campaign.

Experimental assessment of the instrument sensitivity for BLH retrieval

Table 4 and Table 5 show the BLH root-mean square deviation (RMSD) and the percentage of BLH deviation lower than 250 m, $P_{<250m}$, among the ceilometers and the lidar system during CEILINEX. Coloured columns group the ceilometers by model and the darker cells highlight the RMSD(BLH) and the $P_{<250m}$ for the same model of ceilometer (i.e., CHM, CHX, CL31, CL51, CS and LD40).

Regarding the instrument variability, all model of ceilometer show similar behavior since RMSD(BLH) is lower than 550m and $P_{<250m}$ is larger than 80% except for CHX model. In the case of CHX, $P_{<250m}$= 70% and RMSD(BLH) = 960 m, a large variability between the CHX80082 and CHXLMU is clearly highlighted. Retrievals of CHXLMU compare well with retrievals for the rest of ceilometers but a better agreement is found between CHX80082 model and radiosonde retrievals. Therefore, the

ceilometer CHX model database is under investigation to better understand whether these differences come from firmware or hardware issues.

WHOLE DAY	Root-mean square deviation (m)											
	CHM100110	CHM140101	CHX80082	CHXLMU	CL31RAO	CL31RUB	CL51CG	CL51RAO	CS1	CS2	LD40_002	LD40_003
CHM140101	390											
CHX80082	900	930										
CHXLMU	480	460	960									
CL31RAO	700	720	770	730								
CL31RUB	670	680	760	670	490							
CL51CG	590	580	910	570	640	520						
CL51RAO	670	670	950	1110	600	620	550					
CS1	980	1000	630	1020	740	780	900	1110				
CS2	980	1000	650	990	770	820	920	970	540			
LD40_002	1040	1070	480	1110	810	880	1070	1140	680	680		
LD40_003	860	880	390	910	660	730	880	960	650	670	440	
RALPH	570	540	1010	590	840	760	670	830	1080	1070	1160	950

Table 4 – BLH root-mean square deviation (m), RMSD, among the co-located ceilometers during the CEILINEX campaign. Colored columns group the same ceilometer types. Darker colored cell shows the RMSD between the same ceilometer type.

WHOLE DAY	Percentage of BLH deviation lower than 250 m (%)											
	CHM100110	CHM140101	CHX80082	CHXLMU	CL31RAO	CL31RUB	CL51CG	CL51RAO	CS1	CS2	LD40_002	LD40_003
CHM140101	90											
CHX80082	70	70										
CHXLMU	80	80	70									
CL31RAO	70	70	60	70								
CL31RUB	70	70	70	70	80							
CL51CG	80	80	70	80	70	80						
CL51RAO	70	70	60	50	80	70	80					
CS1	60	60	70	60	60	60	60	50				
CS2	60	60	70	60	60	60	60	60	80			
LD40_002	60	60	80	60	60	60	60	60	70	70		
LD40_003	70	70	80	70	70	70	70	60	70	70	80	
RALPH	80	80	60	80	60	60	70	60	60	60	60	60

Table 5 – Percentage of BLH deviation lower than 250 m among the co-located ceilometers during the CEILINEX campaign.

Comparisons performed within groups of same ceilometers model (CHM, CHX (considering CHXLMU), CL51, and CL31 and between CS and LD40) provide a better retrievals agreement (lower $P_{<250m}$ and larger RMSD values). Vaisala and Lufft ceilometers generally provide better results since CL31RAO and CHM100101 showed better agreement with radiosonde than CS and LD40. Deeper analysis (not shown) evidenced that the low signal-to-noise ratio of CS135 and LD40 makes difficult the detection of the BLH candidates. However, as it was afore mentioned, the firmware version used on CS135 was able to affect the measurements and thus, only the LD40 ceilometer can be completely discarded for BLH detection. To improve the variance retrieval, despiking techniques will be included in STRAT+ to provide better BLH retrievals.

DAY	Root-mean square deviation (m)											
	CHM100110	CHM140101	CHX80082	CHXLMU	CL31RAO	CL31RUB	CL51CG	CL51RAO	CS1	CS2	LD40_002	LD40_003
CHM140101	390											
CHX80082	900	930										
CHXLMU	480	460	960									
CL31RAO	700	720	770	730								
CL31RUB	670	680	760	670	490							
CL51CG	590	580	910	570	640	520						
CL51RAO	670	670	950	1110	600	620	550					
CS1	980	1000	630	1020	740	780	900	1110				
CS2	980	1000	650	990	770	820	920	970	540			
LD40_002	1040	1070	480	1110	810	880	1070	1140	680	680		
LD40_003	860	880	390	910	660	730	880	960	650	670	440	
RALPH	570	540	1010	590	840	760	670	830	1080	1070	1160	950
NIGHT	Root-mean square deviation (m)											
	CHM100110	CHM140101	CHX80082	CHXLMU	CL31RAO	CL31RUB	CL51CG	CL51RAO	CS1	CS2	LD40_002	LD40_003
CHM140101	70											
CHX80082	70	70										
CHXLMU	90	90	80									
CL31RAO	100	110	110	140								
CL31RUB	80	80	90	120	90							
CL51CG	70	80	70	100	110	80						
CL51RAO	110	110	110	130	80	120	100					
CS1	80	80	80	100	120	90	60	130				
CS2	100	100	100	100	130	100	100	120	90			
LD40_002	110	90	90	120	130	110	100	110	100	90		
LD40_003	80	80	70	90	120	110	80	110	80	100	80	
RALPH	90	80	80	80	130	110	80	120	80	80	70	70

Table 6 – Day- and night-time BLH root-mean square deviation (m) among the co-located ceilometers during the CEILINEX campaign.

Spatio-temporal assessment of the STRAT+ BLH candidates

The spatio-temporal assessment was carried out by comparing the STRAT+ BLH candidates with those provided by the manufacturer of the CHM15k Nimbus ceilometers (Jenoptik/Lufft). The retrievals were obtained from the three instruments deployed at Granada, Spain; Payern, Switzerland and Palaiseau, France. Table 7 shows the Lufft versus STRAT+ BLH-candidate coincidence percentage for all cases, the percentage for differences lower than 250 meters ($P_{<250m}$), lower than 500 meters ($P_{<500m}$) and lower than 1000 meters.

The agreement between the Lufft and STRAT+ BLH candidates is generally better during day-time for all cases. The CHM15k deployed at Palaiseau station shows better results during day-time ($P_{<250m}$ of 75%). During night-time, Granada shows large discrepancies with the other stations ($P_{<250m}$ = 80, 11 and 26% for Granada, Payerne and Palaiseau, respectively). Ancillary analyses (not shown) prove that it is caused by an artefact due to wrong overlap correction applied to the measurements. This artefact affects the signal between 0 and 400 meters above ground level that is even more relevant for the BLH detection during night-time. Payern station (Switzerland) has developed a technique to retrieve a better overlap correction. Further studies carried out in collaboration with them showed that the correction of the artefact can considerably improve the BLH retrievals.

Period	Station (months)	Coincidence Percentage (%)	Percentage of cases with Lufft-to-STRAT+ BLH candidate difference in absolute terms:		
			< 250 m (%)	< 500 m (%)	> 1000 m (%)
Daytime	Granada (4)	89	53	62	16
	Payerne (3)	91	48	71	4
	Palaiseau (3)	95	75	85	4
Nighttime	Granada (4)	95	88	89	3
	Payerne (3)	95	11	37	2
	Palaiseau (3)	94	26	48	12
Total	Granada (4)	92	67	74	11
	Payerne (3)	93	31	55	3
	Palaiseau (3)	95	52	69	8

Table 7 – Percentage of Lufft-to-STRAT+ BLH candidate differences in absolute terms lower than 250, 500 and larger than 1000 m.

STRAT+ and microwave radiometer BLH comparison

The boundary layer height, BLH, has been determined by means of the CHM15k Nimbus and the HATPRO microwave radiometer (MWR) co-located at SIRTA (Löhnert et al, 2012). STRAT+ and the parcel method were applied to the CHM15k and HATPRO measurements, respectively. Since the parcel method requires convective atmospheric situation, the comparison was applied only to daytime measurements. The parcel method strongly depends on the surface temperature and thus, a high uncertainty in the estimated BLH may result (for example, in situations without a pronounced inversion at the convective boundary layer top).

Station	Coincidence Percentage (%)	Percentage of number of cases with MWR-ALC BLH difference in absolute terms:		RMSD (m)
		< 250 m (%)	< 500 m (%)	
Palaiseau (1 month)	86	41	68	620

Table 8 - Percentage of MWR-to-ALC BLH differences lower than 250, 500 and larger than 1000 m in absolute terms and the root-mean squared deviation, RMSD for 1 month comparison at SIRTA station (Palaiseau, France).

The percentage for BLH differences lower than 250 and 500 meters, $P_{<250m}$ and $P_{<500m}$ and the root-mean squared deviation, RMSD, are shown in Table 8. Despite the smooth temperature profiles derived from MWR measurements and the use of different tracers (aerosols and temperature), almost 70% of the cases presented a difference below 500 m. Thus a general good agreement is obtained. However, the RMSD value is 620 meters and so, large differences are expected in some cases.

System Type	System name & Manufacturer		Firmware, Outputs& BLH support	Characteristics	Purchase Cost
Multi-wavelengths research lidars	Research Labs Gordien Strato Raymetrics	Location	User –based algorithm, output depends on manufacturers, raw data provided	Multi-wavelengths, high range and time resolution, heavy and not mobile	300-1000 k€
Micro or Mini lidars	MPL sigma space	SIRTA, Palaiseau	Manufacturer firmware. Binary or netcdf outputs, include BLH	One channel, 532nm, 50kg	80-150 k€
	CE370 Cimel	-	Manufacturer firmware based on existing open source algorithms	Different configuration one or two channels (532nm) adjustable near field performance, 30-50kg	
	CE376 Cimel	OHP	Manufacturer firmware	One or two wavelengths (532nm or 808nm), w. or wo. polarization, 30-70kg if environmental unit or not., 20-30kg	
	ALS450, Leosphere (not manufactured anymore)	SIRTA	Manufacturer firmware, binary outputs, BLH from manufacturer firmware	2 channels 355nm w. polarization, 50kg	
Ceilometers	CHM15K*# nimbus, Luft	SIRTA, Palaiseau Payerne, Switzerland Ceilinex, Munich GFAT, Granada	Manufacturer firmware. Output telegram data and netcdf format	One wavelength (1064nm), 70kg	30-40 k€
	CL51*#, Vaisala	*Humain*, Belgium Ceilinex, Munich	Telegram output, BLview includes BLH retrieval	One wavelength (910nm), 46kg	20-30 k€
	CL31*#, Vaisala	SIRTA, Palaiseau Ceilinex, Munich	Telegram output, BLview includes BLH retrieval	One wavelength (910nm), 32kg	10-20 k€
	LD40*#, Vaisala	Ceilinex, Munich	Telegram output, BLview includes BLH retrieval	One wavelength (910nm), <30kg	
	CS135*#, Campbell scientific	Ceilinex, Munich	MLH+PCviepoint, BLH from KNMI algorithm.	One wavelength (905nm), 32kg	18-20k€

Table 9 - Costs and characteristics of different automatic lidars and ceilometer systems, background colour represents system that are eyesafe (green) or not (blue). Superscript stars indicate the systems involved in CEILINEX campaign for which performance were compared in retrieving BLH using STRAT+ algorithm. Superscript sharp indicates the systems for which radiosonde BLH retrievals were compared to STRAT+ ones.

Instruments overview

Many automated light detection and ranging systems are manufactured to provide backscatter signals and boundary layer heights. Three categories are shown in Table 9, and classified on cost range criteria. Green and blue color backgrounds discriminate systems that are eye-safe or not. Non eye-safe systems require security operational procedures based on marine radar devices detection. Red background cells indicate systems for which dataset were used to evaluate performances of STRAT+ algorithm for BLH retrievals. Yellow background indicates dataset for which performances of PyBL_ICOS and STRAT+ algorithm were compared. Purple backgrounds indicate systems for which performance of STRAT+ will be checked later on.

Conclusions and perspectives

WP3 produced a unified algorithm to retrieve BLH from lidar measurements but also from ceilometers measurements. The results concerning interpolation of BLH value, exploring the case of instrument non-collocated to ICOS stations, are still not mature enough to be used. More studies should be performed in order to combine radiosonde profiles, optical measurements and model data. In an effort to test the algorithm's 'universality', 12 co-located ceilometers (of six different types) and a lidar system were compared by applying STRAT+ on measurements. The results shows that the detected-to-potential BLH percentage for all systems is between 91% and 97% of the potential 10-min BLH during the whole period of comparison (i.e. 2 months). Experimental assessment of the instrument sensitivity for BLH retrieval shows that comparisons performed within groups of same ceilometers model provide a better BLH retrievals agreement. Regarding the instrument variability, all models of ceilometer show similar behavior since the percentage of BLH deviation lower than 250 m ($P_{<250m}$) among the co-located ceilometers during the comparison is better that 80% and the BLH root-mean square deviation is lower than 550m. Comparison with collocated radiosondes shows $P_{<250m}$ during night time is larger than 80% for all ceilometers. For daytime $P_{<250m}$ drops to 25%-45% due to low signal to noise ratios.

References

U. Löhnert and O. Maier, 2012, *Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges*, Atmos. Meas. Tech., 5, 1121–1134, 2012. Doi:10.5194/amt-5-1121-2012.

S. Pal and M. Haeffelin, 2013, *Exploring a geophysical process-based attribution technique for the determination of the atmospheric boundary layer depth using aerosol lidar and near-surface meteorological measurements*. JGR, 118(16), 9277–9295. Doi: 10.1002/jgrd.50710

S. Pal and M. Haeffelin, 2015, *Forcing mechanisms governing diurnal, seasonal, and interannual variability in the boundary layer depths: Five years of continuous lidar observations over a suburban site near Paris*. JGR, in press. Doi: 10.1002/2015JD023268

TOPROF (COST Action ES1303, www.toprof.imaa.cnr.it): towards operational ground based profiling with ceilometers, doppler lidars and microwave radiometers for improving weather forecasts.

Holzworth, G. C., 1967, *Mixing depths, wind speeds and air pollution potential for selected locations in the United States*, J. Appl. Meteorol., 6, 1039–1044, doi:10.1175/1520-0450(1967)006<1039:MDWSAA>2.0.CO;2.

Kretschmer, R., Gerbig, C., Karstens, U., and Koch, F.-T., 2012, *Error characterization of CO2 vertical mixing in the atmospheric transport model WRF-VPRM*, Atmos. Chem. Phys., 12, 2441-2458, doi:10.5194/acp-12-2441-2012, 2012.

Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, A., Vogel, F., Hammer, S., and Totsche, K. U., 2014: *Impact of optimized mixing heights on simulated regional atmospheric transport of CO2*, Atmos. Chem. Phys., 14, 7149-7172, doi:10.5194/acp-14-7149-2014, 2014.