



# Error estimation for localized signal properties: application to atmospheric mixing height retrievals

G. Biavati, D. G. Feist, C. Gerbig, and R. Kretschmer

Max Planck Institute for Biogeochemistry, Jena, Germany

Correspondence to: G. Biavati (gbiavati@bgc-jena.mpg.de)

Received: 11 March 2015 – Published in Atmos. Meas. Tech. Discuss.: 19 May 2015

Revised: 21 September 2015 – Accepted: 28 September 2015 – Published: 13 October 2015

**Abstract.** The mixing height is a key parameter for many applications that relate surface–atmosphere exchange fluxes to atmospheric mixing ratios, e.g., in atmospheric transport modeling of pollutants. The mixing height can be estimated with various methods: profile measurements from radiosondes as well as remote sensing (e.g., optical backscatter measurements). For quantitative applications, it is important to estimate not only the mixing height itself but also the uncertainty associated with this estimate. However, classical error propagation typically fails on mixing height estimates that use thresholds in vertical profiles of some measured or measurement-derived quantity. Therefore, we propose a method to estimate the uncertainty of an estimation of the mixing height. The uncertainty we calculate is related not to the physics of the boundary layer (e.g., entrainment zone thickness) but to the quality of the analyzed signals. The method relies on the concept of statistical confidence and on the knowledge of the measurement errors. It can also be applied to problems outside atmospheric mixing height retrievals where properties have to be assigned to a specific position, e.g., the location of a local extreme.

localization of a local property. Examples of local properties for a signal are maximum and minimum values. A more general example can be seen as the property to have a certain value or threshold. This is also the case for the location of mixing height (MH), which can be defined by local properties of the data used for its estimation.

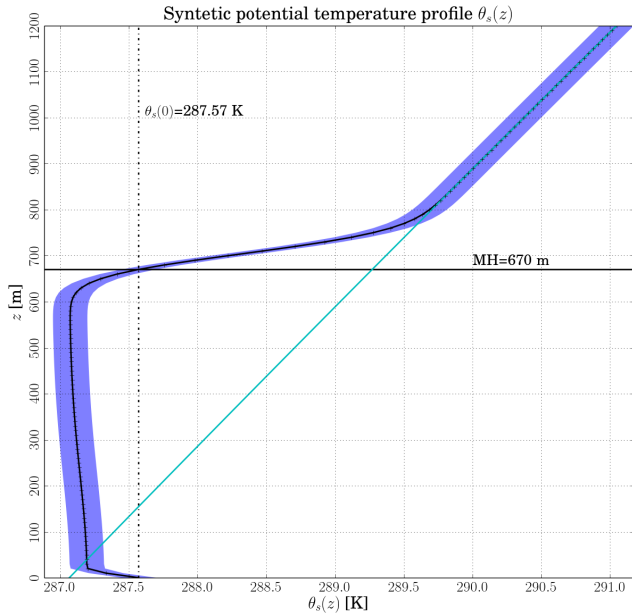
The top of the mixed layer or MH is the thickness of the layer adjacent to the ground where any pollutants or constituent emitted within it or entrained from above will be vertically mixed by convection or mechanical turbulence on a reasonably short timescale. This timescale is about 1 h or less according to Seibert et al. (1998).

The mixed layer is a sublayer of the planetary boundary layer (PBL), which is the atmospheric layer that is closest to the ground. In the PBL, several processes control exchange of energy, water and pollutants between the surface and the free atmosphere. The structure of the PBL is variable as detailed by Stull (1988).

The knowledge of MH has been considered fundamental for modeling dispersion of pollution since Holzworth (1964). This is because it defines the volume where ground fluxes are diluted. In more recent times, a strong effort has gone toward the determination of fluxes of greenhouse gases from atmospheric mixing ratio measurements (Gurney et al., 2002; Rödenbeck et al., 2003; Peters et al., 2007). The impact of model errors on MH estimations is considered one of the primary sources of uncertainties in the inverse estimates of regional CO<sub>2</sub> surface–atmosphere fluxes (Gerbig et al., 2008; Kretschmer et al., 2012). The uncertainties or localization error estimated on the MH retrieved from atmospheric profiles can be used as a valuable tool to reduce the model uncertainties as demonstrated in Kretschmer et al. (2014).

## 1 Introduction

In good scientific practice, uncertainties or errors must be provided for all physical quantities which are measured or estimated. Unfortunately, for a wide class of estimations it is not straightforward to apply standard error propagation on the result. This is the case for many applications where thresholds have to be identified in noisy signals. The aim of this work is to provide a rigorous way to estimate uncertainties for this class of operations. This is the general case of the



**Figure 1.** Idealized profile of potential temperature  $\theta_s(z)$ . The tiny black line represents the profile, and the blue area around it represents a hypothetical measurement error of  $\pm 0.125$  K. The cyan line is the potential temperature of standard atmosphere. The horizontal black line represents the MH, and the vertical dotted line the temperature at the ground. The spatial resolution used in this plot is 10 m.

This paper introduces a rigorous method to derive uncertainties in the localization of a property, with a special focus on mixing height retrievals as an example. This allows for more quantitative assessments of the quality of retrievals and can provide useful information especially when comparing different observation-based retrieval methods with one another or with mixing heights diagnosed in weather prediction models. Note that this proposed method does not address uncertainties related to assumptions about the boundary layer physics, or related to its spatial and temporal variability.

In Sect. 2 we introduce the implementation of the parcel method in the form of an algorithm as it is used in this work. The mainly theoretical part introducing the error retrieval method is described in Sect. 3. An application to other MH retrieval methods, in addition to the parcel method, is presented in Sect. 4.

## 2 Localizing the mixing height

Several methods for detecting MH are reported in the literature, depending on meteorological conditions and instrumentation used (Seibert et al., 2000). However, in order to explain our methodology to estimate uncertainties, we use the parcel method as proposed by Holzworth (1964) for its simplicity. According to Holzworth (1964), vigorous vertical mixing is driven by thermal convection. The parcel method was defined by Holzworth (1964, Sect. 2) for maximum mixing depths as

“maximum mixing depths were estimated by extending a dry adiabat from the maximum surface temperature to its intersection with the most recently observed temperature profile.”. Figure 1 provides a clear example of this method.

The driving idea is that warmer air in contact with the ground reaches an altitude where a capping inversion is located. For practical use in convective conditions – when the impact from wind shear can be neglected – the MH is located at the altitude  $h$  where the virtual potential temperature  $\theta_v(h) = \theta_{vh}$  as defined in Eq. (1) is equal to the virtual potential temperature  $\theta_v(0) = \theta_{v0}$  at the surface.

$$\theta_v(z) = T_e(z) \left[ \frac{P_0}{P(z)} \right]^\gamma \approx (T(z) + bMR) \left[ \frac{P_0}{P(z)} \right]^\gamma, \quad (1)$$

where  $T_e$  is the equivalent temperature or the temperature that the air parcel would have if all the water vapor condensed, releasing its latent heat. MR is the mass mixing ratio of water vapor,  $P(z)$  is the pressure at altitude  $z$ ,  $P_0$  is a reference pressure, and  $b$  represents the ratio of latent heat of vaporization to the specific heat of dry air at constant pressure. Taking  $P_0 = P(0)$  then results in  $\theta_v(0) = T_e(0)$ .

We chose this methodology to estimate MH because it uses a smaller number of environmental profiles than the bulk Richardson number method. In our numerical examples, we do not consider humidity, so we will focus on potential temperature  $\theta(z)$ , which can be obtained from Eq. (1) by setting  $MR = 0$ .

Real examples of vertical profiles of virtual potential temperature are presented in Sect. 4. Here instead we present a synthetically generated profile in Fig. 1.

We found convenient the use of an analytical function to describe the potential temperature profile, because in this way we can control the more relevant aspects of the profile, which are the excess temperature at the ground and the uniformity of potential temperature within the mixed layer. The use of an analytical function helps also to study effects of spatial resolution and smoothing. The profile  $\theta_s(z)$  presented in Fig. 1 can be seen as an almost neutral profile of potential temperature. The black solid curve represents the ideal signal, and the blue area around the signal represents the  $\pm 1\sigma$  error. The error and the excess temperature at the ground are chosen together for explanatory purposes and are not directly related to the physics of the mixed layer.

Looking at Fig. 1, we can see how the parcel method works. The MH is located at the altitude  $h$  where the potential temperature  $\theta_v(h)$  equals the ground potential temperature  $\theta_v(0)$ . The idealized values depicted by the black curve are affected by measurement errors. The uncertainty is represented by a blue region around the signal. Assuming normal errors, the amplitude of the blue region at a fixed altitude represents the area where we expect a probability of 68 % to measure  $\theta_v(h)$ . This implies that, when we attempt to estimate MH on a noisy signal, we will detect an altitude around the region of interest and not the exact location.

**Table 1.** Vaisala RS92-SGP (Vaisala, 2015) sensors specifications in range pressure of 1080–100 hPa.

RS92-SGP		
Quantity	Range	Error
Pressure	1080 to 100 hPa	0.2 hPa
Temperature	−90 to +60 °C	0.1 °C
Relative humidity	0 to 100 %	2.5 %
Wind speed		0.15 m s <sup>−1</sup>
Wind direction		2 °

Under steady conditions, we would get an estimated MH and a properly estimated uncertainty by repeating the measurements many times. However, in the real world, the conditions are typically not steady and the measurements cannot be repeated often enough (if at all) to obtain a statistically consistent set of estimates. Therefore, a methodology is needed that retrieves the localization error from a single profile. Our methodology requires the knowledge of the errors of the measured profiles, so that it is possible to propagate it onto the signals we want to analyze. The error propagation on potential temperature and on the bulk Richardson number profiles are provided in the Appendix.

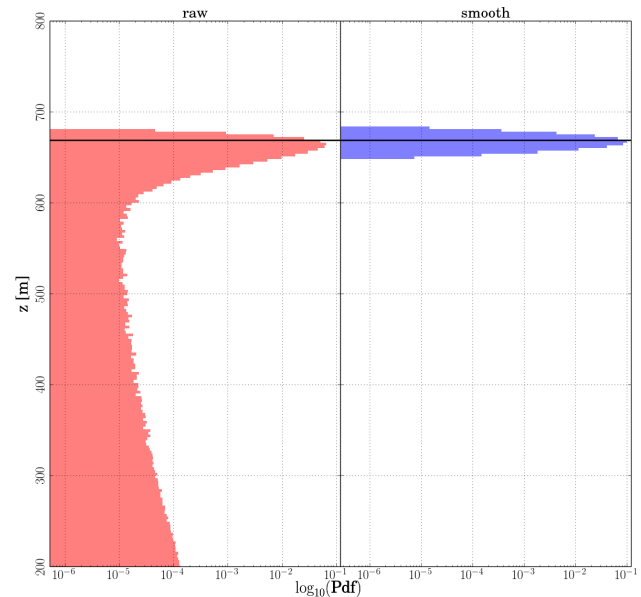
The meteorological quantities observed by radiosondes are pressure, temperature, relative humidity, wind speed, and wind direction. The data for the practical examples used in this work are part of the data set of radiosonde data of the Lindenberg Meteorological Observatory in Germany (WMO station 10393). The data are collected regularly every 6 h. The measurements are extensively described in Beyrich and Leps (2012). The model of radiosondes used is the Vaisala RS92-SGP (Vaisala, 2015). The technical specifications of the measurements are described in Table 1.

In practice, the method that is used more widely to produce estimates of MH is the so-called bulk Richardson number method. In Sect. 4 we apply our method to assess the localization error of two variants of the bulk Richardson number method described by Vogelezang and Holtslag (1996).

### 2.1 Defining an algorithm for the parcel method

After the choice of a methodology to detect MH on meteorological profiles, we have many options for implementing it as an algorithm. Again, the parcel method defines the MH at the altitude where the virtual potential temperature equals  $\theta_v(0)$ . From an operational point of view, the parcel method can be seen in many different ways.

From an abstract point of view – not related to the actual meteorological concept – the core of the method is detecting the location where a certain threshold value is reached. This is a very common task in signal analysis, commonly called threshold detection. To implement a threshold detection, one must consider different properties of the signal. The signal



**Figure 2.** Results of a Monte Carlo for the parcel method as detailed in algorithms presented in Sect. 2.1. The first panel presents the probability density function of the Monte Carlo using the data directly. In the second panel the results are for data smoothed with a window of three points instead. The algorithms were applied to the synthetic profile of Fig. 1 after application of random noise. The number of runs performed is 1 000 000. The probability density is presented on a logarithmic scale to allow the view of results far from the expected MH. The expected MH is presented as a black horizontal line.

noise is the main source of erroneous and multiple detections, especially for non-monotonic signals.

As an algorithm for applying the parcel method, we decided to use the location of the last data point (starting from the bottom) that is still smaller than  $\theta_v(0)$ . We think that this is the closest way to apply the method as described by Holzworth (1964).

From the more physical point of view, the parcel method can be implemented as the simple parcel method introduced by Holzworth (1964) or by considering an excess temperature at the ground as calculated by Troen and Mahrt (1986). One advantage of using a synthetic profile is that we control the excess temperature, so that we do not need to estimate it.

Referring to the synthetic profile of Fig. 1, we can apply the algorithm and evaluate the performances and see the probability distribution of the results. So we created  $10^6$  profiles, applying to the synthetic profile Gaussian random noise with a standard deviation (SD) of 0.125 K. In Fig. 2, we see how the estimated MHs are distributed. Where  $\theta_v(h)$  is close to  $\theta_v(0)$ , the algorithm has a high probability of retrieving results. In order to reduce the chance of retrieving MH very close to the ground in conditions closed to neutrality, we added a constraint to the algorithm in this example: only consider data above 200 m. In general Holzworth (1964)

suggests using the parcel method in the case of vigorous convection. Instead, the example introduced here presents weak convection or almost neutral conditions. We decided to use this as an example for better illustrating the uncertainty in the localization, i.e., in the determination of the MH. However, when applying the algorithms on smoothed data with a three-point window, the mode of the distribution of the results is closer to the expected MH (670 m). The smoothed profiles had reduced noise, which reduced the probability of a false detection.

We must point out that the parcel method as it is implemented can be considered just an algorithm for threshold detection in a signal. So all the considerations that we made could be applied to other methods, for example the bulk Richardson number explained in Sect. 4.

### 3 Calculating the localization error

So far, we have used a simple Monte Carlo (MC) simulation to illustrate the impact of measurement noise on the error in the retrieved MH. However, for application to large data sets this is too expensive to perform, and a more analytical method is needed.

In a continuous signal, a property can be defined as local when it occurs in an arbitrarily small neighborhood of points. However, real signals are not continuous but rather discrete data series of ordered points. For such discrete data series, the neighborhood concept must be adapted since it is not possible to consider arbitrarily small neighborhoods. Instead, a neighborhood would be a set of contiguous points. It contains a reference data point and some other points in its vicinity.

Two measurements can be considered equivalent when their difference is smaller than their errors. The degree of equivalence is commonly called confidence. Confidence is rigorously defined in several textbooks. It is used to verify a hypothesis or, in other words, to see if an estimated value agrees with a theoretical expectation. The most general case is presented in Eq. (2), where the concept is used to check if two estimated values can refer to the same quantity.

A local property on an ordered data series can be shared between data points. This is due to the fact that data have errors, which has the consequence that different data values at different points can be differentiated from each other only within a certain degree of confidence. This sharing of properties by contiguous data points is the key to defining a rigorous concept of localization error.

To give an example, a data point located at 400 m (see Fig. 1) is located in a region of uniform and constant  $\theta_v$ . For such a point, a local property is almost impossible to define, given the uncertainty of the data. On the other hand, at 700 m altitude the difference of consecutive values is such that the localization can be easily performed.

The formal description of the method requires the introduction of some symbols. An ordered data series  $y_i$  where  $i \in \{1, \dots, N\}$  is associated with a series of locations  $x_i$  and with a series of errors  $\epsilon_{y_i}$ .

When a local property in a signal can be defined, there are two choices to define its location: the local property can be located exactly at a data point or between two data points. The second possibility will not be discussed. Instead, for simplicity, we assume that the localization is located at the first data point that defines the interval where the property is detected.

The general assumption of the method is that the measurement errors are known, and they are normally distributed and uncorrelated. We focus on data points that have neighbors on both sides – not the end points of a series.

### 3.1 Definitions

The method relies on one main idea: (a) the results of an algorithm are expected to fall in a neighborhood of the true location, and (b) this neighborhood can be seen as a set of data points that have similar values within the errors of the measurements. The similarity of values is measured with the quantity commonly called confidence.

#### 3.1.1 Confidence

The confidence  $\zeta$  of two values  $y_i$  and  $y_j$  with errors  $\epsilon_{y_i}$  and  $\epsilon_{y_j}$ , respectively. It is expressed as

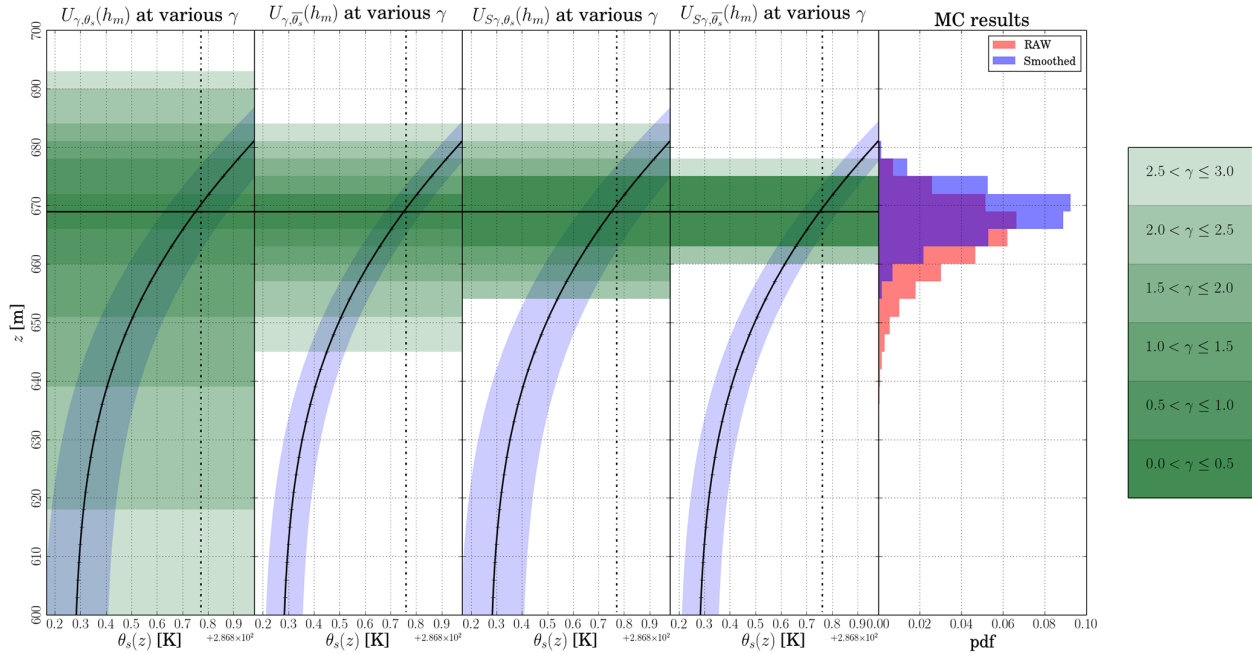
$$\zeta(f_i, f_j) = \frac{|y_i - y_j|}{\sqrt{\epsilon_{y_i}^2 + \epsilon_{y_j}^2}}. \quad (2)$$

Equation (2) is a modification of Welch's t test (Welch, 1947; their Eq. 25). Instead of the errors of the estimated mean used by Welch (1947), we used the measurement errors.

Welch's t test and other similar tests are typically used to evaluate hypotheses. In this particular case, we try to verify the null hypothesis that two estimations  $y_i$  and  $y_j$  are equal by taking their difference.

For a normal distribution, confidence intervals are typically defined as a distance in units of the SD  $\sigma$ : 68.27 % for  $\pm 1\sigma$ , 95.45 %  $\pm 2\sigma$ , and 99.73 % for  $\pm 3\sigma$ . If  $y_i$  and  $y_j$  are normally distributed, the denominator of Eq. (2) is equivalent to the SD of the distribution  $y_i - y_j$ . The function  $\zeta(y_i, y_j)$  can then be interpreted as an absolute distance in units of  $\sqrt{\epsilon_{y_i}^2 + \epsilon_{y_j}^2}$ . This provides a natural scale for accepting or rejecting the hypothesis:

1. optimal confidence:  $\zeta(y_i, y_j) \leq 1$ ;
2. good confidence:  $1 < \zeta(y_i, y_j) \leq 2$ ;
3. acceptable confidence:  $2 < \zeta(y_i, y_j) \leq 3$ ;
4. bad confidence:  $\zeta(y_i, y_j) > 3$ .



**Figure 3.** All plots are obtained starting from the synthetic profile  $\theta_s(z)$  introduced in Fig. 1 and calculated with spatial resolution of 3 m. The first panel presents confidence neighborhoods as from Eq. (4) for various thresholds  $\gamma$ ; second panel presents the confidence neighborhoods for the smoothed profile  $\bar{\theta}_s(z)$ ; third panel: strict confidence neighborhoods as from Eq. (5); fourth panel: strict confidence neighborhoods as from Eq. (5) for the smoothed profile  $\bar{\theta}_s(z)$ ; fifth panel: probability density function of the Monte Carlo results for the algorithm introduced in Sect. 2.1 for raw data (red) and smoothed (blue). The confidence neighborhoods are depicted as green areas, following the color scale on the right. The horizontal black line represent the location on the MH =  $h_m = 669$  m for this spatial resolution.

### 3.1.2 Discrete neighborhoods

Given a series of locations  $x_i$  where  $i \in \{1, \dots, N\}$   $N \in \mathbb{N}$ , a discrete neighborhood or simply a neighborhood of a point  $x_m$  of the series is defined as the set  $V_{x_m}$  that respects the following relation:

$$V_{x_m} = \{x_{m-l_1}, \dots, x_{m+l_2} : m-l_1 \geq 1, m+l_2 \leq N \in \mathbb{N}\}. \quad (3)$$

This reflects the idea that the neighborhood  $V_{x_m}$  of a point  $x_m$  extends from a point  $x_{m-l_1}$  to the left of  $x_m$  to a point  $x_{m+l_2}$  to the right of  $x_m$ . The neighborhood must include at least one point in each direction, so  $l_1, l_2 \geq 1$ .

To estimate a local property in an ordered data series  $y_i$ , we consider the value  $y_m$  at a specific data point  $x_m$  and the values at data points in a neighborhood around the specific point  $\{y_{m-l_1}, \dots, y_{m+l_2}\}$ .

### 3.1.3 Confidence neighborhood

By merging the concept of the discrete neighborhood expressed in Eq. (3) with the one of confidence expressed in Eq. (2), we can define the key tool for our methodology: the confidence neighborhood.

To refer to confidence neighborhoods, we use the following notation:  $U_{\gamma,y}(x_m)$  is the confidence neighborhood of the data point located at  $x_m$ , with the respective data series  $y_i$  of the property  $y$  and the confidence threshold  $\gamma$ .

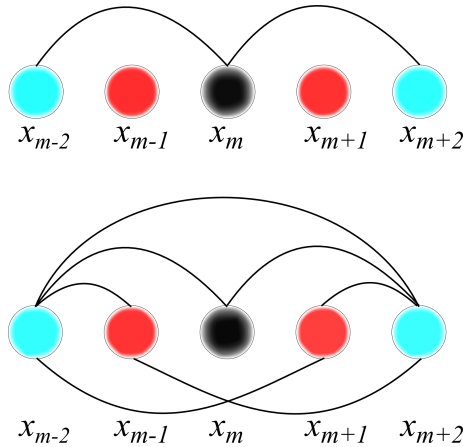
We take a monotonic series of locations  $x_i \in \mathbb{R}$  with an associated ordered data series of values  $y_i \in \mathbb{R}$  and corresponding errors  $\epsilon_{y_i}$  with  $i \in \{1, \dots, N\}$ . Given a real constant  $\gamma > 0$ , we define the confidence neighborhood of  $x_m$  as the neighborhood of data points  $U_{\gamma,y}(x_m)$  that respects the following relation:

$$U_{\gamma,y}(x_m) = \{x_{m-1}, x_m, x_{m+1}\} \cup \{x_j : \zeta(y_m, y_j) \leq \gamma\}, \quad (4)$$

where confidence  $\zeta$  is defined by Eq. (2).  $U_{\gamma}(x_m)$  is the contiguous set of locations surrounding the point  $x_m$  that share the relation of confidence to the specific point  $y_m$  with respect to the constant  $\gamma$ . The confidence neighborhood is defined as a neighborhood (Eq. 3). So a confidence neighborhood must respect both Eq. (3) and Eq. (4). In this way contiguity and confidence relation are achieved together.

Referring to the test function presented in Fig. 1, we can estimate  $U_{\gamma,\theta_v(h_i)}(h_m)$ . According to the algorithms defined in Sect. 2.1, the MH is  $h_m = 670$  m. In the first panel of Fig. 3,  $U_{3,\theta_v(h_i)}(h_m)$  extends downward to an altitude of 0 m above ground. The confidence intervals are larger for larger values of  $\gamma$ . From the example, we can also see that they are not symmetric with respect to  $h_m$ . This asymmetric behavior arises from the fact that  $\theta_v(h)$  is nonlinear.

In Fig. 3, the different amplitudes of the confidence neighborhoods reflect the idea that the algorithm could provide



**Figure 4.** This diagram shows the difference between confidence neighborhood (above) and strict confidence neighborhood (below). The lines connecting the points represent where the relation of confidence Eq. (2) must be met for a fixed  $\gamma$ . The red points are included within the confidence neighborhood of every kind even if they do not respect the relation of confidence with  $x_m$ .

MH at any altitude below the true one. This is expected when considering the probability density function for the raw data in Fig. 2.

This definition of confidence neighborhood is more than a mathematical abstraction. From the physical point of view it reflects the idea of probability to obtain an estimation of a local property starting from a signal which has its own uncertainties. Qualitatively, some properties of the distributions of results can also be inferred. In particular, the skewness or asymmetry of the distributions is captured by differences between the leftward  $l_1$  and rightward  $l_2$ .

Smoothing the data with a window of three points produces a profile whose error is  $\sigma_{\overline{\theta_s(z)}} \approx 0.072\text{k}$ . It is  $1/\sqrt{3}$  times the error of the original profile  $\sigma_{\theta_s(z)} = 0.125\text{k}$ . The effect of noise reduction by smoothing can be clearly seen in the first and second panel of Fig. 3. In this example, the skewness of the distribution of the results can still be seen from the confidence neighborhoods at various values of  $\gamma$ . It is also clear that in signals with smaller errors the localization of a property is more precise.

### 3.1.4 Strict confidence neighborhood

When an algorithm defines a location  $x_m$ , we already know that other nearby points could have been chosen if the random noise had manifested differently. The points that have higher probability to be chosen would all share the property that caused the choice. In the given definition of confidence neighborhood, we looked only at the points that respect the equivalence relation Eq. (2) with the single data point  $y_m$ . Because of this simple definition, the confidence neighborhood in the first panel of Fig. 3 tends to grow also to points with a rather low probability – at least for higher values of

$\gamma$ . However, if we assume that the selected location  $x_m$  has a confidence neighborhood in which the true value is located, in this neighborhood all the points can be considered candidates. This leads to the definition of strict confidence neighborhood: a neighborhood where all points must respect the equivalence relation Eq. (2) between each other (not only with  $y_m$ ). The strict confidence neighborhood  $U_{S\gamma,y}(x_m)$  is the neighborhood of  $x_m$  that satisfies Eq. (2) for all of its points:

$$U_{S\gamma,y}(x_m) = \{x_{m-1}, x_m, x_{m+1}\} \cup \{x_i : \zeta(y_i, y_j) \leq \gamma \forall i, j \in [m-l_1, \dots, m+l_2]\}. \quad (5)$$

According to the definition in Eq. (4), the confidence neighborhood has to include the point itself as well as its direct neighbors. Therefore,  $l_1$  and  $l_2$  have to be greater or equal to 1 – even if there is no confidence between  $y_{m-1}$ ,  $y_m$ , and  $y_{m+1}$ .

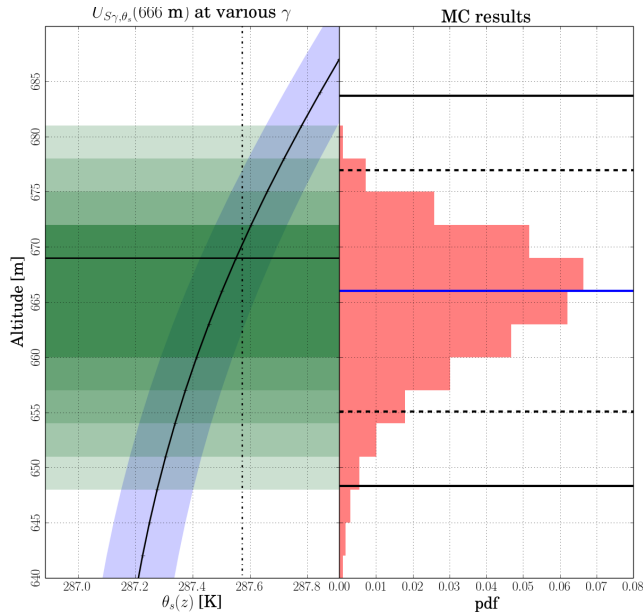
Comparing the confidence neighborhoods and the strict confidence neighborhoods in Fig. 3,  $U_{S\gamma,y}(h_m)$  is much more symmetric than  $U_{\gamma,y}(h_m)$  also for larger values of  $\gamma$ . This is due to the stronger constraint for the strict confidence neighborhood, as can be seen in Fig. 4.

### 3.1.5 Practical determination of the strict confidence neighborhood

Despite the definition of confidence neighborhood, the strict confidence neighborhood is not always unique. This can be understood by examining the process that is used to estimate it.

We start from the first three points  $x_{m-1}$ ,  $x_m$ , and  $x_{m+1}$ . If they do not match the confidence relation for the chosen  $\gamma$ , the strict confidence neighborhood contains only these three points. Otherwise, if they do match the confidence relation, an iterative process can be used to see if other contiguous points can be added to the neighborhood. This can be achieved by checking for couples of points directly left and right of the borders of the neighborhood. These are checked to see if they match the confidence relation with the neighborhood. Each element of the couple can agree with the already-defined confidence neighborhood. However, it can also be the case that the two points do not agree with each other.

When checking for couples of points that agree with all the other points, but not with each other, a choice must be made and one of the two points must be rejected. Rejecting a point means that the confidence neighborhood stops growing in that direction. For the other direction, one can still add points as long as the confidence relation remains satisfied. A practical way to determine which one of two conflicting data points should be kept is to select the one that has a better confidence with  $y_m$ . This is the criterion that we adopted.



**Figure 5.** Left panel: the strict confidence neighborhoods  $U_{S\gamma\theta_s}(x_m)$  and the true MH as a horizontal black line. For color scales for different  $\gamma$  values see Fig. 3. Right panel: the distribution of the Monte Carlo output as a probability density function (pdf). The blue line is the median of the distribution, the dashed lines define the localization error of the median as from Eq. (7), and the solid lines represent the square root of the second-order moment about the median of the distribution.

### 3.2 From confidence neighborhood to localization error

The width of the confidence neighborhood is a measure for the quality of a localization. In particular, the introduced left width  $l_1$  and the right width  $l_2$  give quantitative information about the range of possible results of an algorithm. Moreover,  $l_1$  and  $l_2$  provide a qualitative estimation of the skewness of the distribution of possible results.

In Fig. 3, we showed the confidence neighborhood for the true MH. That was possible because we know the true MH for the synthetic profile that we used. Now we want to focus on the confidence neighborhood of the results retrieved from a localization algorithm working on noisy data. We expect that the result should fall into the confidence neighborhood of the true MH.

One way to measure localization error would be to use the SD of the output distribution of a Monte Carlo. However, the Monte Carlo distribution overlaps well with the confidence neighborhoods for different values of  $\gamma$  in Fig. 3. Because the output distributions are not necessarily normal, the mean result of the Monte Carlo is not the most likely one. By definition, the mode (maximum value) of the distribution has the maximum probability.

As a measure for uncertainty we used the square root of the second moment about the mode. For normal distributions

this is also called SD. So we can calculate the square root of the second-order moment about the mode of the distribution of the results.

#### 3.2.1 Localization error

The second-order moment expresses clearly the uncertainty of the localization. However, its calculation requires performing a Monte Carlo experiment. Moreover, it depends strongly on the algorithm used. If the data have reasonably small errors and the algorithms provide a useful estimate of the target quantity, the results will have good confidence with respect to the true target quantity.

The confidence neighborhoods as from Eq. (4) for  $\gamma \approx 2$  and the second-order moment of the distribution correspond at least qualitatively. Therefore, we could use the second-order moment about a retrieved location  $x_m$  to give a first definition of the localization error  $\sigma_{x_m}$ :

$$\sigma_{x_m} = \sqrt{\frac{1}{l_1 + l_2} \sum_{x_k \in U_{\gamma,y}(x_m)} (x_m - x_k)^2}, \quad (6)$$

where  $l_1$  and  $l_2$  are respectively the left and right extensions of the confidence neighborhood  $U_{\gamma,y}(x_m)$  defined in Eq. (4).

We found that Eq. (6) is a good choice for the definition of localization error because for uniformly sampled data series, in the case of the smallest confidence neighborhood  $U_{\gamma,y} = (x_{m-1}, x_m, x_{m+1})$ , the localization error is just the amplitude of the sampling, which in common practice is often used as uncertainty.

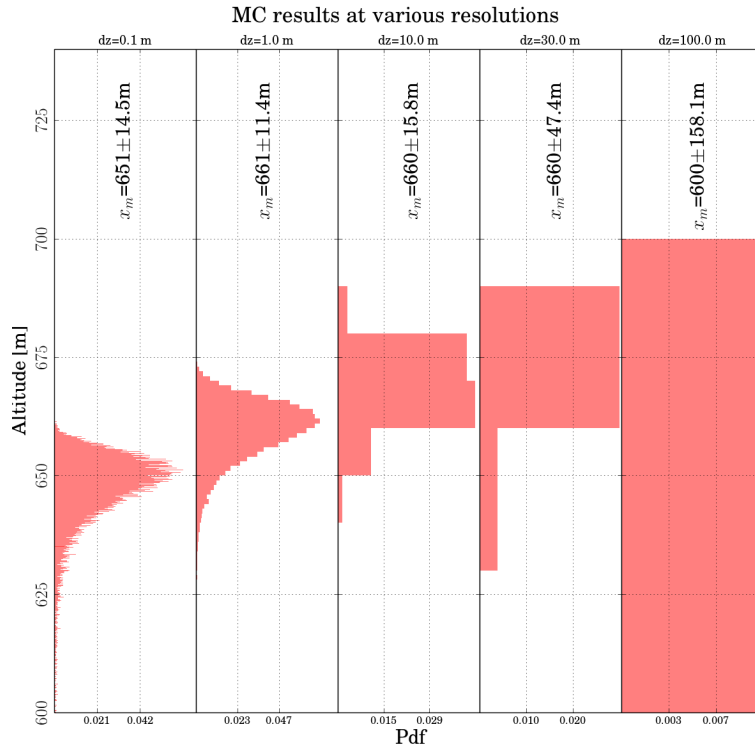
This definition of  $\sigma_{x_m}$  is subject to the choice of the confidence threshold  $\gamma$ . From studying many cases of Monte Carlo results, we found that for practical purposes the use of  $\gamma = 2$  will provide reasonable uncertainties (Kretschmer et al., 2014). For arguments supporting the use of  $\gamma = 2$  see the Supplement. Note that here we use the strict confidence neighborhood for the localization error, which is less subject to the choice of  $\gamma$  (Fig. 3).

For our final definition of localization error, we take Eq. (6) and substitute  $U_{\gamma,y}(x_m)$  with the strict definition  $U_{S\gamma,y}(x_m)$ . Then we consider which value of  $\gamma$  is most representative of the distribution of the results. Clearly  $\gamma = 3$  captures most of the results as we can see in Fig. 5. This is true for the example: in particular, 97 % of the results are captured.

With  $\gamma = 3$  and the strict confidence neighborhood our final definition of the localization error  $\sigma_{x_m}$  is

$$\sigma_{x_m} = \sqrt{\frac{1}{l_1 + l_2} \sum_{x_k \in U_{S3,y}(x_m)} (x_m - x_k)^2}. \quad (7)$$

The strict confidence neighborhood  $U_{S\gamma,y}(x_m)$  is in general more symmetric than the simple confidence neighborhood. However, it can still be unbalanced in the case of very



**Figure 6.** Probability density functions of five MCs performed 100 000 times each. From left to right we created  $\theta_s(z)$  with the following resolutions:  $dz = 0.1, 1, 10, 30,$  and  $100$  m. On each plot, the median of the distribution  $x_m$  and its localization error as defined in Eq. (7) are printed.

poor localization, for example when the property is located in a region where the signal is uniform within the uncertainties.

### 3.2.2 Symmetry of the localization error

In general, a confidence neighborhood can be asymmetric if the signal does not depend on the location in a linear way. Especially if there is no change in signal towards one side of  $x_m$ , the confidence neighborhood will extend into that region. Instead, a confidence neighborhood estimated on a linear trend extends almost equally into both directions.

The localization error as defined in Eq. (7) has two distinct contributions:  $e_1$  from the points before  $x_m$  and  $e_2$  from the points after.

$$\sigma_{x_m}^2 = \frac{1}{l_1 + l_2} \sum_{x_k \in [x_{m-l_1}, \dots, x_m]} (x_m - x_k)^2 + \frac{1}{l_1 + l_2} \sum_{x_k \in [x_m, \dots, x_{m+l_2}]} (x_m - x_k)^2 = e_1^2 + e_2^2 \quad (8)$$

In the example shown in Fig. 5, the localization error  $\sigma_{x_m}$  is 10.9 m, while the contributions from either side are  $e_1 = 8.6$  m and  $e_2 = 6.7$  m.

### 3.3 Localization depending on multiple data series

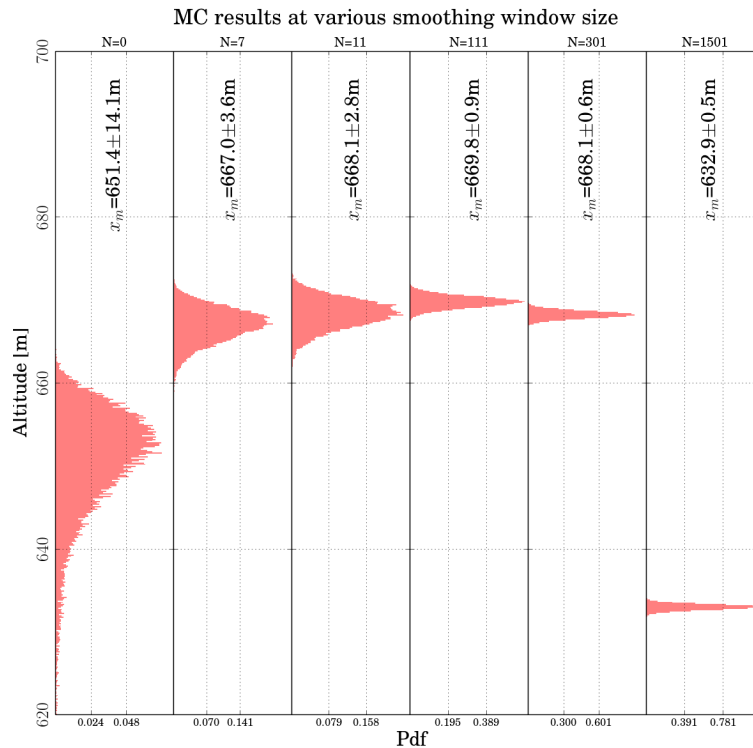
Often a local property can be defined as a location where more than one condition must be fulfilled. The localization might depend on different data series defined in the same series of locations  $x_i$ . In this case, we define the confidence neighborhood as the intersection of the respective confidence neighborhoods of each data series. So in the end, the confidence neighborhood for multiple data series is identical to the smallest confidence neighborhood of  $x_m$  for these data series. This definition also holds for strict confidence neighborhoods.

### 3.4 Effects of signal resolution and smoothing

The location vector  $x_i$  can be unevenly spaced. In the previous examples, we used a fixed resolution of 3 m. However, the resolution of a radiosonde profile usually varies with altitude. To assess how the resolution affects the localization error, MC simulations were performed on signals at various spatial resolutions  $dz$ .

The results clearly show that the resolution has an impact (Fig. 6). In particular, for higher resolutions the chosen algorithm underestimates MH. The skewness of the distribution is kept for different resolutions. Reducing the resolution increases the localization error. This is true for all resolutions





**Figure 7.** Probability density functions of six MCs performed 100 000 times each. For this example we used a fixed spatial resolution of 0.1 m. The different smoothing window with size  $N$  is indicated at the top of each panel, and the median of the distribution  $x_m$  and its localization error as defined in Eq. (7) are printed.

starting from 1 to 100 m. The localization error for 0.1 m is larger because the mode in this case is located closer to the region of an almost constant signal.

Note that in Sect. 2.1 we decided to use the first location  $x_i$  that satisfies the algorithm, not an intermediate point. Therefore, for the coarse resolution of 100 m most of the probable MHs are located within 600 to 700 m, although  $x_m = 600$  m was chosen.

From the example of the rightmost panel of Fig. 3, we have already seen some of the effect of smoothing. It drastically reduces the size of the confidence neighborhoods and the spread of the possible results. This is because smoothed profiles have smaller errors as can be easily inferred from standard error propagation (Eq. A1). If the signal has an uncorrelated and constant error  $\sigma_{y_0}$  and the running average is performed on  $N$  samples, the resulting error for the smoothed signal is  $\sigma_{y_0}/\sqrt{N}$ .

There are two main effects when running averages are used before an algorithm is applied:

- the increased window size reduces the localization error,
- the median of the results changes.

Both effects can be seen in Fig. 7. To better appreciate the effects of smoothing, we have chosen a very high resolution synthetic profile with  $d_z = 0.1$  m. It is clear that the smooth-

ing affects the median like the reduction of resolution. Simultaneously, it reduces the localization error of the median.

From Fig. 7, we can see that this reduction of localization error is limited. Increasing the window size beyond  $N > 501$ , which corresponds to a smoothing interval of  $\pm 25$  m, does not reduce the localization error any further. When exceeding reasonable limits, the smoothing affects the algorithm output negatively. This is because it modifies the signal and corrupts it for large windows.

An interesting effect of increasing resolution can be seen in Fig. 6. Increasing the resolution provides results with a mode lower than the expected result. This is a combined effect of the choice of the algorithm described in Sect. 2.1 and of increasing the resolution. The algorithm points to the first point that satisfies the threshold relation. When increasing the number of points before the true one, we also increase the probability to have outliers before the true value. When choosing a different algorithm, e.g., the last point that is lower than the threshold, we will end with a distribution of results that will overestimate the true value.

#### 4 Application to mixing height retrievals

The estimation of MH comes with many dubious aspects. The first problem is the choice of a method to detect MH. The second problem is the choice of an algorithm to apply the

method. It is well known (Vogelezang and Holtslag, 1996; Seibert et al., 2000; Beyrich and Leps, 2012) that different methods produce different results in most cases. There are a few exceptions, like when the vertical profiles look like examples from a textbook (Beyrich and Leps, 2012).

In this context, we do not want to evaluate the uncertainties that MH has due to the choice of a method. This was done successfully by Beyrich and Leps (2012) by comparing the results of various methods to detect MH. We just estimate the uncertainty on an individual retrieval.

The retrieved uncertainty can be used for several purposes, e.g., to compare two different methods to see the degree of confidence by using Eq. (2).

Kretschmer et al. (2014) successfully used the qualitative localization error (Eq. 6) to filter data. The observations of the symmetry of  $U_\gamma(x_x)$  were used to reject the worst MH estimates performed on the Integrated Global Radiosonde Archive (IGRA) (Durre Imke et al., 2006) over Europe. The retrieved MH errors were then used to propagate the errors in a geostatistical interpolation that extended the observations to a grid over the domain.

As a practical application, we used the described methods to retrieve the uncertainties of MH from radiosonde data. Together with the already-introduced parcel method, we applied two variants of the bulk Richardson number method as described by Vogelezang and Holtslag (1996).

#### 4.1 Bulk Richardson number methods

Vogelezang and Holtslag (1996) analyzed three methods for stable and neutral conditions. All of them imply the use of different dimensionless profiles. However, in this work we use only the first two methods proposed by Vogelezang and Holtslag (1996). The third method they propose is not applicable in this work because it requires additional data sources not available from the radiosonde data themselves. The dimensionless profiles we analyzed are defined by the symbols  $Ri_b$  and  $Ri_g$ .

The first definition  $Ri_b$  should be used under stable conditions when wind is weak:

$$Ri_b(\theta_{v0}, \theta_{vh}, h, V_h) = \frac{gh}{\theta_{v0}} \frac{\theta_{vh} - \theta_{v0}}{V_h^2}, \quad (9)$$

where  $\theta_{v0}$  and  $\theta_{vh}$  are the virtual potential temperatures at the surface and at the top, respectively;  $g$  is the gravitational acceleration; and  $V_h$  is the wind speed. The variable  $h$  is the altitude above the ground, which is about 112 m a.s.l. for Lindenberg.

The second definition  $Ri_g$  is more appropriate for stable conditions with high winds:

$$Ri_g = Ri_g(\theta_{vs}, \theta_{vh}, h, u_h, v_h, u_s, v_s, z_s) \\ = \frac{g(h - z_s)}{\theta_{vs}} \frac{\theta_{vh} - \theta_{vs}}{(u_h - u_s)^2 + (v_h - v_s)^2}, \quad (10)$$

where the subscript  $s$  denotes a reference level. The variables  $v$  and  $u$  are the two horizontal wind components.

The reference level as used by Vogelezang and Holtslag (1996) is located on a meteorological tower, not on the radiosonde itself. They studied the effects of using different reference levels but observed no larger changes as long as the chosen reference level is close to 20 m. In this work, we did not have tower data available, so we used the second data point of each radiosonde profile as the reference level, which was located at approximately 20 m above ground.

To locate the MH, an appropriate critical or threshold value for  $Ri$  has to be selected. The MH is located where this threshold value is reached. A typical value for the threshold for the first method  $Ri_b$  is 0.25 (Seibert et al., 2000). For  $Ri_g$ , we used a critical value of 0.28, taken from Vogelezang and Holtslag (1996, their Table 1).

The critical number for  $Ri_g$  was taken from Vogelezang and Holtslag (1996, their Table 1), considering a general case with a reference level of 20 m. We followed Vogelezang and Holtslag (1996) using as reference layer the data taken closest to 20 m above the ground.

Examples of the profiles of  $Ri_b$  and  $Ri_g$  are plotted in Figs. 8 and 9.

#### 4.2 Localization error for a data set of high-resolution radiosonde profiles

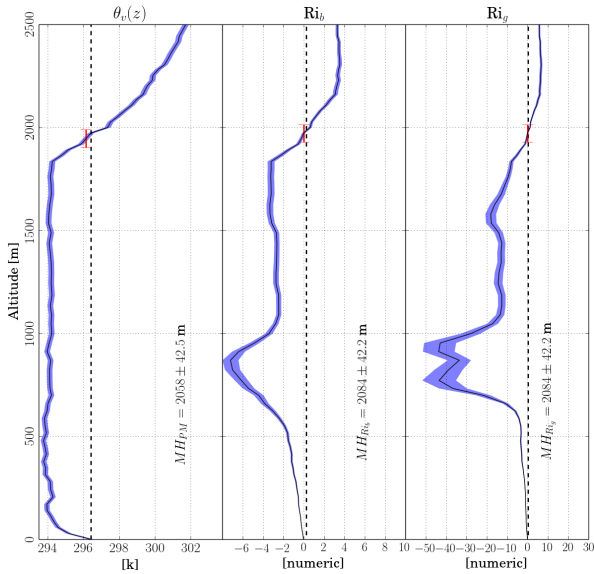
As introduced in Sect. 2, we applied our methodology to a subset of a well-known data set (Beyrich and Leps, 2012): the month of June 2010 at the Lindenberg Meteorological Observatory in Germany. During this time period, the presence of many different meteorological conditions allowed us to see the behavior of the localization error.

Using the radiosonde data, we calculated  $\theta_{vz}$ ,  $Ri_b(z)$ , and  $Ri_g$  as defined by Eqs. (1), (9), and (10), respectively.

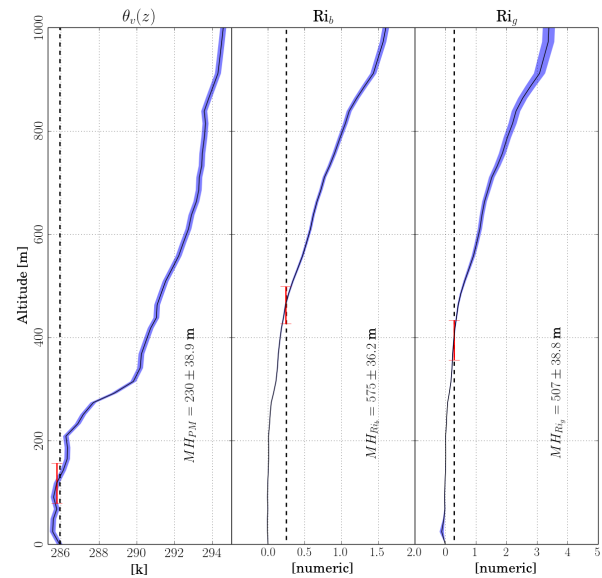
The following steps were taken to describe the process that we used to estimate  $\sigma_{x_m}$ :

1. propagate the errors using Eqs. (A4), (A9), and (A10);
2. use the algorithm presented in Sect. 2.1 for threshold detection;
3. estimate the  $U_{S3,y}(x_m)$  as defined in Sect. 3.1.5 for the retrieved MHs;
4. use Eq. (7) to calculate  $\sigma_{x_m}$ .

The result of such an analysis are presented in Fig. 10.



**Figure 8.** Profiles used for estimating MH on 24 June 2010, 12:00 UTC. The  $\theta_v$ ,  $Ri_b$  and  $Ri_g$  are depicted surrounded by a blue region that represents the error of the profile. The vertical dashed lines represent the threshold that locates the MH. The estimated MH with localization error is shown in red.



**Figure 9.** Profiles used for estimating MH on 3 June 2010, 06:00 UTC. The plot is analogous to Fig. 8.

### 5 Discussion

The methodologies for retrieving MH should be applied in proper meteorological conditions. The use of a wrong methodology directly results in a large localization error. This is clear from the time series of results in Fig. 10. For the radiosonde data collected at 18:00 UTC, the error bars are clearly larger than for other times of the day. Usually at 18:00 UTC at the Lindenberg station, the atmospheric profiles experience a transition from convective via neutral to stable conditions. This makes the parcel method unusable and also affects the Richardson number methods. The problems related to the 18:00 UTC MH retrieval are well known at this site. In particular, Beyrich and Leps (2012) found the largest differences between different MH retrieval methods for that time of the day.

Another reason for high uncertainties is that the wind speed might not be strong enough to justify the use of either  $Ri_b$  or  $Ri_g$ . Low wind speed combined with uncertainties in wind speed translates into large uncertainties in  $Ri_b$  and  $Ri_g$ , as wind speed appears with large exponents in the denominators of Eqs. (A9) and (A10). This can be clearly seen in Figs. 11 and 12. If the virtual potential temperature has only small changes with altitude, the parcel method will not produce a good localization.

We consider the points with small error bars in Figs. 8 and 9. Here the error is similar to the one encountered studying the effects of resolution in Sect. 3.4. In particular, our data set has a varying resolution with a mean of about 30 m. Looking at Fig. 6, we see that the expected error (47 m) for a good

localization is very similar to what we obtained for the well-developed convective case in Fig. 8. The uncertainty for well-localized points in our data set is generally around 40 m.

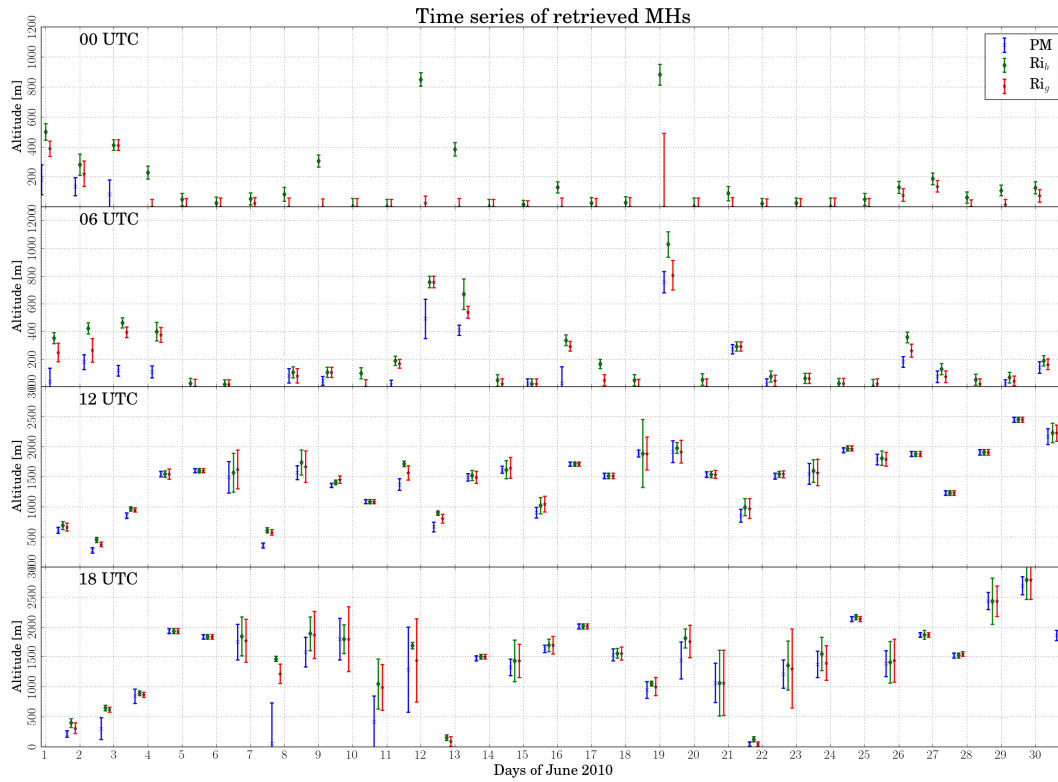
In the examples for bad localization (Figs. 11 and 12), both wind speed and temperature contribute to the large localization error (red error bar). However, we must distinguish between the cases of Figs. 11 and 12 because the large resulting errors are of a different nature. The case of Fig. 11 shows a strongly asymmetric confidence neighborhood, while the one presented in Fig. 12 is fairly symmetric.

The values of  $e_1$  and  $e_2$  are plotted in Figs. 11 and 12 as black error bars. In the asymmetric example of 20 June 2010 at 18:00 UTC, the contribution to  $\sigma_{x_m}$  comes mostly from the  $e_1$  component. So if we repeated the measurement, we would expect most of the values from below the retrieved MH and only few from above. For the symmetric example of 6 June 2010 at 18:00 UTC, the components  $e_1$  and  $e_2$  are much more similar in magnitude for all the retrieval methods.

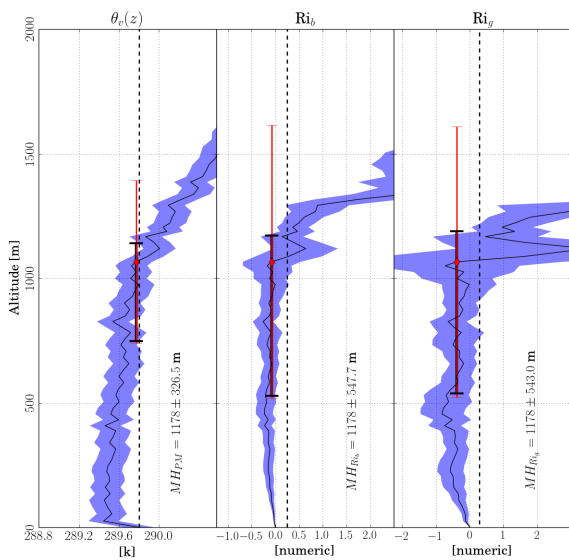
In fact, a more symmetric error or confidence neighborhood can allow us to use the localization error like an ordinary Gaussian error. However, when the distribution of the results is not Gaussian, the retrieved localization errors must be used with care. In particular, the  $e_1$  and  $e_2$  contributions should be considered separately.

### 6 Conclusions

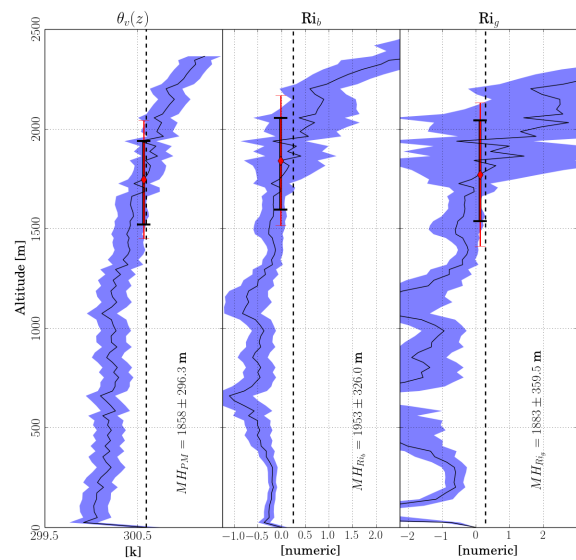
We defined a rigorous method for evaluating uncertainties in estimated quantities where standard error propagation cannot be used. It is particularly useful when a complex algorithm is necessary to find the location of a certain property and a Monte Carlo approach is too expensive. We call this un-



**Figure 10.** Time series of MHs calculated with three methods: blue is the parcel method (PM), green the  $Ri_b$ , and red the  $Ri_g$ . It presents the results obtained from the raw data without smoothing. The error bars represent the localization error calculated with Eq. (7).



**Figure 11.** Potential temperature,  $Ri_b$  and  $Ri_g$  profiles at the Lindenberg Meteorological Observatory on 20 June 2010, 18:00 UT. The propagated errors are depicted as light blue areas around the profiles. On each plot, the localization error  $\sigma_{x_m}$  is depicted as a red error bar. Black bars represent the distinct contributions  $e_1$  and  $e_2$  to  $\sigma_{x_m}$ .



**Figure 12.** Potential temperature,  $Ri_b$  and  $Ri_g$  profiles at the Lindenberg Meteorological Observatory on 6 June 2010, 18:00 UTC. The propagated errors are depicted as light blue areas around the profiles. On each plot, the localization error  $\sigma_{x_m}$  is depicted as a red error bar. Black bars represent the distinct contributions  $e_1$  and  $e_2$  to  $\sigma_{x_m}$ .

certainty of a localization the localization error  $\sigma_{x_m}$ . Its definition descends from the statistical concept of confidence (Eq. 2), which is considered in this work as a relation of equivalence of data points. To reach the mathematical definition of localization error in Eq. (7), we defined the strict confidence neighborhood in Eq. (5) as a contiguous set of equivalent data points.

Our work has proved to be a useful tool for qualitative analyses, and in particular for filtering data by quality of the retrieval (Kretschmer et al., 2014). When the confidence neighborhood is symmetric, the localization error can be considered analogous to the SD. Like the SD, the interval defined by the localization error will include the largest part of possible results if the measurements are repeated.

Depending on the actual signal, the interval defined by the localization error may be strongly asymmetric. In such a case, the expected distribution of possible results is not Gaussian, and the distribution will have non-negligible skewness. The correct use of the localization error then requires considering the left and right contributions to  $\sigma_{x_m}$  separately.

Our methodology was applied to compare different methods for retrieving mixing height from radiosonde data. This was not done to provide a better algorithm, nor to perform a general study on the best way to estimate mixing height. We rather provided a tool that can be used to better and more quantitatively compare different algorithms.

All the methodologies described in the literature provide values of MH without a specific error estimate. Instead, the uncertainty was estimated on the basis of the spatial and temporal variability of large data sets or by comparing results of different methods. Our goal was to estimate a reasonable uncertainty for one singular estimate of MH that depends only on the signals used and their uncertainties. The uncertainties that we retrieve this way are consistent with the climatological results of Beyrich and Leps (2012).

Our method is not limited to mixing height retrieval or atmospheric science at all. It can be applied to many problems where data points in a signal have to be localized: for example to find minima or maxima, or values exceeding a certain threshold. The localization error provided by our method can be used for error propagation in almost the same way as SD. It opens the possibility to check hypotheses by use of Welch's t test (Welch, 1947) or similar methods also for small data sets.

## Appendix

The errors of meteorological quantities retrieved by measured ones like potential temperature, virtual potential temperature and the different bulk Richardson number profile formulations are displayed here.

### A1 Error propagation

When combining different measurements  $y_k$  where  $K \in 1, \dots, M$  into a derived quantity  $f(y_1, \dots, y_M)$ , the proper way to calculate the error of the retrieved quantity  $\epsilon_f$  is the following:

$$\epsilon_f = \sqrt{\sum_{k \in \{1, \dots, M\}} \epsilon_{y_k}^2 \left( \frac{\partial f}{\partial y_k} \right)^2}, \quad (\text{A1})$$

where  $\epsilon_{y_k}$  is the error of the measurement of  $y_k$ .

### A2 Potential temperature

Potential temperature represents the temperature that an air parcel would have if adiabatically translated from a certain pressure level  $P$  to a reference level  $P_0$ .

$$\theta(z) = T(z) \left[ \frac{P_0}{P(z)} \right]^\gamma, \quad (\text{A2})$$

where  $\gamma$  is the adiabatic lapse rate

$$\begin{aligned} \sigma_\theta^2(z) = & \sigma_T(z)^2 \left[ \frac{P_0}{P(z)} \right]^{2\gamma} \\ & + \sigma_P^2(z) \frac{\gamma^2 P_0^2 T(z)^2 \left( \frac{P_0}{P(z)} \right)^{2\gamma-2}}{P(z)^4} \\ & + \sigma_{P_0}^2 \frac{\gamma^2 T(z)^2 \left[ \frac{P_0}{P(z)} \right]^{2\gamma-2}}{P(z)^2}. \end{aligned} \quad (\text{A3})$$

In this error formulation we are considering that the adiabatic lapse rate has no error. The value  $P_0$  and its error  $\sigma_{P_0}$  are provided for sake of generalities, but in general  $P_0$  is fixed to 1000 hPa and has no error.

### A3 Virtual potential temperature

Virtual potential temperature represents the temperature that an air parcel would have if adiabatically translated from a certain pressure level  $P$  to a reference level  $P_0$  and if all the water vapor condensed, so that the all condensation latent heat is released. If no liquid water content is considered, follow the definition below.

In order to retrieve the error propagation formula in this case we can use Eq. (A3), substituting the temperature and its error with  $T_e$  and its error:

$$\sigma_{T_e}^2(z) = \sigma_T^2(z) + \sigma_{\text{MR}}^2(z) b^2. \quad (\text{A4})$$

Before we give a formulation on the error on the MR, we have to point out that the coefficient  $b$  is not really a constant, being dependent on the water vapor's latent heat of condensation temperature, but in this context we are keeping it constant and neglecting any error associated with its value.

Radiosondes are measuring relative humidity, so MR must be calculated from this quantity. This task can be performed by calculating the water vapor pressure directly from RH and pressure after the estimation of water vapor saturation pressure.

$$\text{MR}(z) = \frac{e(z)}{P(z) - e(z)} = \frac{\text{RHe}_{\text{sat}}(z)}{P(z) - \text{RHe}_{\text{sat}}(z)}, \quad (\text{A5})$$

where  $e(z)$  is the water vapor pressure at the altitude  $z$ . The relative humidity in this expression and in the following is intended to be divided by 100, and  $e_{\text{sat}}(z)$  is the saturation pressure of water vapor. For this last quantity many empirical formulations can be found as a function of temperature. Due to the large number of expressions we have chosen to refer to Bolton's approximation (Bolton, 1980):

$$e_{\text{sat}} = 6.112 e^{17.67T/(T+243.5)}, \quad (\text{A6})$$

where  $T$  is expressed in °C and pressure in hPa. For this expression the error propagation is

$$\sigma_{e_{\text{sat}}}^2 = \sigma_T^2 37.357 \left[ \frac{17.67}{T+243.5} - \frac{17.67T}{(T+243.5)^2} \right]^2 e^{\frac{35.34T}{T+243.5}}. \quad (\text{A7})$$

Propagating the errors on the MR expression, we obtain

$$\begin{aligned} \sigma_{\text{MR}}^2(z) = & \sigma_{e_{\text{sat}}}^2 \left[ \frac{\text{RH}}{P - \text{RHe}_{\text{sat}}} + \frac{\text{RH}^2 e_{\text{sat}}}{(P - \text{RHe}_{\text{sat}})^2} \right]^2 \\ & + \sigma_{\text{RH}}^2 \left[ \frac{e_{\text{sat}}}{P - \text{RHe}_{\text{sat}}} + \frac{\text{RHe}_{\text{sat}}^2}{(P - \text{RHe}_{\text{sat}})^2} \right]^2 \\ & + \sigma_P^2 \frac{\text{RH}^2 e_{\text{sat}}^2}{(P - \text{RHe}_{\text{sat}})^4}. \end{aligned} \quad (\text{A8})$$

Substituting Eq. (A7) in this last expression, we obtain an expression on the error of MR that consider only the quantities directly measured by the RS92-SGPD.

### A4 Error propagation on bulk Richardson number methods

Equation (9) gives us the following expression for the errors:

$$\begin{aligned} \sigma_{\text{Ri}_b}^2 = & \sigma_{\theta_{v0}}^2 g^2 h^2 \left( \frac{1}{\theta_{v0} V_h^2} - \frac{\theta_{v0} - \theta_v}{\theta_{v0}^2 V_h^2} \right)^2 + \sigma_{\theta_v}^2 \frac{g^2 h^2}{\theta_{v0}^2 V_h^4} \\ & + \sigma_h^2 g^2 \frac{(\theta_{v0} - \theta_v)^2}{\theta_{v0}^2 V_h^4} + 4\sigma_V^2 g^2 h^2 \frac{(\theta_{v0} - \theta_v)^2}{\theta_{v0}^2 V_h^6}. \end{aligned} \quad (\text{A9})$$

It must be noted that the variable  $h$  is defined as the difference between the measured altitude over mean sea level and

the altitude of the balloon release area; for this reason, on a nonuniform terrain the uncertainty of this quantity can be relevant. Propagation of the  $Ri_g$  adds to Eq. (A9) also the dependence on the wind direction used to calculate the wind components, as well the errors in the reference levels.

$$\begin{aligned}
\sigma_{Ri_g}^2 = & \sigma_{\theta_{vs}}^2 g^2 \left\{ \frac{h - z_s}{\theta_{vs} [(u_h - u_s)^2 + (v_h - v_s)^2]} \right. \\
& \left. + \frac{(h - z_s)(\theta_{vh} - \theta_{vs})}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]} \right\}^2 \\
& + \sigma_{\theta_{vh}}^2 g^2 \frac{(h - z_s)^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^2} \\
& + \sigma_h^2 g^2 \frac{(\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^2} \\
& + \sigma_{z_s}^2 g^2 \frac{(\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^2} \\
& + \sigma_{u_h}^2 g^2 \frac{(2u_h - 2u_s)^2 (h - z_s)^2 (\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^4} \\
& + \sigma_{v_h}^2 g^2 \frac{(2v_h - 2v_s)^2 (h - z_s)^2 (\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^4} \\
& + \sigma_{u_s}^2 g^2 \frac{(2u_h - 2u_s)^2 (h - z_s)^2 (\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^4} \\
& + \sigma_{v_s}^2 g^2 \frac{(2v_h - 2v_s)^2 (h - z_s)^2 (\theta_{vh} - \theta_{vs})^2}{\theta_{vs}^2 [(u_h - u_s)^2 + (v_h - v_s)^2]^4} \quad (A10)
\end{aligned}$$

The Supplement related to this article is available online at doi:10.5194/amt-8-4215-2015-supplement.

*Acknowledgements.* This work is a contribution to the FP7 project ICOS–INWIRE, funded under grant agreement no. 313169. We are much obliged to Frank Beyrich of the Deutscher Wetterdienst (DWD), who provided us with the high-resolution radiosonde profiles.

Edited by: L. Bianco

The article processing charges for this open-access publication were covered by the Max Planck Society.

## References

- Beyrich, F. and Leps, J.-P.: An operational mixing height data set from routine radiosoundings at Lindenberg: Methodology, *Meteorol. Z.*, 21, 337–348, doi:10.1127/0941-2948/2012/0333, 2012.
- Bolton, D.: The computation of equivalent potential temperature, *Mon. Weather Rev.*, 108, 1046–1053, doi:10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2, 1980.
- Durre Imke, Vose Russell, S., and Wuertz David, B.: Overview of the integrated global radiosonde archive, *J. Climate*, 19, 53–68, doi:10.1175/JCLI3594.1, 2006.
- Gerbig, C., Körner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, *Atmos. Chem. Phys.*, 8, 591–602, doi:10.5194/acp-8-591-2008, 2008.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, doi:10.1038/415626a, 2002.
- Holzworth, G. C.: Estimates of Mean Maximum Mixing Depths in the Contiguous United States, *Mon. Weather Rev.*, 92, p. 235, doi:10.1175/1520-0493(1964)092<0235:EOMMMD>2.3.CO;2, 1964.
- Kretschmer, R., Gerbig, C., Karstens, U., and Koch, F.-T.: Error characterization of CO<sub>2</sub> 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.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO<sub>2</sub>, *Atmos. Chem. Phys.*, 14, 7149–7172, doi:10.5194/acp-14-7149-2014, 2014.
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *P. Natl. Acad. Sci. USA*, 104, 18925–18930, doi:10.1073/pnas.0708986104, 2007.
- Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: CO<sub>2</sub> flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, 3, 1919–1964, doi:10.5194/acp-3-1919-2003, 2003.
- Seibert, P., Beyrich, F., Gryning, S., Joffre, S., Rasmussen, A., and Tercier, P.: Mixing layer depth determination for dispersion modelling. European Commission, in: COST Action 710-Final Report. Harmonisation of the pre-processing of meteorological data for atmospheric dispersion models. L-2985 Luxembourg: European Commission, EUR 18195 EN, edited by: Fisher, B., Erbrin, J., Finardi, K. S., Jeannot, P., Joffre, S., Morselli, M., Pechinger, U., Seibert, P., and Thomson, D., Office for Official Publications of the European Communities, <http://bookshop.europa.eu/en/cost-action-710-pbCGNA18195/> (last access: 15 May 2015), 1998.
- Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., and Tercier, P.: Review and intercomparison of operational methods for the determination of the mixing height, *Atmos. Environ.*, 34, 1001–1027, doi:10.1016/S1352-2310(99)00349-0, 2000.
- Stull, R.: An Introduction to Boundary Layer Meteorology, vol. 13 of Atmospheric and Oceanographic Sciences Library, Kluwer Academic Publishers, Dordrecht, the Netherlands, 670 pp., doi:10.1007/978-94-009-3027-8, 1988.
- Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation, *Bound.-Lay. Meteorol.*, 37, 129–148, doi:10.1007/BF00122760, 1986.
- Vaisala: Vaisala Radiosonde RS92-SGP, available at: , last access: 15 May, 2015.
- Vogelezang, D. and Holtslag, A.: Evaluation and model impacts of alternative boundary-layer height formulations, *Bound.-Lay. Meteorol.*, 81, 245–269, doi:10.1007/BF02430331, 1996.
- Welch, B. L.: The generalization of Student's problem when several different population variances are involved, *Biometrika*, 34, 28–35, available at: <http://biomet.oxfordjournals.org/content/34/1-2/28.short> (last access: 9 October 2015), 1947.