Reviewer 3 (RC1)

"The paper presents an interesting technique to infer cloud base height from the MISR standard cloud product. It demonstrated a valuable skill with this technique that uses the 15 percentile threshold to the vertical distribution of MISR cloud heights in a 10-km domain. The algorithm can be readily applied to all MISR cloud data for seasonal and global statistics of cloud base height."

We thank the reviewer for her or his constructive comments and suggestions.

"The technique is perhaps valid for broken cloud scenes in the 10km domain, but would fail if clouds are 100 optically thick or overcast in the domain. This is often the case over land, but not necessary over ocean. The authors should acknowledge this limitation in the abstract and conclusion."

Yes, in case of complete overcast, MIBase cannot derive the cloud base. The algorithm requires at least one MISR z retrieval within the MIBase cell which is a surface pixel according to the stereo derived cloud mask. This is stated in the abstract: "[...] It can be applied if some cloud gaps occur within the chosen distance [...]". The spatial distribution of the frequency of occurrence of apparent overcast situations is now included in the revised manuscript (Fig. 10d).

Other reviewers noted the limitation for different cloud types, too. Therefore, some additional investigations have been performed and abstract and conclusions have been revised accordingly.

"For the sensitivity calculations summarized in Table 3, the results might be dependent on roughness of terrain since MISR cloud height retrievals would correlate worse with ceilometer base height if the site is surrounded by mountains. What would be the results if only those sites with flat terrain (in 10, 20, or 30 km radius) are included in statistics?"

We addressed this interesting point by an additional analysis. To derive a quantity to judge what is "flat" terrain, we use the standard deviation of the average scene elevation (ASE). The ASE is taken from the MISR Ancillary Geographic Product (Bull et al, 2011) at 1.1 km resolution (like MISR cloud top height product). We proceed as follows:

- use ASE within MIBase cell around each ceilometer station (R_c = 5 km, 10 km, 15 km, 20 km, 30 km) - Note, we changed the notation from field of view to MIBase cell and R_fv to R_c
- calculate standard deviation of ASE for each ceilometer station and each R_c
- plot density of standard deviations for each R_c
- filter out stations with higher standard deviation to calculate statistics (correlation r, bias, RMSE)

A figure showing the results has been added to additionally compiled supplement (Fig. S1). Skill scores (correlation, bias, RMSE) were calculated for subsamples with standard deviations (sd) below 20 m, 30 m, 40 m, 50 m, and 60 m corresponding to the 74th, 84th, 89th, 92th, and 94th percentile of the distribution of sd, respectively, for the smallest radius (5 km). For the largest radius (30 km), the corresponding percentiles are 30th, 49th, 60th, 66th, and 71th. In other words, for a larger radius, more stations are excluded for the same sd threshold.

The correlation increases slightly if the more varying terrain is excluded but only until sd=50m. If an even stricter threshold is applied, the correlation decreases again. The bias improves more or less monotonically with a stricter sd threshold, i.e. for "flatter" terrain. The RMSE improves as well.

This is discussed in the new Section 4.1 of the revised manuscript. In summary, limiting the comparison to sites showing more homogeneous terrain improves the comparison only slightly.

"Some minor issues and English: p4, line 12: MISR cloud motion vector in L2TCSP file is determined at 17.6 km resolution, and is used to derive H_SDCM by correcting the wind-induced parallax effect. As noted in Mueller et al. (2013, 2016), the cloud height and along-track wind errors are correlated. p17, line 21 .. shows a higher number of ... p.17, line 27 .. seasons .. p20, line 20 ... mentioned ..."

Thank you for these points. We took them into account and adjusted the manuscript accordingly.

Reviewer 4 (RC2)

"This paper describes an interesting technique to infer cloud base height from MISR measurements within selected areas. The paper is well written. The technique is described well. I recommend the paper to be accepted for publication after some minor suggested additions and corrections listed below."

Thank you for your constructive feedback. We have addressed your comments in the following way:

"The authors should describe better to which kind of cloud fields this method can be applied. The abstract states it can be applied if some cloud gaps occur within the chosen distance of typically 10 km. However, cirrus are excluded in the evaluation, because it probably would not work on cirrus. I also do not expect the technique to work particularly well on areas dominated by deep convection and congestus, for example. Please discuss the expected limitations of the technique related to cloud types."

We agree that MIBase has limitations in respect to cloud types. Therefore, we introduced two new sections "3.4 Scene limitation" and "4.1 Scene structure influence" into the manuscript and modified abstract and conclusions accordingly. In short: A bias towards a certain cloud type can be introduced by two sources: the MISR cloud top product, yielding more valid retrievals for specific cloud types, or the MIBase algorithm. Therefore, these two aspects would have to be investigated separately. Many approaches to distinguish different cloud types from satellite data have been proposed, e.g. the cloud optical depth / top height approach by the International Satellite Cloud Climatology Project (ISCCP). However, this kind of classification is not unique, depends on the horizontal resolution and likely needs additional data products. Therefore, it goes beyond the presented study.

To investigate the performance of the MIBase algorithm in dependence on parameters which are also relevant for cloud type classification, we determine RMSE, bias, and the correlation coefficient in dependence on z_{top} and the cloud vertical extent Δz. A figure showing the results can be found in the added supplement material (Fig. 2). For most cloud fields which are observed within this study, z_{top} ranges between 1000 m and 2000 m (supplement Fig. 2a). For a lower z_{top}, the RMSE shows a minimum of approximately 300 m and increases for clouds with higher z_{top}. As we already discussed at the end of section 4, this behavior could be due to the termination of the z_{base} range by the threshold height $h_{\min}$. However, in case of even lower z_{top} values, the RMSE increases. As these low z_{top} values approach the threshold height, two different cloud scenes are possible: the cloud extents below the threshold height indicating near surface clouds or fog, or the cloud is extremely thin. In this study we excluded scenes for which the ceilometer reported a cloud below the threshold height. Therefore, low clouds or fog should not be included in the statistics, unless the ceilometer did not detect it. In particular for very thin clouds, the RMSE is lowest (supplement Fig. 2f). In conclusion the higher RMSE for very low z_{top} values could indicate that the MIBase algorithm does not perform as well in proximity to the threshold height. This is also indicated by an increasing correlation coefficient with increasing z_{top} (supplement Fig. 2d). High correlation coefficients (supplement Fig. 2h) and low RMSE (supplement Fig. 2f) for cloud thicknesses up to 1000 m indicate that the algorithm works particularly well for thinner clouds.

For cirrus clouds or high clouds in general, we cannot make a robust statement as we tax the accuracy of the METAR reports insufficient for this particular height range.

For overcast situations, z_{base} cannot be retrieved. We added statistics on how many apparent overcast cases are observed to the manuscript (Tab. 5, Fig. 10 of the revised manuscript).

"Page 7, line 7: I agree that MISR cloud top heights are probably superior to those of other passive satellite instruments, but not to those from active instruments, in particular lidar."

Agreed. We added "passive" to the sentence.

"Figure 12: In the caption note that these are anomalies. Also add a Delta in front of the y-axis labels."

Thank you for pointing out that this was not clear enough. We added a Delta to the y-axis label and edited the caption to make it clearer to the reader that we show anomalies here.

"I thought the discussion of multi-layer situations on page 8 was interesting and I suggest to add some words about that in the conclusions."

If we include multi-layer cases we would add 689 cases (10%) to the statistics for the year 2007 (432 cases or 8% in 2008). With these additional cases, the correlation with the ceilometer retrievals decreases slightly from 0.66 to 0.64 (2007 and 2008) and the RMSE increases slightly from 385 m to 395 m (2007) and from 404 m to 418 m (2008). The MISR cloud top height product includes a correction for cloud advection. This is carried out via a cloud motion vector which is determined at a certain cloud feature height at a 17.6 km horizontal resolution. The wind correction is applied to any cloud top height retrieval which is within 840 m distance from the feature height of the cloud motion vector. Collocated cloud motion vectors and their neighbors are considered for the correction of a cloud top height retrieval. We suspect that the wind correction in multi-layer cases, i.e. cases with a wide range of cloud heights, might not be as accurate. At the same time multi-layer cases might also lead to a false comparison with the ceilometer, since it is unclear which layer passed over the ceilometer and which may have not. Therefore, we decided to exclude multi-layer cases from the evaluation and do not mention them in the conclusion.

Reviewer 5 (RC3)

Thank you for your constructive feedback. We have addressed your comments in the following way:

"The attempt to derive cloud base heights from MISR data is interesting, but as far as I can tell they basically take the minimum retrieved cloud height, assume it corresponds to the base height and move on from there. The authors need to state more clearly that this algorithm is only valid over broken clouds, [...]"

As also other reviewers commented on the dependence on cloud types, we introduced two new sections "3.4 Scene limitation" and "4.1 Scene structure influence" into the manuscript and modified abstract and conclusions accordingly. Specifically we added: "The occurrence of a broken cloud field is a basic assumption of MIBase." to Section 3.1.

"[...] indeed I would be very interested in seeing a study of the accuracy of the results as a function of scene structure and degree of brokenness, [...]"

Figure 8 in Section 4.1 of the revised manuscript shows the dependence of RMSE, bias and correlation coefficient on the configuration of the stereo-derived cloud mask. In particular, the dependence on the number of z retrievals marked high confidence cloud within the considered cloud field can serve as a proxy for cloud cover fraction.

"[...] and also as a function of the number of unobscured cloud top and side pixels as available in the MISR TC_ALBEDO product. I am willing to reconsider the paper if the authors perform such a study as I think that would be much more interesting than just the minor algorithm parameters such as R, N and P."

The TC_ALBEDO (MIL2TCAL) product provides the number of unobscured top and side pixels at a resolution of 2.2 km. This means, the number of pixels at the actual MISR resolution (275 m) within a 2.2 km area which observe the same reflecting layer are counted. Therefore, the product might suffice as an indicator of a more or less complex scene structure. As the influence of the scene structure on the MIBase performance has been brought up also by other reviewers, we decided to extent the discussion on this (new Section 4.1). Instead of using the MIL2TCAL product to further investigate the scene structure, we decided to exploit the stereo-derived cloud mask in more detail because it is already included in the MIL2TCSP product which builds the base for MIBase. Furthermore, we investigated the influence of z_{top} and Δz on the performance of MIBase. These parameters are also characterising the scene structure in more detail. An additional figure is included in the supplement (Fig. 2).

"Additionally they need to clarify which MISR product they are using (TC_STEREO or TC_CLOUD), and which type of SDCM (WindCorrected or WithoutWindCorrection). I am unsure if they are using Stereo or Cloud, because they mention the correct short name for Cloud, but list the wind resolution as being 70.6 km and Stereo is at 70.4 km and Cloud retrieves its winds at 17.6 km. I am hoping this is just a typo on their part but Im not sure. It is my opinion that they need to use the WindCorrected heights from the TC_CLOUD product."

Thank you for pointing this out. Unfortunately, a typo occurred which has been corrected. As we state in the manuscript, we are using the MISR Level 2TC Cloud Product (MIL2TCSP) which provides the cloud top height and the stereo-derived cloud mask at a 1.1 km resolution with and without wind correction. Here, we are using only the wind corrected data sets. As stated in the MISR Level 2 Cloud Product Algorithm Theoretical Basis (Mueller et al., 2013) the wind correction is carried out via a Cloud Motion Vector which is determined at a resolution of 17.6 km, like you mention. We added a sentence about the wind correction to Section 2.1 to make this clear.

Reviewer 6 (RC4)

"The authors propose a method to derive cloud base height from MISR measurements. Here, they make use of the 9-angle viewing capabilities of the instrument and derive all possible cloud top heights within a specified area, the (approximately) lowest z_top is then attributed to be the base height of the cloud field within the specified area."

Thank you for your constructive feedback. We have addressed your comments in the following way:

"For this algorithm to work, several preconditions have to be met, as specified by the authors. First, the cloud field has to be inhomogeneous so that MISR can see thin cloudy layers around the cloud fields edges. Second, it should not be used for thin cirrus. Personally I would say it will probably also have problems in regions with very inhomogeneous cloud bases or in regions with strong convective systems which means very inhomogeneous but also very thick clouds. Due to these restrictions I am not convinced that this product will be an easy-to-use tool for the quantitative assessment of cloud base height in climate models as stated in the conclusions."

We agree that MIBase has limitations in respect to cloud types. Thin cirrus will be problematic because the MISR clout top height retrieval method is based on frequencies in the visible light range, for which thin cirrus is hard to detect. Therefore, a height limit of 5 km is used for the global application. Heterogeneous cloud base heights pose a challenging scene as well, since we assume that the lower end of the cloud top height distribution is representative for the cloud base height within the region of interest. However, any kind of retrieval method may have trouble with heterogeneous cloud base heights. In the new Section 4.1 "Scene structure influence", we included an investigation of the MIBase performance in dependence on Δz and z_{top} (supplement Fig 2). In short, MIBase performs best for shallow low clouds.

We agree that some constraints have to be taken into account when using MIBase to evaluate cloud base height in climate models. MIBase could still be a valuable tool, if for example the climate model output is limited to clouds below 5 km and cloud fractions below 1. While the comparison of individual clouds suffers from the large uncertainty, evaluation on seasonal and inter-annual scales should yield robust results. We modified the conclusion accordingly.

"However, the comparison to METAR data shows good results. The article is well written and the method is clearly explained.
Nevertheless, I think it could be improved because a better analysis of the situations in which the retrieval does not perform well would be necessary in order to evaluate its capabilities."

We agree that such an analysis would be beneficial. Therefore, we included the above mentioned new Section 4.1 in which we present further investigations of the scene structure. Besides evaluating the performance of MIBase in dependence on Δz and z_{top}, we also exploited how the configuration of the stereo-derived cloud mask influences the performance. This way we assessed for which scenes the algorithm performs better or worse.

"Also some statistics that quantify, in how many cases the algorithm could not retrieve a cloud base height is missing. These values should be given for each possible retrieval rejection, a too homogeneous cloud cover for instance, in comparison to the number that would have theoretically been possible."

To elaborate on this in more detail, we added Section 3.4 "Scene limitations" to the manuscript. Statistics on the situations for which MIBase cannot retrieve z_{base} are discussed quantitatively. Following the numbers in the new Table 5 and the description in the text, we now allow the reader to comprehend how we ended up with the number of cases which are considered for the calibration and validation of the algorithm. Furthermore, we also extended Section 5.1 by a discussion regarding the number of valid retrievals versus retrieval failure. Figure 10 of the revised manuscript shows the spatial distribution of scenes for which MIBase cannot retrieve z_{base}, i.e. apparent clear sky and apparent overcast.

"In Fig. 9 b), the ITCZ should be more visible in the Atlantic Ocean and over Africa, there are almost no z_top values over 1.4 km. Even if the analysis is restricted to cases with z_top < 5000 m, I would assume that there should be more z_tops higher than 1.4 km. Could you please comment on that?"

In Fig. 9b and Fig. 10a, 10b (Fig. 11a, 11b in the revised manuscript), the ITCZ is revealed by the light turquoise band slightly north of the equator, indicating higher z_{base} and z_{top} compared to the immediate surroundings to the north and south. This band is most pronounced in the Pacific ocean. Over the Atlantic, it can be seen most clearly in the manuscript's Fig. 9c, which shows a band of increased cloud vertical extent in that region. As stated in the manuscript, over continents the diurnal cycle should be kept in mind. MISR has a morning overpass which means, the three year median heights provided here represent the morning heights around 10 a.m. local time. For the Congo Basin, Taylor et al. (2007) investigated the diurnal cycle of cloud top temperature (CTT) retrieved via satellite remote sensing (SEVIRI). According to them, the CTT is lowest around the MISR overpass time with a mean value of about 290 K during late morning hours. If we take the observed z_{top} of about 1200 m and assume a lapse rate of 0.6 $\frac{\mathrm{K}}{100\mathrm{m}}$, the extrapolated surface temperature would be 297 K ($\approx 24°\mathrm{C}$) which seems very plausible.

"And why is z_top restricted to 5000 m, is this threshold not only applied to z_base in order to exclude cirrus?"

We agree, that the limit for z_{top} should not be the same as for z_{base}. Therefore, we reproduced the figures for the global distribution of z_{top} . This time, the median is calculated only for those z_{top} values for which the respective z_{base} is below the 5000 m threshold. We updated Fig. 9 and Fig. 11 of the revised manuscript and their respective captions accordingly. Generally, a threshold is necessary to exclude high clouds from the analysis in order to avoid difficulties associated with cirrus clouds. In our opinion the median of z_{base} and z_{top} provides less valuable information if low and high clouds are mixed together. From our best judgement, 5000 m seems like a good choice for a threshold to ensure that the algorithm works properly. The resulting product is not highly sensitive to this threshold as can be seen in Fig. 4 (supplement).

"Fig. 9 a): Since the number of valid retrievals over the Sahara is so small, it is quite understandable, that the cloud base height jumps between very small and very high values and a warning is given by the authors on page 17. In order to use maps of this kind for a climate model evaluation, many more valid data points would be necessary. This should be noted in the conclusions."

We included a note of this in the conclusion: "This makes MIBase a promising tool for the evaluation of climate models on seasonal and inter-annual time scales in data sparse regions if for example the climate model output is limited to clouds below 5 km and cloud fractions below 1 and if a sufficient amount of MIBase retrievals is provided within the considered region and time period."

"Fig. 9c): Why is the sample size low over Antarctica? Shouldnt it be covered with approx. 50% cloud cover throughout the year?"

MISR's stereo-derived cloud mask shows configurations which indicate apparent clear sky conditions in Antarctica for 60 % to almost 100 % of the cases (Fig. 10c of the revised manuscript, and Fig. 3a of the supplement). This is in agreement with the cloud cover derived from MODIS presented by Suen et al. (2014). Of course, if the clouds are below the threshold height, the retrievals would appear as low or high confidence surface in the stereo-derived cloud mask. We added the following sentence to the manuscript: "However, the boundary layer is typically shallower in polar regions. Therefore, boundary layer clouds occur likely below $h_{\min}$, so that z_{base} cannot be retrieved by the MIBase algorithm."

"p 4, l 19: please specify SDCM in H_SDCM"

SDCM stands for "stereo-derived cloud mask". We added the abbreviation in parenthesis at the first occurrence of this phrase. In particular, H_{SDCM} is the threshold height which is applied to derive the stereo derived cloud mask according to Equation 59 in the Algorithm Theoretical Basis documentation by Mueller et al. (2013). We added that this is a threshold height to the manuscript.

"p 16, l 6: yielding an overall higher something is missing here."

This should be "yielding an overall higher Δz". Thank you for pointing this out.

"P 17, l 1: Do you refer to Fig. 9 c) instead of b)?"

Yes. Thank you and sorry for the confusion!

Reviewer 2 (RC5)

"The authors describe a novel algorithm for the retrieval of cloud base height (CBH) from MISR satellite measurements. Global information on cloud base height is important for many applications and the retrieval approach is interesting and promising. However, the manuscript is not sufficiently convincing in demonstrating the reliability of the new CBH product. Below are a number of major issues to be addressed before this manuscript may be suitable for publication."

We thank the reviewer for the constructive feedback. We have implemented your comments in the following way:

"The algorithm is tuned with METAR observations over the U.S., i.e. extratropical land surface. How representative is this for ocean surfaces and other climate zones, where different cloud types prevail? To show the skill in other regions, some comparisons with independent measurements elsewhere would be required."

Our study includes cloud height retrievals over the continental U.S. over the course of two years (one year for calibration, one for validation). Therefore, various cloud types should be included in the analysis. Clouds within Arctic air masses which typically occur in the northern U.S. during winter, as well as tropical like deep convective clouds which typically occur during the summer in the southeastern U.S. should be included as well as stratocumulus clouds which usually occur at the coast of California. The METAR data set includes maritime island stations in the Gulf of Mexico and near the west coast of the U.S.

The utilized MISR product does not distinguish between land and ocean, but only between cloud and surface by a geometric technique. Furthermore, for each retrieval scene a particular configuration of the stereo-derived cloud mask is provided which characterizes the scene structure. Therefore, if a similar scene structure is found outside the continental U.S. region, for which the calibration and validation has been carried out, MIBase should perform similarly. However, we agree, that additional validation in other regions would be beneficial to backup this statement.

To further investigate MIBase limitations in respect to cloud types, we introduced two new sections 3.4 Scene limitations and 4.1 Scene structure influence into the manuscript and modified abstract and conclusions accordingly. In short: The statistics are rather robust with regard to the configuration of the stereo-derived cloud mask. The bias depends on the number of z retrievals marked high confidence cloud. This indicates that a bias correction might be feasible. Since the origin of the bias is not fully understood yet, we like to leave such potential improvements to future studies.

Furthermore, the new Section 4.1 includes an investigation of the influence of z_{top} and Δz on the MIBase performance. Such parameters are also important for cloud type classification. High correlation coefficients (supplement Fig. 2h) and low RMSE (supplement Fig. 2f) for cloud thicknesses up to 1000 m indicate that the algorithm works particularly well for thinner clouds. For cirrus clouds or high clouds in general, we cannot make a robust statement as we tax the accuracy of the METAR reports insufficient for this particular height range.

"More information on the success rate of the retrieval algorithm is required to evaluate how useful it is. Statistics of the number of samples n_s are given but these are only absolute numbers, not (fractional) success rates. For example, Table 3 indicates that n_s is between 3059 to 7772 depending on R_fv. A rough calculation based on 1510 ceilometers, a MISR revisit time of 6 days, and a cloud fraction of 50% would potentially yield around 45,000 cloudy collocations. This suggests that in only 10% of the cloudy cases, a valid CBH retrieval is obtained. Is this correct? Such statistics, accompanied by the relative occurrence of different causes of retrieval failure, need to be provided, also for the global plots, to evaluate the applicability of the method."

We added statistics regarding the success rate of the algorithm to the manuscript (Section 3.4 and Section 5.1 of the revised manuscript). For the comparison with the retrievals from METAR, we combined the numbers for all sites and present the resulting numbers for the years 2008 (calibration) and 2007 (validation) individually in the new Section 3.4 including new Table 5. For the year 2008, we downloaded data which provided 80454 overpasses. Only 65 % of those contained valid z retrievals at the METAR sites with a corresponding METAR message. Out of those potential cases, about 30 % do not include z retrievals marked high confidence surface. Please, see the revised manuscript for more details. For the global application, the calculations are carried out for each grid cell, so that the spatial distribution of the numbers can be studied (Fig. 10 of the revised manuscript). We added a discussion on the retrieval failure statistics to Section 5.1.

"The calibration of the algorithm is done for z_base smaller than 3000m, because of the limited range of the ceilometers. However, for the global composites an upper threshold of 5000m is used. It is unclear whether this extrapolation outside the range for which the product has been trained, is valid."

One of the questions this manuscript tries to answer is whether the cloud base height can be derived from the MISR cloud top height product. The comparison with the ceilometer demonstrates that this is possible in case some preconditions are met. The reason this works is because typically the cloud top height is heterogeneous leading to geometrically thinner and thicker parts of the cloud. As far as we understand, this concept should not change for different heights within the troposphere. This means, if it can be validated within one height region, it should work in other height regions as well. However, we do agree that the algorithm might perform differently for different cloud types. Then of course, for different heights the distribution of cloud types varies. We discussed the dependence on cloud types in the reply to the first comment.

"The global maps in Figs. 9 and 10 are hard to interpret because upper limits of z_base and z_top have been applied. What does the median of a distribution cut-off at some value tell us?"

Applying a threshold is necessary to exclude high clouds from the analysis. This is appropriate, because we are focusing the study on clouds which occur in the lower troposphere. In our opinion, the median of z_{base} and z_{top} provides less valuable information if low and high clouds are mixed together. From our best judgement, 5000 m seems like a good choice for a threshold to ensure that the algorithm works properly. The resulting product is not highly sensitive to this threshold as can be seen in Fig. 4 (supplement). Inspired by reviewer 6 (RC4) who questioned the height limit for z_{top}, we changed the calculation of the median z_{top} height. We reproduced the figures by calculating the median only for

those z_{top} values for which the respective z_{base} is below the 5000 m threshold. We updated Fig. 9 and Fig. 11 of the revised manuscript and their respective captions accordingly.

"Im also confused by the description of Fig. 9, which says that the ITCZ is clearly visible with higher z_base and z_top. In the plots a brown band can be seen, but these are lower rather than higher values. Can you explain?"

In Fig. 9b and Fig. 10a, 10b (Fig. 11a, 11b in the revised manuscript), the ITCZ is revealed by the light turquoise band slightly north of the equator, indicating higher z_{base} and z_{top} compared to the immediate surroundings to the north and south. This band is most pronounced in the Pacific ocean. Over the Atlantic it can be seen most clearly in the manuscript's Fig. 9c, which shows a band of increased cloud vertical extent in that region. As stated in the manuscript, over continents the diurnal cycle should be kept in mind. MISR has a morning overpass which means, the three year median heights provided here represent the morning heights around 10 a.m. local time. For the Congo Basin, Taylor et al. (2007) investigated the diurnal cycle of cloud top temperature (CTT) retrieved via satellite remote sensing (SEVIRI). According to them, the CTT is lowest around the MISR overpass time with a mean value of about 290 K during late morning hours. If we take the observed z_{top} of about 1200 m and assume a lapse rate of 0.6 $\frac{\mathrm{K}}{100\mathrm{m}}$, the extrapolated surface temperature would be 297 K ($\approx 24°\mathrm{C}$) which seems very plausible.

"Is it also possible that the results in these multi-year median are biased to certain cloud types? For example, in the stratocumulus (Sc) areas west of the continents, cases with closed Sc will probably not yield a valid retrieval, while for open Sc z_base can be retrieved, so that the end result will be biased to the latter."

An inherent bias of the method results by the necessary condition of a MISR z retrieval which is marked high confidence surface by the stereo-derived cloud mask. In other words, a broken cloud scene is required. Therefore, cloud base heights for situations with overcast are not included in the calculation of the three-year median. So yes, the statistic is biased towards particular cloud types. We stated this limitation more clearly in the manuscript by mentioning the limitation in Section 3.1 and additionally in the conclusion. Apparent overcast situations which prevent a valid MIBase retrieval occur mainly in the mid latitudes over ocean and in the subtropical stratocumulus areas which you mentioned as an example (Fig. 10d of the revised manuscript).

"The authors define percentile (P) values of the MISR lowest cloud layer z distribution to obtain z_base and z_top. For z_base one would expect P=0 because z_base should be lower than any MISR-derived cloud-top height. The chosen value P=15 is motivated by the noise in MISR z, which makes sense. However, for z_top I do not understand the chosen value P=95. All MISR z values are actual estimated cloud top heights. The logical way to aggregate these is to average the individual z observations or take the median. In other words, a value P=50 seems natural. The choice of P=95 should thus be motivated."

Without further validation, we apply the 95th percentile rather than the median, as we do not want a height which might be representative for the whole area, but rather an estimate of the highest top of the cloud especially for a heterogeneous cloud top height. The focus of this study is on the z_{base} retrieval method and its validation. The use of

z_{top} retrievals serves only auxiliary purposes. It is a measure to describe the individual cloud scenes better. For instance, it allows a qualitative assessment of the algorithm's performance in dependence on cloud vertical extent. We extended the motivation of the 95th percentile at the end of Section 3.1.

"Cloudsat is, especially in combination with Calipso, arguably the most accurate source of cloud base height (as well as cloud top height) information from space. Surprisingly, Cloudsat is not mentioned at all (except for two remarks in the context of the Desmons et al. paper) in the manuscript. At the least, Cloudsat should be discussed, and it would also be good to make some comparisons with this instrument, even if direct collocations with MISR supposedly only occur at high latitudes."

We agree, that it would be very beneficial to have further data sets to compare the method to. However, CloudSat has limitations in estimating the cloud base height, in particular for low liquid clouds. It does not detect small droplets at the base of the cloud (Sassen and Wang, 2008) due to its detection limit of $\approx -28\,\text{dBz}$. Furthermore, retrievals are degraded in the ground clutter reigon (Tanelli et al., 2008; Marchand et al., 2008). Mülmenstädt et al. (2018) evaluate the 2B-GEOPROF-LIDAR product (Mace and Zhang, 2014) which uses a combination of CloudSat and CALIOP retrievals. From this product, they extracted the LIDAR only and RADAR only subsets and compared the cloud base height retrievals with ceilometer measurements similar to the reference data utilized in this study (their Fig. 9). Within their study, the RADAR does not perform as well as the LIDAR. In fact, they find a correlation of 0.265 and an RMSE of 782 m for the RADAR only subset. Therefore, we believe CloudSat would not be suitable as reference data for our study.

A comparison with Cloudsat would require the identification of collocated measurements, which you mentioned would occur at high latitudes. CloudSat has an afternoon overpass, while MISR on Terra has a morning overpass. Therefore, a comparison would also require a discussion on the impact of the temporal difference. We believe, that this paper should focus on the introduction of this new z_{base} retrieval method, its calibration and validation, statistics on the success rate, and a discussion on its global application. Further comparisons go beyond the scope of this paper.

"P1, Abstract: The abstract does not include any information on the cloud types the algorithm is applicable to. In the manuscript this information is also too limited. Does the method work for cirrus, or for deep convective clouds?"

We added that overcast cloud scenes are not included in the statistics to the abstract. Further, we added "The impacts of the cloud scene structure and macrophysical cloud properties discussed", to alert the reader, that this is an issue.

Our study does not focus on cirrus clouds. Since the MISR z retrieval method is in the visible light range, thin cirrus clouds are probably not included, since they are almost transparent. Further, we cannot validate cirrus clouds, as the METAR data does not include reliable retrievals for high clouds. Deep convection might be problematic, as the thinner cloud edge might not be as pronounced or hidden by the towering cloud.

However, deep convection should not be a major issue, even in tropical regions, because of the morning overpass time of MISR on Terra.

"P2, L8: Stephens et al. (2002) is mainly about Cloudsat. Its not the appropriate reference for CALIPSO."

Thank you for pointing this out. We changed the reference by citing Winker et al., 2010.

"P4, L4: Is aftward correct English?"

We think aftward is correct English. It is also used in the MISR product documentation. However, we will follow whatever guideline the editor suggests.

"P5, L6: measurements: what is measured?"

The signal return is measured. We rephrased the sentence.

"P5, L9: Is the value 5000 ft correct? It seems such a big jump from 100 ft and 200 ft in the two lower height categories, respectively."

Sorry, that is a typo. According to the "Automated Surface Observing System Users Guide"(National Oceanic and Atmospheric Administration, Department of Defense, Federal Aviation Administration, and United States Navy, 1998) it is 500 ft ($\approx$ 150 m).

"P5, L10: This suggests that bins is the same as clusters, which is not the case, I assume."

We delete "bins" from the sentence.

"P8, L1-3: This paragraph looks out of place here. Suggest to move it somewhere else."

We moved it to the end of Section 2.1.

"P8, L7: Suggest to replace the estimated by a typical."

Done.

"P8, L15: what are z pixels?"

It refers to the z retrievals from the MISR cloud product. We replaced "pixel" by "retrieval".

"P8, L25-26: Does this mean that the case in Fig. 4 is not included in the statistics? Isnt it a bit strange to present a case study that is not part of the selection applied furtheron?"

This case was studied before we decided which years we would use for the comparison. The main reason why it is shown, is because it illustrates the way the algorithm works and the parameters which are used. Preferentially, the presented case study should be a multi-layer case so the applied layer distinction can be illustrated as well. However, as mentioned in the manuscript, any multi-layer case will be masked out and not be included

in the selection for further processing in order to calibrate the algorithm.

"P10, L5: Fig. 4 includes only one h_gap."

It is only illustrated once. However, h_{gap} is calculated for any height gap between two z retrievals and then tested against the threshold. If it is greater than 500 m, the retrievals below and above the gap are treated as separate layers. In the case study, an apparent third layer which is about 1000 m above the top of the middle layer, is revealed in the density plot (Fig. 4 in the manuscript).

"P10, L13-14: The second layer detected by MISR has a base height between 5000 and 5500 m a.s.l. The ceilometer detects layer base heights at 853 m, 2286 m, and 7010 m a.g.l. None of these seems to match with the second MISR layer. Can you explain?"

Thank you for pointing this out. From the distribution of the z retrievals (Fig. 4 in the manuscript), we can distinguish three apparent cloud layers. The highest ceilometer retrieval seems to correspond well with the top layer (between 7000 m and 8000 m). The lowest ceilometer retrieval corresponds well with the bottom layer. We corrected the last sentence of Section 3.2 accordingly. The second ceilometer retrieval roughly matches the top of the lowest layer detected by MISR. The connection to the bottom layer detected by MISR might be an indication of a varying cloud base height within this cloud field. It could also be due to the temporal mismatch between the measurements.

"P12, Fig. 6: I assume this figure is for N=10 and R_fv=10 km. Can you add this to the caption for completeness?"

Yes. Done.

"P12, L6: N=10 seems a relatively low number and one could wonder whether P=15 makes sense for such a small N. Can you comment?"

As discussed in the manuscript, $N = 10$ is a compromise. If a higher N is chosen, the performance improves only slightly. At the same time, the algorithm would neglect more potential retrieval scenes. As addressed above, the bias increases for an increasing number of z retrievals marked high confidence cloud N_{HCC} (Fig. 8k of the revised manuscript). This indicates the potential for a bias correction. Another way to decrease the bias could be carried out by defining the selected percentile as a function of N_{HCC}. However, as mentioned above, the origin of the bias is not fully understood yet. Therefore, we like to leave such potential improvements to future studies.

"P13, Fig. 7: Can you add a color bar? Is the scale linear or logarithmic?"

The color indicates a normalized density. It is a linear scale. Contour lines are shown with the corresponding values on them. We modified the caption to point this out.

"P14, Table 4: Do you have any explanation why 2007 has 30% more valid retrievals than 2008?"

We found about 18 % more overpasses with valid z retrievals in the fields around the METAR sites. And out of those we did not have to neglect as many apparent clear sky cases. See the new section 3.4 in the manuscript for more details.

"P14, L5: Certainly the different measurement geometry (point over time versus circular area instantaneous) can cause differences. But why would this lead to a bias, and why to a negative bias of MISR in particular? Cant you tune the overall bias to near zero by increasing P?"

As we added to the manuscript, "the bias obtained in this study can have different sources: the different sample volumes of the defined MIBase cell and the ceilometer, biased MISR z retrievals, various scene characteristics." As of now, we found that the bias seems to depend strongly on N_{HCC}. Simply modifying P to tune the bias to zero is overlooking this relationship. As mentioned earlier, an adaptive P depending on N_{HCC} or an appropriate bias correction would improve the algorithm. However, this goes beyond the scope of this study.

"P15, Fig. 8: Im not sure how useful this distinction in two geometrical thickness classes is, in particular because this thickness is based upon the MISR retrieval itself."

We substituted this distinction by a more in depth discussion about the influence of the scene structure (new Section 4.1). The usage of z_{top} and Δz provides additional information which characterizes the scene structure beyond just the value of z_{base}. Therefore, it is justified, to use this information to further study the performance of the algorithm.

"P15, L3: The termination of the z_base range by the threshold height relates mostly to the lower thickness class, so it would be better to write: The smaller E for clouds with a smaller Delta z ..."

We cut this part out.

"P16, L6: Sentence ends unexpectedly."

Fixed. Thank you.

"P17, L1: The sampling size is in Fig. 9c."

Yes. Fixed.

"P19, L3: effect should be affect."

Fixed

"P19, L32: A mean difference of 500 m is quite large relative to the retrieved z_base in Fig. 11, which appears to vary between 800 and 1200 m for the selected region. Is it reasonable to assume that the model can simulate a reasonable seasonal variation of z_base if it has such a large bias?"

Yes, we think, it is reasonable. The processes responsible for defining the height of the cloud base are different from the processes which produce the seasonal cycle. Models generally underestimate the maritime boundary layer height in the stratocumulus regions. However, the radiation forcing, and the strength of the subsidence which follow an annual cycle can be represented in the model with higher accuracy leading to a realistic seasonal cycle, despite the revealed bias.

Reviewer 7 (RC6)

"As there are already five reviews available (and more referees have accepted the review of the paper) I can be short with my statements. I agree with what has been mentioned by the other reviewers with respect to the pre-conditions (cloud optical thickness, homogeneity), so I can restrict myself to comments mainly related to the ceilometersas this has not yet been covered in detail."

Thank you for your constructive feedback. We have implemented your comments as explained below.

"Section 2.1: The expression "z pixel" might be revised/improved"

We replaced "z pixel" by "z retrieval".

"Section 2.2: Please add 1-2 sentences to describe the type of ceilometers used, and the basic characteristics of the instrument and the cloud height retrieval."

We added that the ceilometers are lidar ceilometers which are operating at a wavelength of 0.9 μm to Section 2.2. Additionally, we mention that the cloud base height retrievals are derived by evaluating the vertical gradient of the backscatter profile.

"Is the very coarse vertical resolution of the METAR-messages an issue?"

Due to the rounding, the given vertical resolution of the METAR $\hat{z}_{\text{base}}$ reports is 100 ft ($\approx$ 30 m) for heights up to 5000 ft ($\approx$ 1500 m) and 500 ft ($\approx$ 150 m) between 5000 ft and 10000 ft ($\approx$ 3000 m). We expect the uncertainty of the MIBase retrievals to be larger than this (RMSE $\approx$ 400 m), so that the resolution of the METAR messages is a small contribution to the total uncertainty.

"What about using backscatter profiles from ceilometer networks, e.g. in Europe: derived cloud base heights are quite reliable and the vertical resolution is in the order of 10 meters. Please comment on this; maybe in the conclusions."

Due to the large homogeneous data set, we focus on the continental U.S. We are aware of harmonisation efforts within Europe. Therefore, we added the following sentence to the conclusion: "Within Europe, the European Cooperation in Science and Technology (COST) activity is expected to harmonize the networks of the different weather services (e.g. Haeffelin et al., 2016 and Illingworth et al., 2018, for further reading) enabling more inter comparisons in the future."

"Is the variability of the 30 s messages used to exclude certain data sets (temporal variability translated to spatial inhomogeneity [taking into account the bins of the messages])?"

As far as we know from the ASOS handbook, no filtering for inhomogeneity is performed.

"The discussion of the implications of the time period of 30 minutes for averaging could be extended."

We added the following sentences to the new Section 3.4: "The METAR reports comprise retrievals over a 30 minute period. During this time, cloud formation and cloud dissipation can alter the cloud scene and cause mismatches between MISR and METAR retrievals."

"Section 3: Better use another word for "field of view" (R_f) here: according to page 7, line 11 it has nothing to do with the optics of the radiometers onboard of MISR as one might expect."

We agree that "field of view" is inappropriate here. We changed it to "MIBase cell" throughout the manuscript. For consistency, we also modified the notation for the radius which defines the size of the MIBase cell from R_{fv} to R_{c}.

"Section 3.2: Taking into account the very poor vertical resolution of the ceilometers and the large "footprint" of the inter-comparison I feel that it is not justified to end up with a $\hat{z}_{\mathrm{base}} \approx 853\,m$ (pretending a one-meter-accuracy). Can you give an uncertainty instead of using "$\approx$"."

As stated above, the binning during the data processing of the ceilometer measurements, leads to a vertical resolution of the METAR retrievals between 100 ft (≈ 30 m) and 500 ft (≈ 150 m). This resolution should suffice for the analysis carried out in this study. The native METAR ceiling report was 2800 ft which is an integer multiple of the measurement resolution. Here we convert to SI units, which leads to values which look not round at all. To avoid the illusion of one meter accuracy, we changed that particular instance to $\hat{z}_{\mathrm{base}} = (853 \pm 15)$ m and added: "The given uncertainty solely represents the resolution of the METAR reports (Tab. 2)."

"Page 10, line 17 states that a cloud base height of 7010 m was retrieved. In section 2.2 it is stated that the ceilometers have a vertical range of up to 3700 m. Please explain."

This height was included in that particular METAR message. This can happen, because a subset of the ceilometers has a higher measurement range. In case of multiple layers, and if at least the lowest retrieval occurs within the reporting range, cloud heights outside this range can be included in the report.

"The caption of Fig. 12 could be misleading. Mention that deviations are shown right at the beginning of the text."

We edited the caption and the axis labels.

"The conclusions of the papers cited in Hannay et al. (2009) are mainly based on thermodynamics. They do not cover pbl-retrievals based on backscatter. This is however relevant for ceilometers (that are used as reference in this paper). Therefore the agreement/disagreement of ceilometer-retrievals with model results should be discussed as well: a lot of papers have recently been published focussing on the potential of ceilometers in general and the determination (and its accuracy) of the mixing layer height (or pbl)."

The reason why we are citing Hannay et al. (2009) is that they provide studies from the area we are interested in, i.e. the southeast Pacific. Their comparison to observations

based on radiosonde data and microwave radiometer retrievals shows that the models underestimate the boundary layer height in this region where stratocumulus clouds prevail. Their conclusion should not be generalized outside this area. To clarify that the study by Hannay et al. is carried out over the southeast Pacific, we updated the manuscript accordingly. We agree, that where available Lidar and ceilometer measurements would be beneficial to validate the mixing layer heights and cloud heights from models. However, we are not aware of such comparisons for this particular region.

"I agree that the MIBase can be a promising tool for remote areas, and for climatological studies with the corresponding (extended) spatiotemporal averages. Nevertheless a few comments on the benefit of the retrieval based on individual observations would be desirable, considering the large uncertainty and the missing coverage of the diurnal cycle. So combination with ground based ceilometer networks (where available) should be envisaged, especially as ceilometers are a very direct and accurate approach (no calibration required, continuous operation) for z_{base}-retrievals."

We added "Depending on the application, the MIBase uncertainty and the missing coverage of the diurnal cycle can be a limitation. However, in combination with ceilometer networks, both temporal and spatial patterns can be investigated." to the conclusion.

Thank you for your time and input!

Sincerely

Christoph Böhm for the authors

Cloud base height retrieval from multi-angle satellite data

Christoph Böhm[1], Odran Sourdeval[2], Johannes Mülmenstädt[2], Johannes Quaas[2], and Susanne Crewell[1]

[1]Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany
[2]Institute of Meteorology, University of Leipzig, Leipzig, Germany

Correspondence to: Christoph Böhm (c.boehm@uni-koeln.de)

Abstract. Clouds are a key modulator of the Earth energy budget at the top of the atmosphere and at the surface. While the cloud top height is operationally retrieved with global coverage, only few methods have been proposed to determine cloud base heights (z_{base}) from satellite measurements. This study presents a new approach to retrieve cloud base heights using the Multi-angle Imaging SpectroRadiometer (MISR) on the Terra satellite. It can be applied if some cloud gaps occur within the chosen distance of typically 10 km. The MISR cloud base height (MIBase) algorithm then determines z_{base} from the ensemble of all MISR cloud top heights retrieved at a 1.1-km horizontal resolution in this area. MIBase is first calibrated using one year of ceilometer data from more than 1500 sites within the continental United States of America. The 15th percentile of the cloud top height distribution within a circular area of 10 km radius provides the best agreement with the ground-based data. The thorough evaluation of the MIBase product z_{base} with further ceilometer data yields a correlation coefficient of about 0.66 ~~.~~ demonstrating the feasibility of this approach to retrieve z_{base}. The impacts of the cloud scene structure and macrophysical cloud properties are discussed. For a three year period, the median z_{base} is generated globally on a $0.25° \times 0.25°$ grid. ~~It shows~~ Even though overcast cloud scenes and high clouds are excluded from the statistics, the median z_{base} retrievals yield plausible results in particular over ~~sea~~ ocean as well as for seasonal differences. The potential of the full 16 years of MISR data is demonstrated for the southeast Pacific revealing ~~inter-annual~~ interannual variability in z_{base} in accordance with reanalysis data. The global cloud base data for the three year period (2007–2009) are available at https://doi.org/10.5880/CRC1211DB.19.

1 Introduction

~~Clouds~~ As Boucher et al. (2013) state in the IPCC Assessment Report 5, clouds and aerosols continue to contribute the largest uncertainty to estimates and interpretations of the Earth's changing energy budget ~~Boucher et al. (IPCC 5th assessment report, 2013)~~. To describe the effect of clouds on the radiation energy budget, the geometric thickness, the vertical location of clouds and, therefore, the cloud base height (z_{base}) are crucial parameters. Furthermore, long term observations of cloud heights would be beneficial to assess the contribution and the response of clouds to climate change. z_{base} is a key parameter for the radiative energy budget at the Earth surface. z_{base} may also have an impact on ecosystems which are supplied with water by the immersion of clouds (Van Beusekom et al., 2017). Aviation is another field which benefits from information on z_{base}.

Various methods to retrieve the z_{base} have been proposed applying different physical concepts, such as active measurements, spectral methods, approaches using an adiabatic cloud model (e.g., Goren et al., 2018), and in-situ measurements.

From the ground, the most accurate and well-established method to derive z_{base} is the backscatter information from a lidar ceilometer, also providing crucial information on visibility for aircraft safety. Thus, ceilometers are employed at airports. Their number has increased in particular in Europe and North America during the past couple of years. A dedicated web page hosted by the Deutscher Wetterdienst shows the distribution of ceilometer stations around the world (http://www.dwd.de/ceilomap). Radiosondes provide in-situ measurements of thermodynamic variables. Costa-Surós et al. (2014) compare different methods to infer z_{base} from radiosonde data. For the best of these methods, 67 % of the considered profiles agree with the utilized reference data regarding number of cloud layers and height category (distinguished are low, middle and high). Cloud radar transmits microwave radiation to derive vertical profiles of radar reflectivity. However, this signal strongly depends on the particle size. Therefore, the occurrence of a few drizzle drops can mask cloud base. ~~Radiosondes~~ Measurements with radiosondes and cloud radars are even less ~~frequent than ceilometersand no global coverage can~~ common than ceilometers, global coverage cannot be achieved from the ground today.

From space, active measurements are carried out by CALIOP (Cloud Aerosol Lidar with Orthogonal Polarization) on the CALIPSO (Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations) satellite ~~(Stephens et al., 2002)~~(Winker et al., 2010) . A valid retrieval of the z_{base} can only be ensured if the signal of CALIOP reaches the Earth's surface, which is only possible in case of low optical thickness. Optically thick clouds will lead to attenuation of the signal. The spatial coverage is limited to the narrow laser beam of CALIOP. The CALIOP cloud base determination has been revisited by Mülmenstädt et al. (2018). They developed an algorithm to extrapolate cloud base retrievals for thin clouds into locations where the CALIOP signal is attenuated within a thicker cloud before it reaches the cloud base.

Passive measurements in the near-infrared exploiting spectral information have been proposed by Ferlay et al. (2010). They suggest an approach to infer the cloud vertical extent from multi-angular POLDER (POLarization and Directionality of the Earth's Reflectances) oxygen A-band measurements. As they point out, the penetration depth of photons into a cloud, and, hence, the height of the reflector, depends on the cloud vertical extent and the viewing geometry. Exploiting the different viewing angles provided by POLDER, Desmons et al. (2013) apply this approach to infer the vertical position of clouds. Their comparison to retrievals from the cloud profiling radar on CloudSat and CALIOP shows that this method works best for liquid clouds over ocean with a retrieval bias of 5 m and a standard deviation of the retrieval differences of 964 m. However, this approach has not been carried out operationally yet. Moreover, an estimate of the cloud top height is required to retrieve the cloud base height from the cloud vertical extent, which introduces additional uncertainty.

Meerkötter and Zinner (2007) suggest a method to derive z_{base} of convective clouds which are not affected by advective motions. An adiabatic cloud model incorporating measurements of cloud optical depth and effective radius is used to calculate the geometric extent of the cloud from the retrieved cloud top height. By introducing a subadiabatic factor, Merk et al. (2016) investigate the adiabatic assumption in more detail. By additionally introducing a factor into the calculations, they account for subadiabaticity due to entrainment of dry air through the cloud edges. As a reference, the cloud vertical extent is derived as the difference between z_{top} (radar) and z_{base} (ceilometer) from ground based measurements. The authors conclude that for their two year data set neither the assumption of an adiabatic cloud nor the assumption of a temporally constant subadiabatic factor is fulfilled.

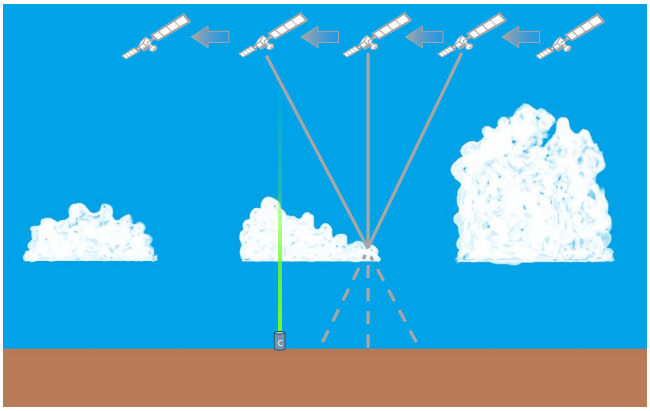

Figure 1. Schematic depiction of a cloud field observed from different viewing angles during the satellite overpass. Ceilometers, here represented as a cylindrical box, provide ground-based measurements of cloud base heights which can be used as reference.

Lau et al. (2012) suggest a new approach to determine z_{base} utilizing the Multi-angle Imaging SpectroRadiometer (MISR) on the Terra satellite. For a preliminary case study, they chose the observations from island Graciosa, Azores, Portugal, for which they compared cloud top height (z) retrievals from MISR to collocated and coincidental ~~Lidar~~ lidar measurements. Under the assumption that the cloud vertical extent varies horizontally within the cloud, they retrieve z_{base} by identifying the lowest cloud top height in the height profile provided by MISR. The reference cloud base height ($\hat{z}_{\text{base}}$) is retrieved from the ~~Lidar~~ lidar signal by visual inspection of the backscatter coefficient in a time-height cross section over a period of about five hours. They selected 12 cases which show a promising agreement between MISR and ~~Lidar~~ lidar retrievals.

We build on the approach proposed by Lau et al. and develop an automatic retrieval method to derive z_{base} from MISR measurements. Parameters employed in the retrieval scheme are derived from coincident ceilometer measurements over one year in the continental United States of America (USA). The performance of the z_{base} algorithm is demonstrated by an evaluation with ceilometer over a longer time period and the potential for application on the global scale and for longer time series is explored.

The paper is structured as follows. In Section 2, the utilized data from MISR and from ceilometers are described. Section 3 introduces the new retrieval method along with a case study for illustration. In Section 4, the evaluation of the algorithm against the ceilometer measurements is shown and the effect of the cloud vertical extent on the performance of the algorithm is discussed. Section 5 includes two applications of the algorithm: the median z_{base} is presented globally for a three-year period, and regionally over the southeast Pacific for a 16-year period. Finally, Section 6 concludes the study.

2 Data

2.1 MISR cloud product

MISR is carried on board the Terra satellite and provides sun-synchronous (equatorial overpass at around 10:30 ~~a.m. LT~~local solar time) global products of cloud properties with a 1.1 km horizontal resolution. With an across-track swath width of 380 km,

MISR takes two (poles) to nine (equator) days for repeated observations of the same site. The MISR Level 2TC Cloud Product (MIL2TCSP; Diner, 2012; Moroney and Mueller, 2012; Mueller et al., 2013) is used in this study to provide retrievals of cloud top height z and a stereo-derived cloud mask (SDCM). Three years of global data (~~2007–2009~~2007–2009) are utilized here. The MISR Ancillary Geographic Product (Bull et al., 2011) is additionally used to assign corresponding spatial coordinates and the average scene elevation for each pixel. Here, we give a brief summary on how the operational MISR z product is derived. More in-depth descriptions can be found in Moroney et al. (2002) and in Marchand et al. (2007).

A cloud field is schematically depicted in Fig. 1. MISR hosts cameras providing a total of nine viewing angles. Besides the nadir viewing camera (0°), there are four forward and four aftward viewing cameras set up at 26.1°, 45.6°, 60.0° and 70.5° angles, respectively. During an overpass, each camera of MISR records the reflected radiances at its particular viewing angle. A pattern matching routine which compares the radiances recorded at a wavelength of 670 nm identifies equal cloud features in the images of the different viewing angles. Pixels with the least deviation from each other are matched. This way, a detected cloud feature is observed from multiple satellite positions with its respective time and viewing angle. If at least three images can be attributed to the same cloud feature, the cloud motion vector along with the horizontal and vertical position of the cloud feature can be inferred geometrically. This process is not sensitive to absolute values of the radiances so that this retrieval method is not sensitive to calibration.

The cloud motion vector is determined at a ~~70.6 km~~ 17.6 km resolution. For each of these coarser grid boxes, the cloud motion vector is then used to determine z at 1.1 km resolution, which is carried out for two camera pairs individually: one pair (FWD) consisting of the nadir and 26.1° forward viewing cameras and the other (AFT) consisting of the nadir and 26.1° aftward viewing cameras. This way, two z values for the same location are available, and the mean of the two values yields the final z. In case only one camera pair provides a valid z, it is taken as the final z at its specific location. To derive the stereo-derived cloud mask, the two individual z values undergo the following comparison. The retrieval of each camera pair is classified as surface or cloud retrieval according to the threshold height $h_{\min}$ (Equation 1). This is Equation 59 in the Algorithm Theoretical Basis documentation by Mueller et al. (2013), where the threshold height for flat terrain H_{SDCM} is 560 m, H is the terrain height and σ_{h} is the variance of the the terrain height listed in the Ancillary Geographic Product. Within the MISR Level 2TC Cloud Product, the cloud top height and the stereo-derived cloud mask are also provided without wind correction. Here, we use the the wind corrected data sets.

$$h_{\min} = H_{\mathrm{SDCM}} + H + 2\sigma_{\mathrm{h}} \tag{1}$$

The use of two camera pairs allows ~~to attribute~~ attribution of a confidence level to the retrieved z. If the mean of the two values is above or below the threshold, the pixel will be classified as cloud or surface, respectively. If only one camera pair provides a valid retrieval, it is tested against the threshold and classified accordingly. In case only one camera pair provides a valid retrieval and in case of two valid retrievals which disagree upon their individual classification, the z~~pixel~~ retrieval is marked low confidence. If two retrievals are available which agree upon their individual classification, the z~~pixel~~ retrieval is

marked high confidence. Any other case leads to a non-retrieval. Table 1 summarizes possible combinations of retrievals from the two camera pairs and their corresponding attribution within the stereo-derived cloud mask.

MISR z is given in meters above the World Geodetic System 1984 (WGS 84) surface. To calculate the height above ground level, we subtract the average scene elevation which is provided within the Ancillary Geographic Product for each pixel.

The MISR z product is expected to be superior to z products from other passive instruments. It does not depend on any auxiliary data and it is not sensitive to calibration. Therefore, it is not granted that the application of MIBase to $z$retrieved by techniques other than the geometric approach would yield similar results.

Table 1. Classification scenarios of MISR retrievals. The cloud height obtained using the nadir and the 26.1° forward viewing camera pair (denoted by FWD) and the cloud height obtained using the nadir and the 26.1° aftward viewing camera pair (AFT) are tested against the threshold height $h_{\min}$ (Equation 1) individually and then compared to one another to determine the Stereo-Derived Cloud Mask (SDCM) attribute.

condition	**SDCM attribute**
FWD and AFT above threshold	high confidence cloud
FWD and AFT disagree, mean(FWD, AFT) above threshold	low confidence cloud
only one camera pair, retrieval above threshold	low confidence cloud
FWD and AFT below threshold	high confidence surface
FWD and AFT disagree, mean (FWD, AFT) below threshold	low confidence surface
only one camera pair, retrieval below threshold	low confidence surface

2.2 METAR data

Aerodrome routine meteorological reports (METAR) (WMO; World Meteorological Organization, 2013) contain weather observations at airports worldwide, including measurements of z_{base}. METARs from airports from the continental USA provide z_{base} determined by the Automated Surface Observing System (ASOS; National Oceanic and Atmospheric Administration, Department of Defense, Federal Aviation Administration, and United States Navy, 1998). ASOS utilizes ~~ceilometers which~~ lidar ceilometers which operate at a wavelength of 0.9 µm and have a vertical range of 12000 ft (≈ 3700 m). ~~The ceilometer provides continuous measurements with~~ Cloud base heights are routinely retrieved by evaluating the vertical gradient of the detected backscatter profile with a temporal resolution of 30 seconds. These individual ~~measurements~~ retrievals are stored in different bins by rounding to the nearest 100 ft (≈ 30 m) for heights between the surface and 5000 ft (≈ 1500 m)~~and~~ ; to the nearest 200 ft (≈ 60 m) for heights between 5000 ft (≈ 1500 m) ~~,~~ and 10000 ft (≈ 3000 m); and to the nearest ~~5000 ft (≈ 1500 m~~500 ft (≈ 150 m) for heights above 10000 ft (≈ 3000 m). If there are more than five bins filled with measurements during a 30 minute period, the cloud heights are clustered into layers until only five ~~bins or~~ cluster remain. Finally, all cluster heights are rounded according to the rules given in Tab. 2. The lowest three layers are passed on to the METAR message.

Table 2. The ceilometer $\hat{z}_{\text{base}}$ retrievals are rounded to different values depending on their height window according to ~~the National Oceanic and Atmospheric Administration, Department of Defense, Federal Aviation Administration, and United States Navy (1998)~~ ASOS User Guide (National Oceanic and Atmospheric Administration, Department of Defense, Federal Aviation Administration, and United States Navy). The values are originally given in feet and are converted to meters here.

height [ft]	**rounded to nearest value [ft]**	**rounded to nearest value [m]**
< 5000	100	30.5
5000 to 10000	500	152
> 10000	1000	305

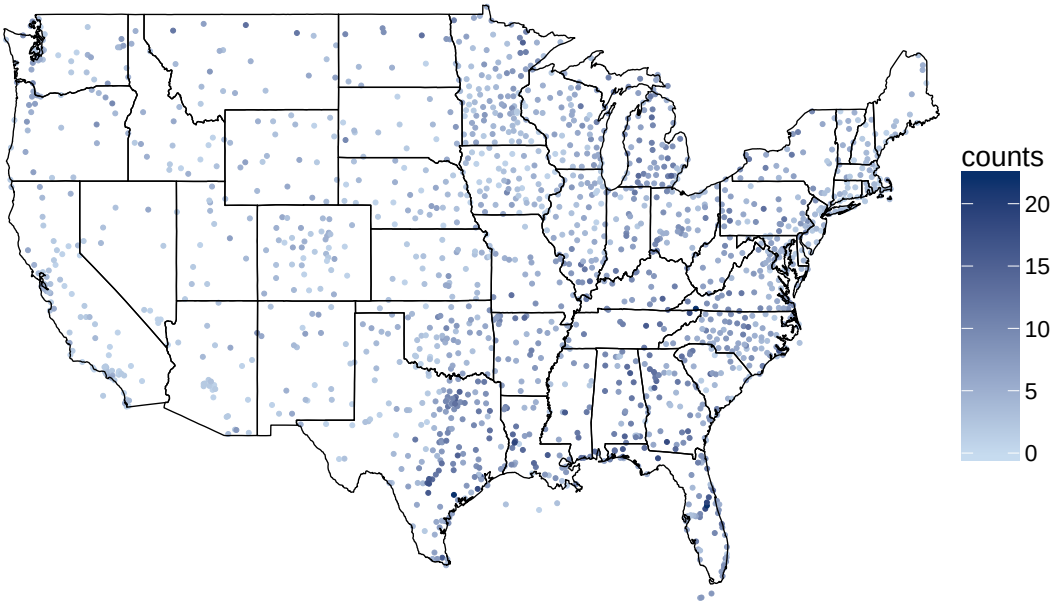

Figure 2. Locations of ceilometer stations utilized in this study across the continental ~~United States of America~~USA. Data from these stations for the years 2008 and 2007 are used for the calibration of the z_{base} retrieval algorithm and a subsequent evaluation, respectively. Blue shading indicates the number of valid coincidental retrievals from MISR and ceilometers which have been utilized for the calibration (year 2008) and are within the constraints described in the text.

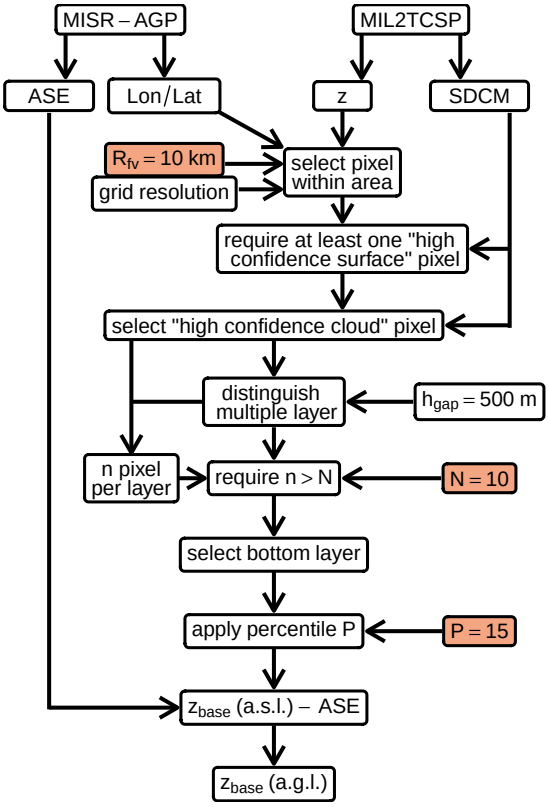

Figure 3. Flow chart of the z_{base} retrieval algorithm. MISR's MIL2TCSP cloud product provides z and the Stereo-Derived Cloud Mask (SDCM). MISR's Ancillary Geographic Product (MISR-AGP) provides the Average Scene Elevation (ASE) and the longitude and latitude coordinates for each pixel. Starting from these products, the depicted processing steps are undergone to derive z_{base}. The parameters which have been optimized during the calibration are highlighted in orange.

We extract the ceilometer cloud base height $\hat{z}_{\text{base}}$ from METAR data for a total of 1510 ceilometer sites around the continental USA to benefit from the homogeneity of the automated measurements and the standardized reporting range. $\hat{z}_{\text{base}}$ serves as reference data to which the z_{base} derived from the satellite cloud heights is compared. First, METAR data from 2008 are used to estimate parameters used in the z_{base} retrieval algorithm to create the MISR Cloud Base height algorithm (MIBase). Second, to validate the "tuned" algorithm, METAR data from 2007 are applied for a statistically independent comparison. For a total of 1510 ceilometer stations, collocated and coincidental satellite-based z_{base} retrievals could be found (see below for exact definition). A distribution of the locations can be seen in Fig. 2.

3 Cloud Base height retrieval

The MISR Cloud Base height retrieval (MIBase) algorithm, which derives z_{base} from the MISR z product, is developed and calibrated with collocated METAR data for defining the involved parameters and preconditions. The first section of this chapter introduces the retrieval principle on the basis of a case study. By comparison with ~~ceilometer measurements (METAR)~~ METAR ceilometer measurements from 2008, parameters used within MIBase are estimated, namely the radius R_c of the ~~field of view~~MIBase retrieval cell, the minimum number of valid cloud pixel N and the percentile P of the z distribution.

3.1 Method

We assume that the information on the z_{base} is included in the distribution of the z retrievals from the MISR cloud product for a specific area of limited size. This assumption is valid in a cloud scene with a homogeneous z_{base} and a heterogeneous z similar to the one schematically depicted in Fig. 1. Especially at the edge of a cloud where the cloud is thinner, z can serve as a proxy for z_{base}. To ensure that the thinner ~~part~~ edge of the cloud is within the observed ~~field of view~~MIBase retrieval cell, the considered area needs to be large enough and the cloud field needs to be broken. The inherent assumption of a homogeneous z_{base} over a certain area presupposes a horizontally constant lifting condensation level. This is in particular given in a well mixed boundary layer or a homogeneous air mass away from the proximity of a frontal zone, where advective motion could introduce temperature or humidity gradients across the horizontal plane.

~~The MISR product is expected to be superior to products from other instruments. It does not depend on any auxiliary data and it is not sensitive to calibration, as mentioned in Section 2.1. Therefore, it is not granted that the application to from other satellite instruments would yield similar results.~~

In order to derive z_{base} from the z product, the following steps, which are outlined in Fig. 3, are undertaken. First, a ~~field of view~~ retrieval cell has to be defined. For the comparison to the ceilometer measurements, we consider a circular area with the radius R_c around its midpoint at a ceilometer station. In order to estimate the magnitude of R_c, we consider the following: METAR $\hat{z}_{\text{base}}$ retrievals are representative for a time window of 30 minutes. Within this time window and at ~~the estimated~~ a typical wind speed of approximately $10\,\text{ms}^{-1}$, a cloud would shift its position about 20 km in the wind direction. Therefore, the magnitude of R_c should be on the order of kilometers. The impact of R_c on the retrieved z_{base} and, therefore, the deviation from the ceilometer $\hat{z}_{\text{base}}$ is discussed below. When we apply the algorithm to retrieve a global estimate of z_{base}, we use a regular lat-lon grid of 0.25° (cf. Section 5). This grid size corresponds to a meridional length of the grid boxes of about 28 km and a zonal length ranging between 25 km (25°N) and 18 km (50°N), taking the continental U.S.A. as an example. A greater ~~field of view~~ MIBase cell increases the chance of seeing the thinner part of the cloud. This could lead to a more realistic z_{base} retrieval. In turn, for a smaller ~~field of view~~ MIBase cell the assumption of a homogeneous z_{base} is more realistic.

For each grid ~~box or circular field of view around the ceilometer station~~cell or circular MIBase cell, the enclosed z~~pixels~~ retrievals from the MISR cloud product are processed further. MIBase only selects those z~~pixels which are~~ retrievals which are marked high confidence cloud (hcc) ~~pixels~~ according to the stereo-derived cloud mask. A consideration of retrievals marked low confidence cloud ~~pixels~~ has shown a decrease of the correlation with the ceilometer $\hat{z}_{\text{base}}$. An example of a cloud field with

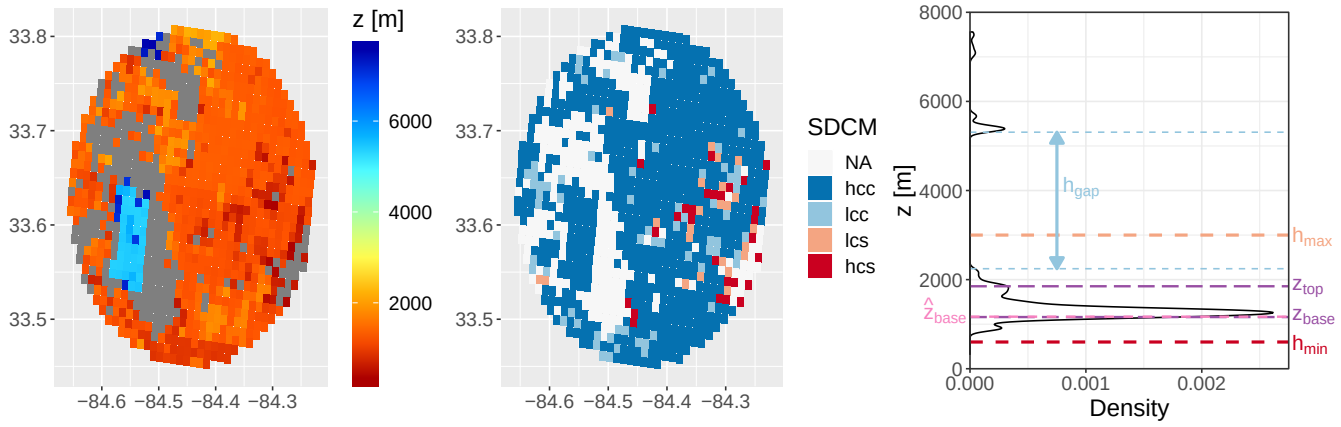

Figure 4. MISR observations within a 20 km radius within the vicinity of Atlanta, Georgia, USA (ICAO:KATL) on 21 August 2015 at around 16:30 UTC. Left: z. Middle: Corresponding Stereo Derived Cloud Mask (SDCM) distinguishing non-retrievals (NA), high confidence cloud (hcc), low confidence cloud (lcc), low confidence surface (lcs) and high confidence surface (hcs). Right: Density of z measurements with illustration of certain parameters: height between two layers (h_{gap}) which is the height difference between the highest retrieval of the bottom layer and the lowest retrieval of the next higher layer (dashed blue lines), upper cut-off height (dashed orange) for z_{base} retrievals (h_{max}) which is based on the ceilometer granularity, lower cut-off height (dashed red), which is based on the MISR threshold height to distinguish between cloud and surface retrieval (h_{min}), and the ceilometer retrieval $\hat{z}_{\mathrm{base}}$ from 16:52 UTC (dashed pink). z_{top} and z_{base} (dashed purple) are inferred by applying the 15th and 95th percentile to the distribution of z of the lowest cloud layer, respectively. Heights are above sea level.

z~~pixels~~ retrievals and the corresponding stereo-derived cloud mask for 21 August 2015 at the International Airport of Atlanta, Georgia, USA, is presented in Fig. 4 (left, middle).

For some scenes, the distribution of z reveals extended height ranges with no z retrievals between two or more local maxima. Such cases suggest ~~multilayer~~ multi-layer cloud scenes if the apparent gap between adjacent z retrievals is of sufficient size. If such a gap h_{gap} is greater than 500 m, the algorithm distinguishes between the cloud layer above and below the gap (cf. Fig. 4 (right) for the aforementioned example). The value for this threshold has been chosen to be close to the specified accuracy of MISR (560 m). By evaluating different vertical cloud layers individually, a z_{base} retrieval for each layer can be derived. Since for most ~~application~~ applications the lowest z_{base} is of interest, the lowest detected cloud layer is processed here. For the comparison with $\hat{z}_{\mathrm{base}}$, we restrict ourselves to scenes for which MISR detects only one cloud layer.

The occurrence of a broken cloud field is ~~the basic assumption~~ a basic requirement of MIBase. Therefore, at least one ~~pixel labeled "~~$z$retrieval marked high confidence surface ~~"~~ needs to be within the ~~field of view~~MIBase cell. A complete cloud cover or a high rate of non-retrievals can prevent this ~~criteria~~ criterion from being met. Both scenarios suggest doubtful z_{base} retrievals. Hence, they are not considered.

For each grid ~~box or circular area around~~ cell or circular cell surrounding the ceilometer station, z_{base} is diagnosed from the height distribution of z using a certain percentile P. In principle, P should be as low as possible. However, as a certain

measurement noise ~~in the field of view~~ is expected and a robust result should be achieved, a choice substantially larger than zero is necessary. Another parameter which describes the distribution of z for each scene is the number of valid ~~cloud pixels~~ $z$retrievals marked high confidence cloud n. A higher n implies a higher observed cloud cover within the ~~field of view~~MIBase cell. In order to take a meaningful percentile of the z distribution, a minimum $n > N$ is required. A cloud which is horizontally more extended (higher cloud cover) is more likely to pass over the ceilometer, so that there is a higher chance that both instruments observe the same cloud. Therefore, the deviation of z_{base} from $\hat{z}_{\text{base}}$ is expected to decrease for a higher n. The impact of the threshold for N is studied later on.

For certain applications, the cloud vertical extent Δz might be of interest. Therefore, an estimate of the cloud top height z_{top} is required. In principle, $P = 100$ should yield the highest point of the cloud. However, analogously to the retrieval of z_{base}, a certain ~~variability in the field of view~~ measurement noise is expected, so that P is not chosen to be the extreme value. ~~Here, we choose $P = 95$ to a meaningful choice for~~ Without further validation, we apply the 95th percentile rather than the median, as we do not want a height which might be representative for the whole area, but rather an estimate of the highest top of the cloud especially for a heterogeneous cloud top height to estimate Δz at its most extensive point.

3.2 Case study

One of the ~~utilized~~ utilised ceilometer stations is located at the Hartsfield–Jackson Atlanta International Airport. To illustrate the functionality of the presented algorithm, we investigate a particular MISR overpass over this station on 21 August 2015 at around 16:30 UTC. Figure 4 shows the z retrievals for all pixels which are within the circular ~~field of view R_{fv}~~MIBase cell defined by R_{c}. Here, we exemplarily use ~~$R_{\text{fv}} = 20\,\text{km}$~~ $R_{\text{c}} = 20\,\text{km}$ with its midpoint at the ceilometer station. z is given above the WGS 84 surface, which is approximately equal to sea level. The spatial distribution shows a low cloud layer with z between 800 m and 2000 m, which covers most of the area. Another cloud layer appears between 5 km and 6 km. Some pixels with heights above 7 km indicate the presence of a third layer (Fig. 4, left). For a few pixels, MISR was not able to determine z. This might be due to the viewing geometry. A retrieval requires valid images from two different cameras, one camera viewing nadir and the other viewing at a 26.1° angle. In the case studied here, the most missing retrievals are closely attached to high clouds which might lead to shading effects (Fig. 4, middle).

The density of the z distribution shows the aforementioned three cloud layers. They are distinguished according to the threshold value for h_{gap} (Fig. 4, right) as illustrated for the bottom and middle layer. For the bottom layer, which is selected for further processing, the number of ~~hcc pixels~~ $z$retrievals marked high confidence cloud is determined to be $n = 621$. This number is well above the threshold N which is defined later. z_{base} is then calculated using $P = 15$ as the preliminary percentile of the z distribution. This yields $z_{\text{base}} \approx 1160\,\text{m}$ above the WGS 84 surface. The mean average scene elevation for the given area is subtracted from the retrieval to obtain $z_{\text{base}} \approx 927\,\text{m}$ above ground level. The closest METAR report for this day is from 16:52 UTC. Three heights were reported at 2800 ft (≈ 853 m), 7500 ft (≈ 2286 m) and 23000 ft (≈ 7010 m) above ground level. By adding the station elevation (315 m), the corresponding height above sea level is obtained. This yields $\hat{z}_{\text{base}} \approx 1168\,\text{m}$ and is denoted in Fig 4 (right). In conclusion, using the preliminary values for P the z_{base} retrieval from MISR is about 927 m above ground level which is 74 m higher than the ceilometer retrieval ($\hat{z}_{\text{base}}$~~≈ 853 m). Note, that the second layer detected~~

= (853 ± 15) m). The given uncertainty solely represents the resolution of the METAR reports (Tab. 2). Note that the third layer detected around 7000 m by MISR has also been detected by the ceilometer~~with rather good agreement between the two~~.

3.3 Parameter optimization

For each considered ceilometer station (Fig. 2), collocated and coincidental MISR overpasses from the year 2008 are identified. ~~Then the algorithm is applied as it is described exemplarily~~ The algorithm is then applied as described in the case study (Sec. 3.2 to retrieve z_{base}. All pairs of MIBase z_{base} and ceilometer ~~measured cloud base height~~ $\hat{z}_{\mathrm{base}}$ are evaluated to investigate the influence of R_{c}, N and P on the performance of the z_{base} retrieval algorithm and to estimate the ~~best suited~~ most suitable values. For this purpose, the following statistical measures are considered: ~~The~~ the slope and intercept of a linear regression, which are ideally 1 and 0, respectively~~, ;~~ the Pearson correlation coefficient r (ideally unity)~~, the Root Mean Square Error ;~~ the root mean square error (RMSE) E defined as

$$E = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(z_{\mathrm{base},i} - \hat{z}_{\mathrm{base},i}\right)^2}\,; \tag{2}$$

and the retrieval bias B defined as

$$B = \frac{1}{n}\sum_{i=1}^{n}\left(z_{\mathrm{base},i} - \hat{z}_{\mathrm{base},i}\right). \tag{3}$$

Table 3. Slope, intercept, correlation coefficient r, RMSE E, bias B and number of samples n_s resulting from comparing z_{base} and $\hat{z}_{\mathrm{base}}$ retrievals for different radii of the MISR circular area around the ceilometer stations. These values are obtained for the year 2008 applying a required minimum number of cloud pixels of $N = 10$ and the 15th percentile to the z distribution.

R_{c} [km]	slope	intercept [m]	r	E [m]	B [m]	n_{s}
5	0.65	371	0.66	392	-71	3059
10	0.62	412	0.66	404	-75	5120
15	0.60	433	0.65	413	-77	6140
20	0.58	464	0.63	423	-74	6895
30	0.54	515	0.60	437	-71	7772

MISR can only detect clouds above the threshold height according to Equation 1. To prevent this obvious limitation from introducing a bias into the statistics, we only consider cloud scenes for which the ceilometer retrieval is above h_{min}. In addition, only z_{base} retrievals below a maximum height h_{max} of 3000 m are considered to focus on a cloud range for which the ceilometer retrievals are ~~granulated finer~~ more finely granulated (below 10000 ft according to Tab. 2).

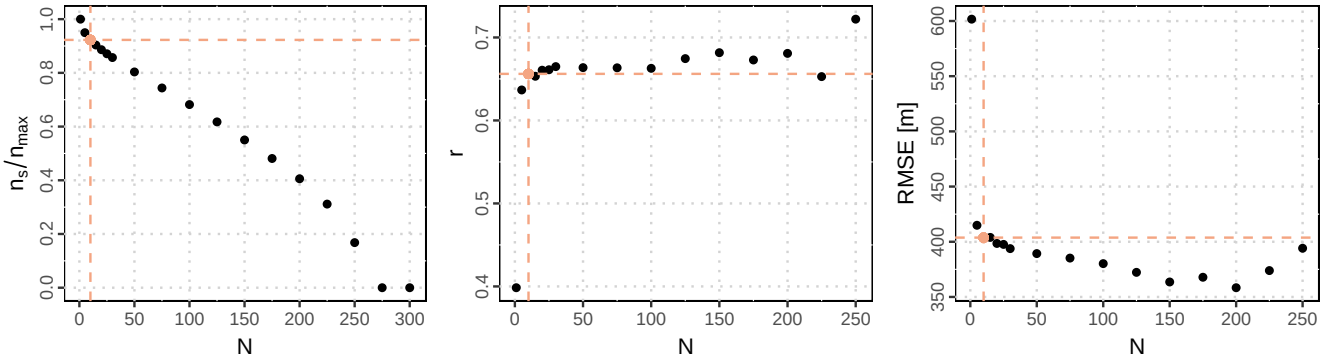

Figure 5. Evaluation of minimum number of valid pixels N within a cloud layer detected by MISR for the year 2008. Left: The normalized number of events $\frac{n_s}{n_{max}}$ for which z_{base} and $\hat{z}_{base}$ could both be retrieved. n_{max} is the maximum number of events, which is found for $N = 1$. Middle: The linear correlation coefficient r between z_{base} and $\hat{z}_{base}$. Right: The RMSE between z_{base} and $\hat{z}_{base}$. MISR z_{base} is retrieved using the 15th percentile of the z distribution for a 10 km radius around the individual ceilometer measurements. The chosen value for N is highlighted in orange. For further details see text.

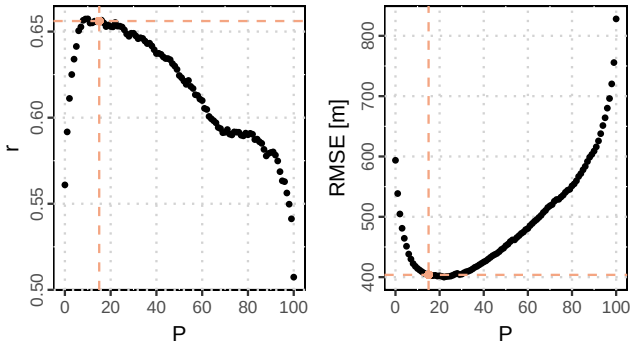

Figure 6. Evaluation of the percentile P which is applied to retrieve z_{base} from the distribution of z for the year 2008 ~~,~~with $N = 10$ and $R_c = 10$km. Left: The linear correlation coefficient r between z_{base} and $\hat{z}_{base}$. Right: The RMSE between z_{base} and $\hat{z}_{base}$. The chosen value for P is highlighted in orange.

First, we investigate the influence of the ~~field of view chosen to match MISR and ceilometer measurements~~size of the MIBase cell on the comparison of MIBase and ceilometer retrievals. For this purpose, ~~R_{fv}~~ R_c is varied between 5 and 30 km while the other parameters are set to the preliminary values $P = 15$ and $N = 10$. With a decreased ~~radius~~R_c, the correlation between z_{base} and $\hat{z}_{base}$ increases and E decreases (Tab. 3). This is to be expected as the representativity should increase. However, for a lower R_c, the retrieval algorithm encounters more situations where at least one of the requirements (at least one high confidence surface pixel is visible ~~,~~and at least 10 valid cloud pixel per layer) cannot be fulfilled, as the decrease in the total number of retrievals indicates. The better agreement between z_{base} and $\hat{z}_{base}$ for lower R_c might be due to a relatively larger overlap of the measurement sampling areas of the two instruments and to a better ~~fulfillment~~ fulfilment of the assumption of a homogeneous

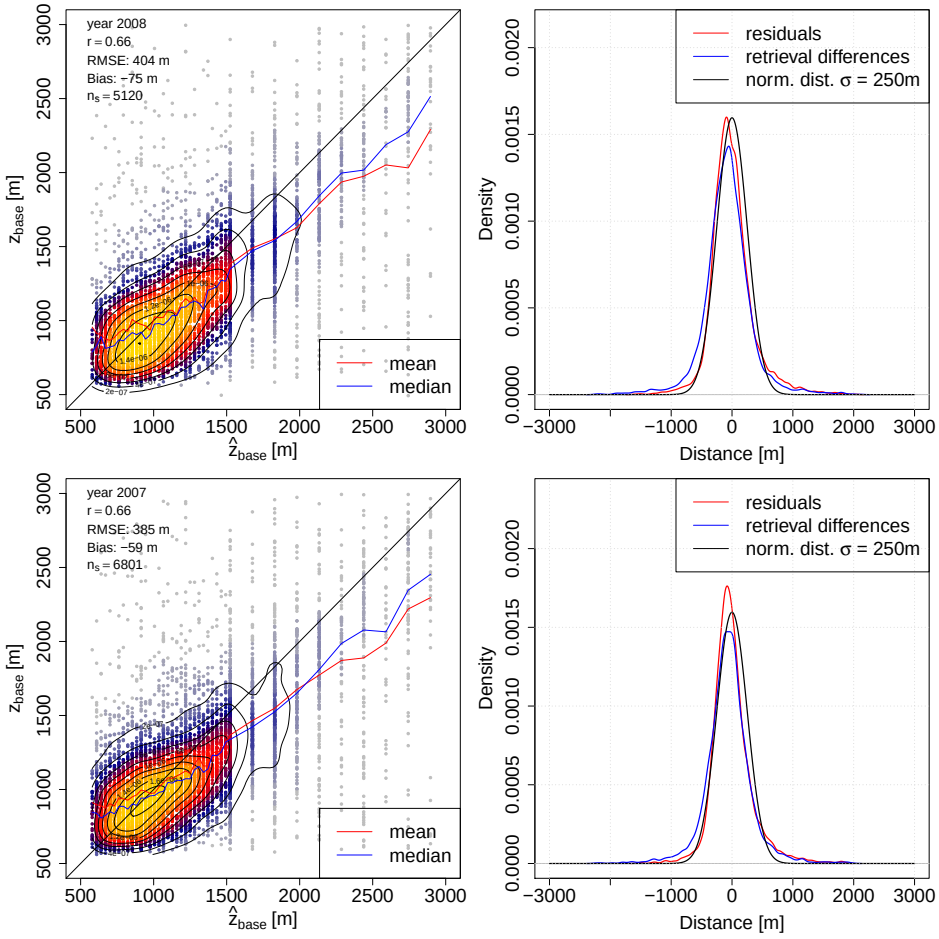

Figure 7. Left: Joint density of z_{base} and $\hat{z}_{\text{base}}$ for the year 2008 (top) which is used to estimate parameters of the algorithm and for the year 2007 (bottom) which is used to validate the stability of the algorithm with the estimated parameters. The value of the normalized density is indicated by color (maximum values in light yellow) and contour lines with corresponding values on them (linear scale). For each ceilometer height bin the mean (red) and median (blue) of the MISR z_{base} is shown. Right: Probability density functions of the residuals after a linear fit (red), the retrieval differences (blue) and a normal distribution with a standard deviation of 250 m (black).

z_{base} over smaller areas. For further evaluation, a radius of 10 km is chosen as a compromise between a good agreement in terms of ~~r^2~~ r and E and without having to discard too many retrieval scenes.

Second, the effect of the minimum number of of valid z_{base} retrievals is studied which strongly limits the number of samples for the comparison (Fig. 5). With increasing N, initially a slight increase to $N = 10$ improves the correlation between z_{base} and $\hat{z}_{base}$ and E significantly to a correlation coefficient of about 0.66. A further increase only yields slight improvement of the correlation and E. This slight increase can be explained by the elimination of more complex scenes from the comparison. However, for a higher N the trade off is a lower total number of z_{base} retrievals. For instance, for $N = 50$ only 80 % of possible retrievals yield a valid z_{base} (Fig. 5, left). Therefore, we select $N = 10$.

Finally, we consider the percentile threshold used to diagnose z_{base} from the z distribution. Figure 6 shows an evaluation of different percentiles which are applied to derive z_{base}. Percentiles between the 10th and the 15th give the best correlation. The lowest E is achieved for percentiles between the 15th and the 25th. Therefore, $P = 15$ is chosen for further processing. The fact that very clear and ~~localized~~ localised minima (maxima) for E (r) are found supports the hypothesis that the z distribution contains information on z_{base}.

In summary, the comparison yields the estimated parameters ~~for the field of view~~ $R_c = 10$ km, the minimum number $N = 10$ and the percentile $P = 15$. While the latter two are kept fixed in MIBase, R_c is ~~optimized~~ optimised for the intercomparison with point data, i.e. ceilometer measurements. The algorithm can also be applied to larger grids. However, no data for validating extended areas are available.

Table 4. Slope, intercept, correlation coefficient r, RMSE E, bias B and number of retrievals n_s resulting from a comparison of z_{base} and $\hat{z}_{base}$ for data obtained 2008 (calibration) and 2007 (validation). These values are obtained with $N = 10$ and $P = 15$.

data	pixel/grid definition	slope	intercept [m]	r	E [m]	B [m]	n_s
2008	$R_c = 10$ km	0.62	412	0.66	404	-75	5120
2007	$R_c = 10$ km	0.61	419	0.66	385	-59	6801
2007	$0.25° \times 0.25°$	0.58	455	0.64	398	-60	7970
2007	$0.75° \times 0.75°$	0.49	579	0.55	446	-56	10474

3.4 Scene limitations

This section investigates the applicability of MIBase by quantifying the amount of cases for which the concurrent conditions allow the successful derivation of a z_{base} retrieval. First, we filter for cases which fulfill the following two conditions: i) The number of valid z retrievals within the MIBase cell N_{val} must be > 0 and ii) METAR data must be available for the calibration and validation. These requirements are fulfilled for about two thirds of a all considered MISR overpasses over the ceilometer sites (Table 5). Furthermore, there are two main conditions which prevent the derivation of a z_{base} retrieval. These are namely apparent clear sky conditions and apparent overcast which is only a limitation for MIBase. Here, we use the phrases

Table 5. Number of cases for different conditions of the cloud field observed by MISR and reported in METAR messages for the considered METAR sites. The number of z retrievals labeled "high confidence cloud" (N_{HCC}) or "high confidence surface" (N_{HCS}) according to MISR's stereo-derived cloud mask is used to characterize the cloud field. The size of the scene is defined by $R_c = 10$ km. * indicates apparent conditions. See text for details, including the meaning of boldface font.

description of the situation		2008	2008	2007	2007
MISR	METAR		[%]		[%]
overpasses over METAR sites		80454	154.1	89782	145.9
valid z retrievals	message available	**52215**	**100.0**	**61531**	**100.0**
$N_{HCC} = 0$; $N_{HCS} > 0$ (clear sky*)		19507	37.4	20300	33.0
	clear sky*	26983	51.7	30037	48.8
clear sky*	clear sky*	**16982**	**32.5**	**17374**	**28.2**
$N_{HCC} = 0$; $N_{HCS} > 0$ (clear sky*)	$\hat{z}_{base}$ retrieval	**2525**	**4.8**	**2926**	**4.8**
$N_{HCC} = 0$; $N_{HCS} > 0$ (clear sky*)	$\hat{z}_{base} > h_{min}$	2106	4.0	2520	4.1
$N_{HCC} > 0$; $N_{HCS} > 0$	clear sky*	**6800**	**13.0**	**8511**	**13.8**
$N_{HCC} > 0$; $N_{HCS} = 0$ (overcast*)		**15945**	**30.5**	**19725**	**32.1**
$N_{HCC} > 0$; $N_{HCS} = 0$ (overcast*)	$\hat{z}_{base}$ retrieval	12769	24.5	15600	25.4
$N_{HCC} > 0$; $N_{HCS} = 0$ (overcast*)	clear sky*	3176	6.1	4125	6.7
$N_{HCC} = 0$; $N_{HCS} = 0$		**51**	**0.1**	**51**	**0.1**
$N_{HCC} > 0$; $N_{HCS} > 0$	$\hat{z}_{base}$ retrieval	**9912**	**19.0**	**12944**	**21.0**
$N_{HCC} \geq N = 10$; $N_{HCS} > 0$	$\hat{z}_{base}$ retrieval	8603	16.5	11387	18.5
z_{base} retrieval	$\hat{z}_{base}$ retrieval	8535	16.3	11319	18.4
z_{base} retrieval; single layer	$\hat{z}_{base}$ retrieval	7863	15.1	10251	16.7
$z_{base} < h_{max} = 3$ km; single layer	$\hat{z}_{base}$ retrieval	7206	13.8	9407	15.3
$z_{base} < h_{max}$; single layer	$\hat{z}_{base} < h_{max}$	7043	13.5	9227	15.0
$z_{base} < h_{max}$; single layer	$h_{min} < \hat{z}_{base} < h_{max}$	**5120**	**9.8**	**6801**	**11.1**

"apparent clear sky" and "apparent overcast" rather than "clear sky" and "overcast", respectively, to account for the fact that this attribution is based on instrumental indications rather than known actual sky condition.

For METAR, apparent clear sky is indicated if a METAR message is available, but does not provide a valid retrieval. Note that in case the lowest cloud is above the METAR reporting range (typically 3700 m), it is possible that no retrieval is issued. Here, such cases would also be attributed apparent clear sky.

For MIBase, we attribute apparent clear sky to the following configuration of the SDCM: MISR sees the surface with high confidence ($N_{HCS} > 0$), and has no high confidence cloud in the view ($N_{HCC} = 0$). This does not have to be an actual clear sky case since it could include low confidence surface or low confidence cloud retrievals for which the declaration is less certain. In case of invalid $z$retrievals, it is also uncertain whether clouds are present or not.

Out of all MISR apparent clear sky cases, 87 % are also classified as clear sky by METAR while the remaining 13 % yield a METAR cloud height retrieval. Mismatches in attributing apparent clear sky cases are due to METAR retrievals below the threshold height $h_{\min}$(17 %) and other reasons, such as the temporal offset between MISR and METAR measurement. The METAR reports comprise retrievals over a 30 minute period. During this time, cloud formation and cloud dissipation can alter the cloud scene and cause mismatches between MISR and METAR retrievals.

Furthermore, for MIBase, we attribute apparent overcast to the following configuration of the SDCM: MISR observes a cloud with high confidence ($N_{\mathrm{HCC}} > 0$) and does not observe any surface retrievals with high confidence ($N_{\mathrm{HCS}} = 0$). Again, the scene could include invalid retrievals, or retrievals of low confidence. In about 20 % of all the MISR apparent overcast cases, the corresponding METAR report yields an apparent clear sky case. These could be cases where the cloud cover is mainly above the reporting range of the ceilometer.

Out of all cases with valid z retrievals within the MIBase cell ($N_{\mathrm{val}} > 0$) and a corresponding METAR retrieval, 19 % are processed further. The main reasons why cases are excluded are apparent clear sky scenes for MISR (37.4 %), apparent overcast for MISR (30.5 %) and apparent clear sky for METAR when valid z retrievals are within the MIBase cell (13 %). Additional requirements, such as the minimum number of z retrievals marked high confidence cloud ($N_{\mathrm{HCC}} > N$), single layer situations, z_{base} and $\hat{z}_{\mathrm{base}}$ retrievals below $h_{\max}$ and METAR retrievals above the MISR threshold height ($\hat{z}_{\mathrm{base}} > h_{\min}$), lead to a further reduction of the number of cases which are used to derive the statistics. Further numbers for specific cases are presented in Table 5.

4 MIBase Evaluation

With the parameters $R_{\mathrm{c}} = 10\,\mathrm{km}$, $N = 10$ and ~~P~~ $P = 15$ derived in the previous section, MIBase is applied to MISR retrievals which are coincident with ceilometer retrievals from the year 2007. These data have not been used for calibration. The joint density of z_{base} retrieved from MISR and ceilometer is shown in Fig. 7. For lower z_{base}, MISR yields higher heights than the ceilometers. This can possibly be attributed to the threshold height (Equation 1) constraining z_{base} retrievals at the lower end of the height distribution. For z_{base} greater than 1000 m, mean and median MISR heights are lower than the ceilometer. Overall, the bias B is slightly negative (about 60 m; cf. Tab. 4) and the density of the retrieval differences is shifted slightly towards negative values (Fig. 7, right). Thus, MISR z_{base} retrievals are generally lower than the ceilometer retrievals. This could be due to the different ~~fields of view~~sample volumes. On the one hand, the ceilometer only records point measurements over a period of time, so that the measured sample of the cloud depends on the velocity of the wind. On the other hand, MISR observes the entire circular area defined by R_{c} around the ceilometer location. Chances are that MISR can observe a cloud with a lower base which does not pass over the ceilometer.

The joint density and the density of the ~~retrievals~~ retrieval differences appear similar for both the 2007 and the 2008 data sets (Fig. 7). Slope, intercept, r^2, E, and B resulting from the z_{base} retrieval comparisons for the year 2008 (calibration) and the year 2007 (validation) appear very similar~~proving~~, demonstrating the stability of the algorithm with the chosen parameters (Tab.~~4)~~ 4) to interannual variability in cloud properties. Changing the ~~field of view~~ MIBase cell to a $0.25° \times 0.25°$ ~~latitude~~

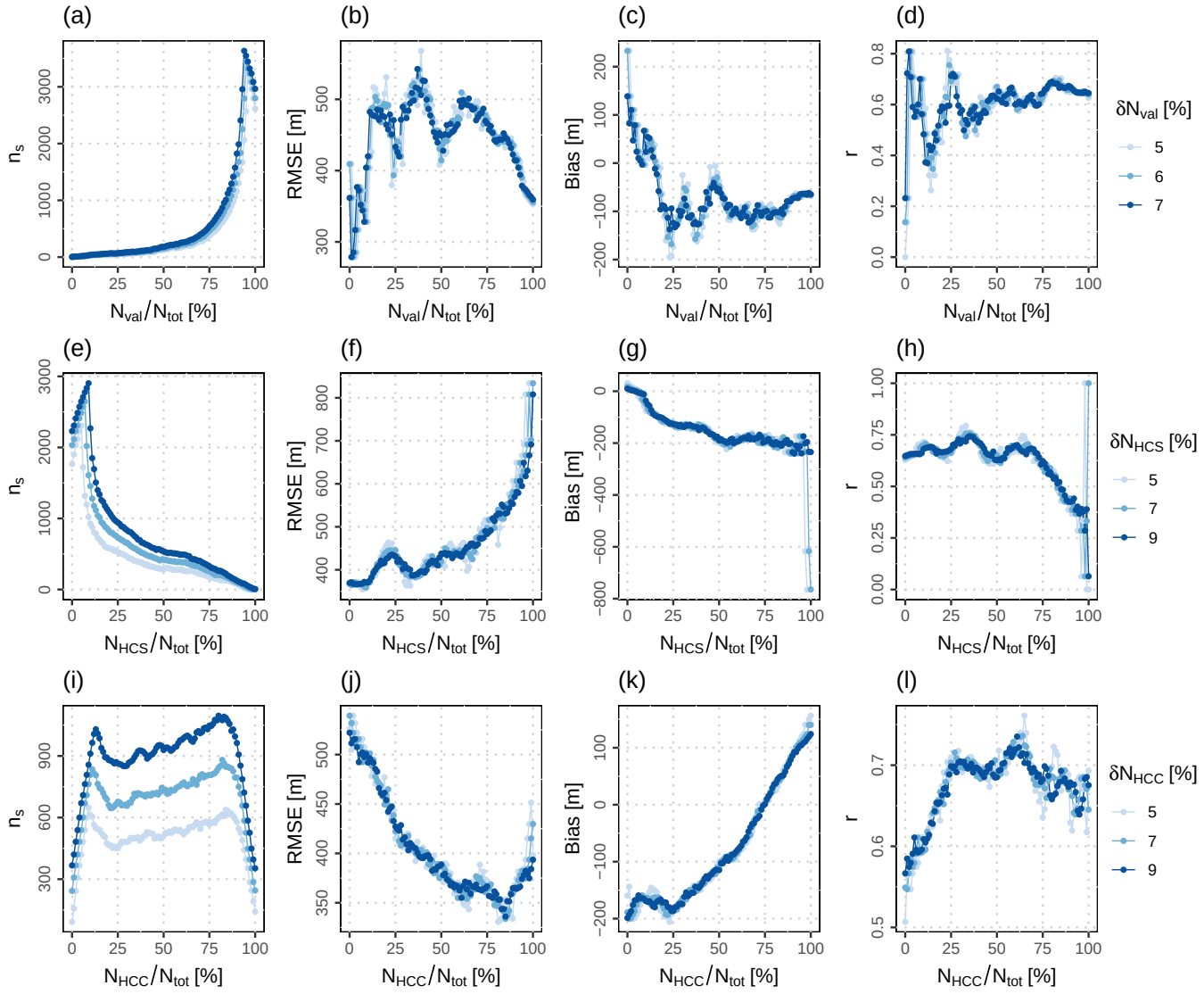

Figure 8. From left to right: number of samples n_s, RMSE, bias and correlation coefficient r for the comparison of MIBase and ceilometer retrievals as a function of the number of valid z retrievals N_{val} (top row), the number of retrievals marked high confidence surface N_{HCS} (middle row) and the number of retrievals marked high confidence cloud N_{HCC} (bottom row). Each data point is calculated for a sub sample which includes only $N_{val} \pm \delta N_{val}$, $N_{HCS} \pm \delta N_{HCS}$ and $N_{HCC} \pm \delta N_{HCC}$, respectively. The various widths of the considered N_{val} and N_{HCC} windows are indicated by the blue shading. All values are normalized by the total number of pixels within the MIBase cell N_{tot}. Data are for the year 2008 with $R_c = 10$ km, $P = 15$ and $N = 10$.

~~longitude~~ latitude–longitude grid results in a slightly lower correlation coefficient accompanied by a higher E. An even coarser grid size of $0.75° \times 0.75°$, which is applied later for a comparison with ERA-Interim cloud heights, results in an even lower correlation and higher E. A decreasing agreement between z_{base} and $\hat{z}_{\text{base}}$ for a ~~greater field of view~~ larger MIBase cell has already been ~~disclosed~~ described when studying the influence of R_c (see discussion in Section 3.3).

~~To address the question whether the algorithm performs differently for different cloud types, a simple approach is carried out. The retrievals are split into two groups by their cloud vertical extent Δz, i.e. above and below the median Δz. Here, we define Δz as the difference of z_{top} (95th percentile) and z_{base} (15th percentile). z_{base} retrievals for clouds with a lower Δz are expected to include more stratiform cloud types which are typically thin in respect to their vertical extent. z_{base} retrievals for clouds with higher Δz are expected to include more convective clouds as these show stronger vertical development. For stratiform cloud types, z is typically more uniform and such scenes are generally less broken, thus hiding the z_{base}. Therefore, these types might pose a more challenging scene to retrieve z_{base}. However, for thin clouds both forward and aft viewing cameras from MISR have a better chance to observe a cloud feature, and thereby the height retrieval is improved. For thicker clouds, one camera pair might not be able to see a lower cloud feature hidden by the cloud itself leading to a low confidence retrieval which is not considered by the algorithm.~~

~~The median Δz of all cloud scenes for which a MISR and a ceilometer retrieval could be matched is about 727 m. The joint density of MISR and ceilometer z_{base} retrievals for thinner clouds is narrower with the majority of the z_{base} retrievals ranging between 500 m and 1000 m (Fig. **??**, left). Clouds with a Δz greater than the median (Fig. **??**, right) have a greater range of z_{base} and a wider joint density. This behavior is also reflected in E, which is greater for clouds with a higher Δz (446 m compared to 312 m). The greater E for clouds with a greater Δz could be due to the termination of the z_{base} range by the threshold height. The threshold height constitutes a lower bound to z_{base} and thus limits E in particular for lower clouds, such as stratiform clouds with a low Δz.~~

4.1 Scene structure influence

To estimate the influence of the the scene structure on the performance of MIBase, we further exploit the MISR cloud top height product and the MISR Ancillary Geographic Product to investigate characteristics of the terrain height and the cloud field.

To derive a quantity to estimate the variability of the terrain height, we calculate the standard deviation of the average scene elevation, which is provided by the ancillary product at 1.1 km resolution. For each METAR site, the standard deviation is calculated for an area defined by different R_c (5 km, 10 km, 15 km, 20 km and 30 km). Typical standard deviations range around a few tens of meters with overall higher standard deviations for greater R_c (Fig. S1 a). When METAR sites with a higher standard deviation of the average scene elevation are excluded from the comparison of MIBase and METAR cloud base height retrievals, the RMSE decreases slightly, the bias slightly increases (towards 0), while the correlation is hardly affected (Fig. S1 b,c,d). Thus, the variability of the terrain height has a very small effect on the accuracy of the MIBase algorithm, with a slightly better performance over more homogeneous terrain.

To further investigate the performance of the MIBase algorithm as a function of parameters related to cloud types, we determine RMSE, bias, and the correlation coefficient as a function of z_{top} and the cloud vertical extent Δz (Fig. S2). The best correlation is obtained for cloud vertical extents up to 1000 m. The RMSE is also smaller for lower Δz and for lower z_{top}. However, the RMSE increases with decreasing z_{top} below about 1000 m. We conclude that MIBase performs best for shallow low clouds. However, further analyses are necessary to increase the sample size of thicker clouds and to include more medium high and high clouds for a more robust analysis of such cloud types. Furthermore, the increased RMSE for very low z_{top} indicates that, for very shallow low clouds in the proximity of the threshold height, MIBase retrievals do not agree as well with the METAR retrievals. This might be due to cases for which MIBase detects a shallow low cloud with z_{base} and z_{top} close the h_{min} when, in fact, the actual cloud base is below h_{min}. MIBase would miss this actual cloud base height because the retrievals below h_{min} would not be marked high confidence cloud. For that matter, we require that the ceilometer retrieval is above the threshold height ($\hat{z}_{base} > h_{min}$). However, if such a near surface cloud was not detected by the ceilometer, a mismatch would result leading to a higher RMSE.

Additionally, we exploit the stereo-derived cloud mask as a proxy of cloud cover fraction to investigate the sensitivity of the MIBase performance to the number of valid z retrievals N_{val}, the number of z retrievals marked high confidence surface N_{HCS}, and the number of z retrievals marked high confidence cloud N_{HCC} within the MIBase cell. We determine RMSE, bias, and the correlation coefficient as a function of N_{val}, N_{HCS} and N_{HCC} normalized by the total number of pixels N_{tot} which the MIBase cell encloses (Fig. 8). For example, for R_c = 10 km, a total of N_{tot} = 265 pixel is processed by MIBase to obtain a unique z_{base} retrieval. For the continental USA, most cases comprise a high portion of valid z retrievals within the MIBase cell. The RMSE, bias, and the correlation coefficient are robust under different choices of N_{val} and N_{HCS}. This suggests that MIBase generally does not depend much on cloud cover fraction. However, for cases which suggest almost apparent clear sky, indicated by high N_{HCS}, RMSE increases and r decreases. This could be due to a lower chance of observing the same cloud in case of less extended clouds. This bias appears to strongly depend on the portion of z retrievals marked high confidence cloud (Fig. 8). The increased bias for higher N_{HCC} could be explained by the decreasing portion of the thin edge of the cloud compared to the thicker part of the cloud with greater horizontal extent. For instance, the edge of a larger cloud might only be partly within the MIBase cell, whereas the edge of a smaller cloud might be fully processed by MIBase. The clear increase of the bias with increasing N_{HCC} shows potential for a bias correction in the future after a better understanding of the underlying reasons. The bias obtained in this study can have different sources: the different sample volumes of the defined MIBase cell and the ceilometer, biased MISR z retrievals, various scene characteristics.

5 MIBase Application

5.1 Global cloud height distribution

MIBase has been applied for a three year period between 2007 and 2009 to determine the z_{base} from MISR globally. Herein, z data from each individual orbit have been sorted into a $0.25° \times 0.25°$ longitude by latitude grid. For each orbit and each grid box z_{base} has been retrieved as described above and the median over the three year period has been calculated. Only cloud

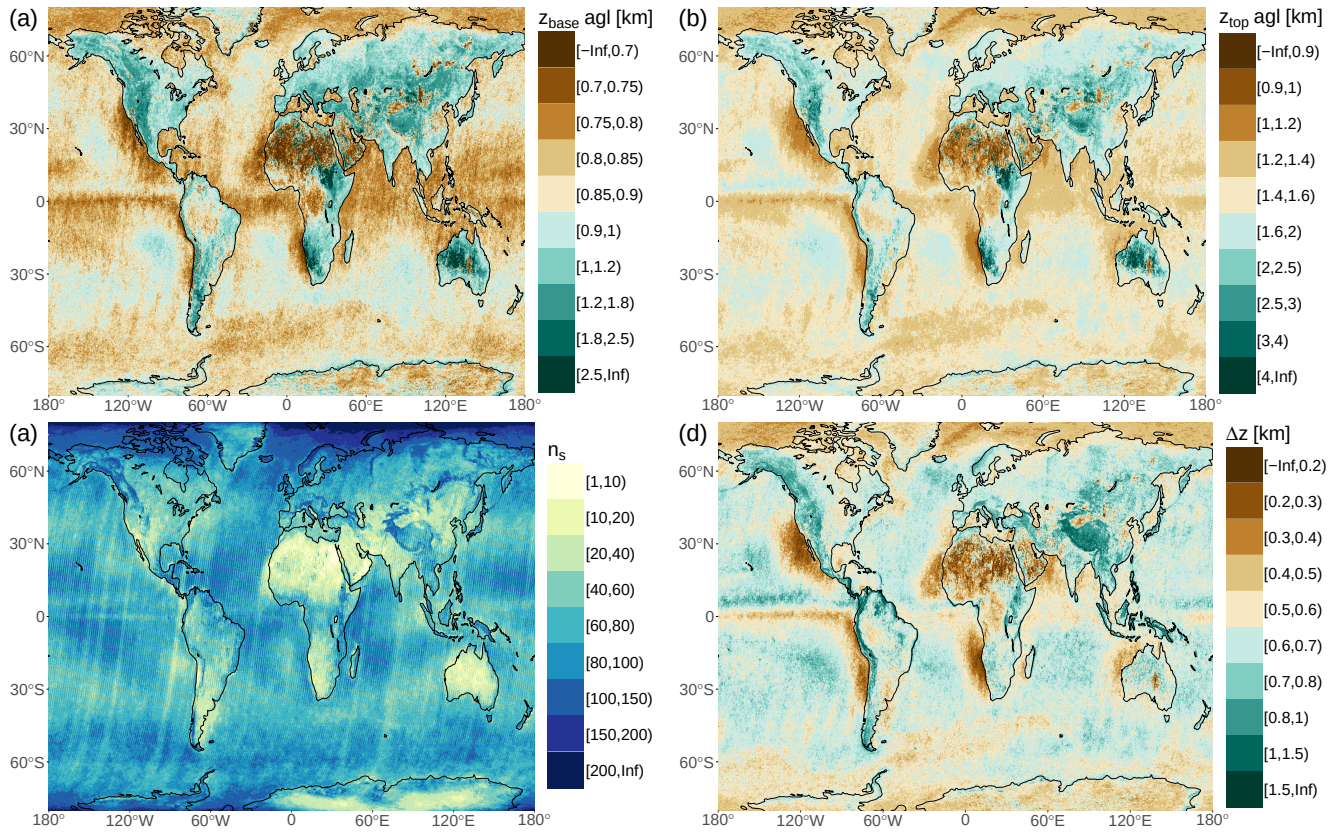

Figure 9. Global distribution of median cloud heights for a 3-year period (2007 ~~–~~2009). Shown are z_{base} (a), z_{top} (b), and cloud vertical extent (d) on a 0.25°×0.25° latitude ~~by~~ longitude grid. z_{base} and z_{top} are above ground level (agl). z_{base} and z_{top} retrievals are only included in the statistic if ~~they are~~ z_{base} is below 5000 m. The number of retrievals n_s (c) represents the number of valid z_{base} retrievals within this 3-year period.

height retrievals below 5000 m are considered to exclude cirrus clouds from the statistics. z_{top} is retrieved analogously to z_{base} by applying the 95th percentile on the z distribution. Taking the difference between z_{top} and z_{base} for each observed cloud scene yields Δz. The medians of these measures are shown in Fig. 9.

A sharp and steep gradient of the z_{base} can be seen at most coast lines with a higher z_{base} over land. This seems plausible as boundary layers above oceans are known to be shallower. Exceptions to this rule are the Congo Basin and the Amazon Basin. These regions are moisture sinks characterized by high precipitation and excessive surface run-off. The maritime stratus cloud regions are clearly visible at the subtropical eastern boundaries of the Pacific, Atlantic and Indian ocean. These regions are characterized by prevailing high pressure due to the location at the subsiding branch of the Hadley circulation and cold ocean currents creating a temperature inversion on top of the boundary layer. For these regions cloud formation is limited to the well mixed maritime boundary layer. The Intertropical Convergence Zone (ITCZ) is clearly visible in particular for the tropical Pacific ocean with a higher z_{base} and even higher z_{top} yielding an overall higher Δz slightly north of the equator. Over land, this

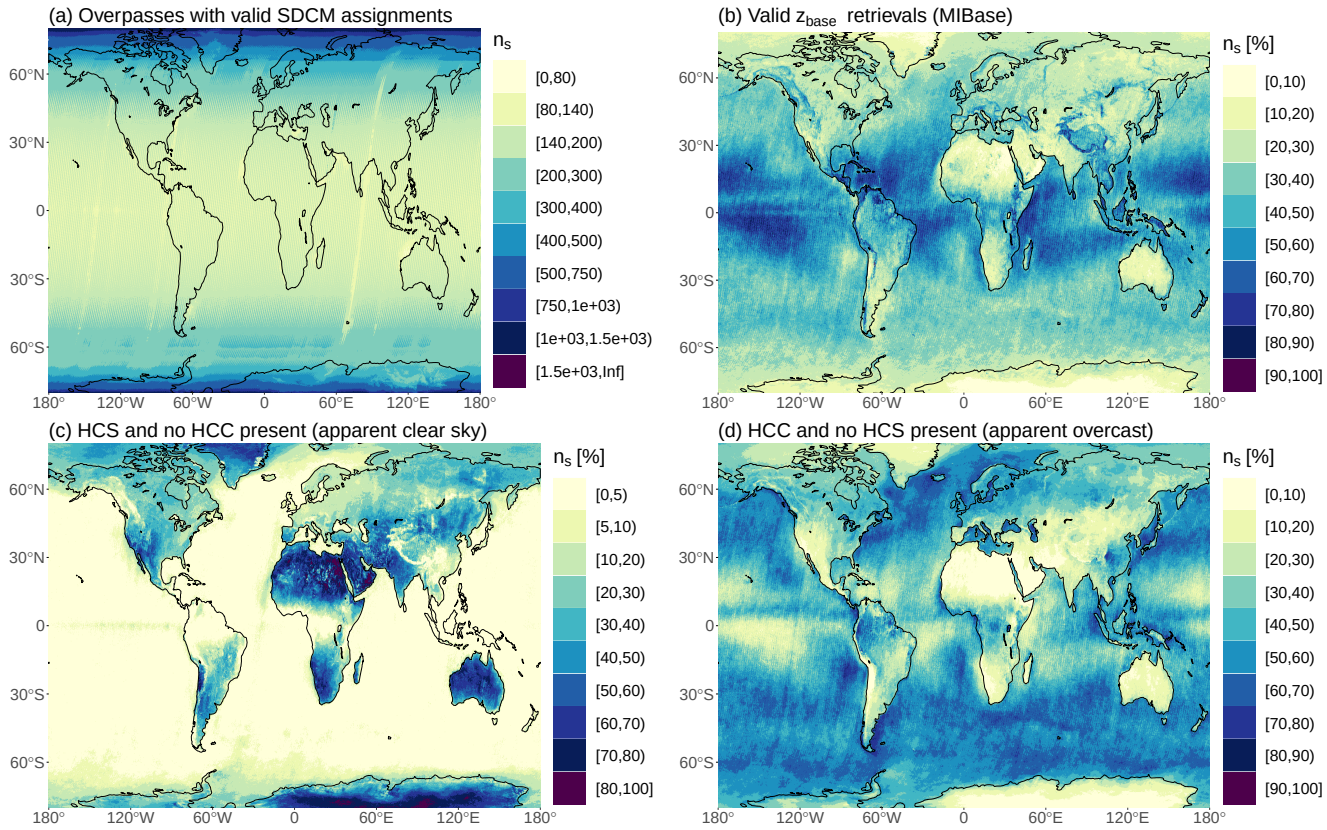

Figure 10. Relative occurrences of different stereo-derived cloud mask (SDCM) configurations within the three-year period (2007–2009). The reference sample size n_s given in (a) corresponds to 100 % and includes all overpasses per grid cell which contain valid z retrievals. (b) Relative number for which MIBase successfully retrieved z_{base}. (c) through (d) show the relative number of occurrence of cloud scenes which include z retrievals of specific SDCM labels within a grid cell. These configurations are: (c) No high confidence cloud (HCS). These cases are apparent clear sky cases. (d) No high confidence cloud (HCS). These cases are apparent overcast cases.

phenomenon is not as clear. There, the diurnal cycle of surface heating becomes important. MISR on the Terra satellite has a morning overpass over the equator when cloud formation just begins. Taylor et al. (2017) show the diurnal cycle of cloud top temperature (CTT) derived from SEVIRI measurements indicating that the lowest z_{top} occurs between 9~~a.m. and 1 p.m.~~ :00 and 13:00 local time with the lowest mean CTT at 11~~a.m.~~ :00. and the lowest median CTT at ~~noon~~12:00, close to the overpass time of MISR.

The sampling size varies spatially with a higher number of retrievals in the Arctic region. (Fig.~~9 (b))shows higher number of retrievals in polar regions~~ 9 (c)). This is expected for a polar orbiting satellite with more frequent MISR overpasses in ~~this region.The highest sampling sizes are obtained in~~ polar regions (Fig. 10 (a)). Generally, the causes for retrieval failure are apparent clear sky and apparent overcast situations as discussed in Section 3.4. The frequency of occurrence of such situations varies spatially. For continental dry regions in the subtropics and continental polar regions apparent clear sky conditions

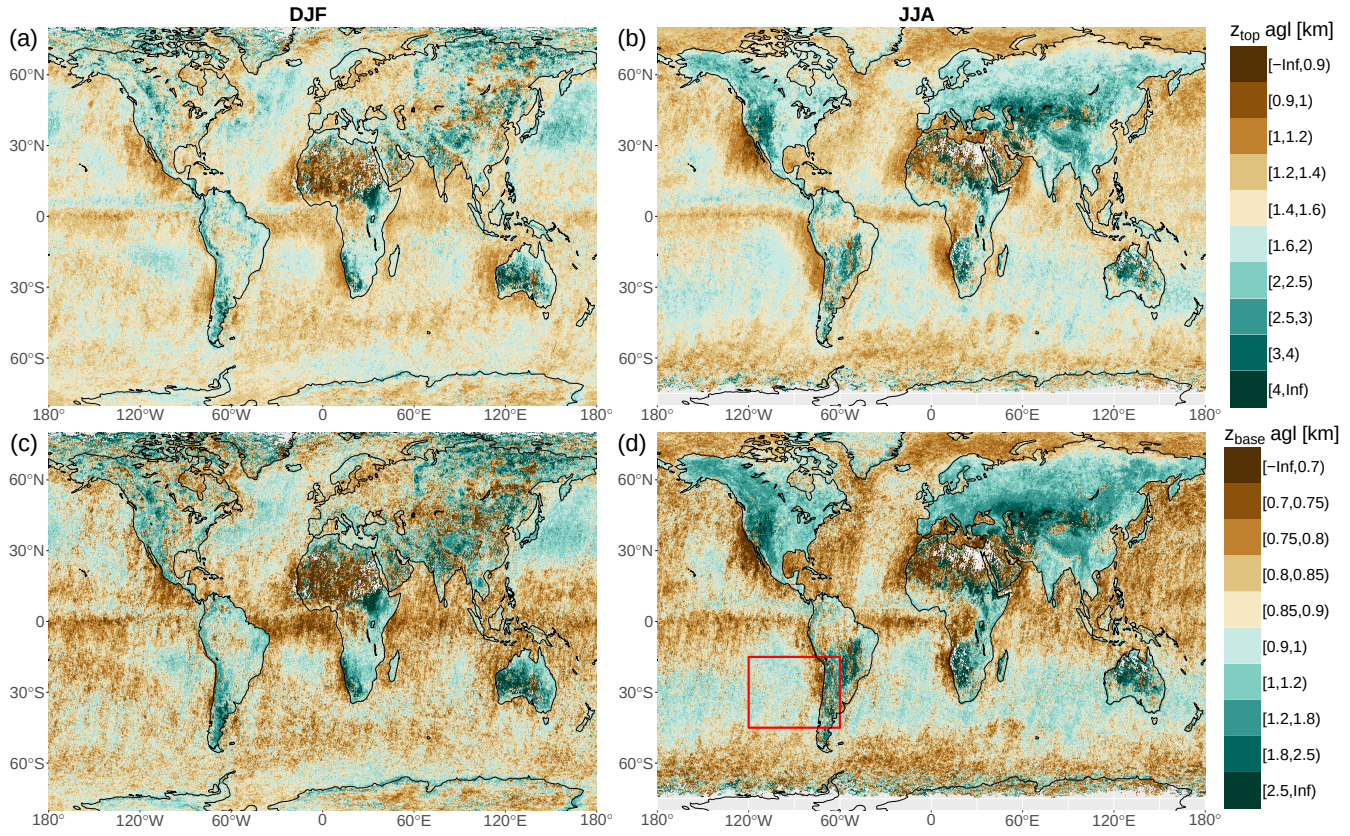

Figure 11. Global distribution of seasonal median cloud heights for a 3-year period (2007 ~~–~~2009). Shown are z_{top} (a, b), and z_{base} (c, d) for December, January, February (a, c) and June, July, August (b, d) on a $0.25° \times 0.25°$ latitude ~~by~~ longitude grid. z_{base} and z_{top} are above ground level (agl). z_{base} and z_{top} retrievals are only included in the statistic if ~~they are~~ z_{base} is below 5000 m. The red rectangle in (d) frames the region for which results over a 16-year period are presented in Fig. 12.

predominantly limit the number of z_{base} retrievals (Fig. 10 (c)). The continental polar regions yield a high number of cases for which the grid cell comprises only high confidence surface retrievals (N_{HCS}=N_{tot}, Fig. S3). This poses an even more robust indication of apparent clear sky conditions. However, the boundary layer is typically shallower in polar regions. Therefore, boundary layer clouds occur likely below h_{min}, so that z_{base} cannot be retrieved by the MIBase algorithm. Predominant apparent overcast conditions limit the number of z_{base} retrievals for midlatitude regions over ocean and stratocumulus regions on the western boundaries of continents in the subtropics. In midlatitude continental regions, a mix of apparent clear sky and apparent overcast conditions limits the ~~Arctic.The smallest sampling sizes are obtained over subtropical continental areas, such as the Sahara desert, the Namib desert or Australia. These regions are characterized by a low frequency of cloud occurrence. Results should be interpreted with care as the~~ number of ~~samples is low.~~ z_{base} retrievals. In the trade cumulus regions within 30°N and

30°S, very high success rates occur (Fig. 10 (b)). A visual comparison to the 2011 mean cloud cover fraction derived from MODIS (Suen et al., 2014) indicates the plausibility of the attribution of apparent clear sky and apparent overcast.

To further investigate the plausibility of the seasonal variability of cloud heights, composites over the three year period are presented in Fig. ~~11. Distinguished are~~ 11. We distinguish boreal winter season ~~including~~ comprising December, January and February (DJF) and boreal summer season ~~including~~ comprising June, July and August (JJA). Over land and between 30°N and 70°N, z_{base} and z_{top} are lower during winter, when stratiform clouds ~~are prevailing~~prevail. In contrast, z_{base} and z_{top} are higher during summer, when more convective clouds are typically present. Boundary layer clouds are also lower during winter season since the boundary layer is ~~of lesser extent~~ shallower during the cold season. Over ocean an inverse pattern can be observed on both hemispheres. During winter z_{base} and z_{top} are higher than during the summer. Sea surface temperatures show less seasonal variation than ~~higher tropospheric air~~ air temperatures due to the higher heat capacity of the water. This causes additional instability during winter enhancing convective cloud formation which can result in higher cloud heights. Additionally, the instability during winter can be attributed to storm tracks. During summer, the influence of high pressure systems can limit convection to the maritime boundary layer causing cloud heights to be lower.

5.2 Southeast Pacific

The southeast Pacific hosts one of the largest and most persistent stratocumulus cloud decks on Earth as shown by Wood (2012) using data from the combined land-ocean cloud atlas database (Hahn and Warren, 2007). In this region, cloud cover and cloud thickness have major impacts on the net cloud radiative effect, which raises the importance of studying the heights of these clouds.

Orographically induced fog at the coastal cliff ranging from Peru to northern Chile is the major source of moisture for this region (Pinto et al., 2006). z_{base} and z_{top} of the stratocumulus clouds near the coast determine the areas where fog can provide water to the environment at the coastal cliff. The cloud heights also ~~effect~~ affect the ability of the fog to be advected further inland across the cliff. Here, we apply the z_{base} retrieval algorithm to determine the spatial and seasonal variability of z_{base} and z_{top} for the region (see red rectangle in Fig. 9 (bottom right)). We extend the time window to the full 16-year record of available MISR data (~~2001-2016~~2001–2016). Furthermore, we investigate how well the temporal changes are represented in the global reanalysis ERA-Interim.

5.2.1 Spatial and seasonal variability of z_{base} and z_{top}

For the 16-year period, the medians of z_{base} and z_{top} over the southeast Pacific are shown in Fig. 12. Distinguished are summer and winter season. Over ocean the median z_{base} ranges from 600 m near the the coast to about 1200 m further west. During austral summer (DJF) the lowest z_{base} is observed near the coast between 30°S and 35°S. During austral winter the region of low z_{base} shifts to the north between 20°S and 30°S. This shift is ~~coherent~~ in phase with the direction of the seasonal shift of the Hadley cell. It appears that the region of lowest z_{base} corresponds to the strongest subsidence. During austral summer the highest z_{base} clearly appear in the north, whereas during austral winter a ~~north-south~~ north–south gradient is hardly visible between 120°W and 80°W. Over land, z_{base} is generally higher except for the coastal line north of 35°S, where cloud heights

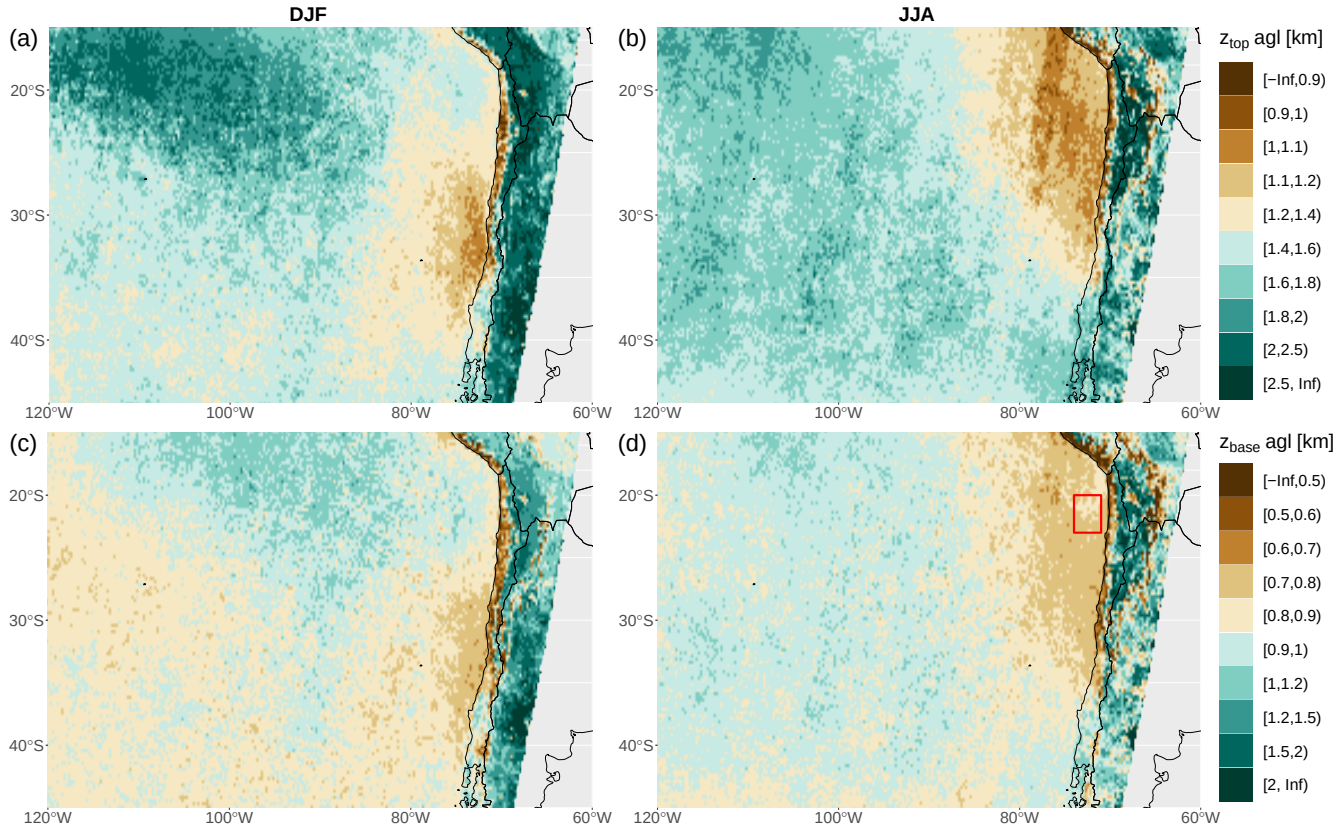

Figure 12. Median of z_{top} (~~bottom row~~a, b), and z_{base} (~~top row~~c, d) over a 16-year period (~~2001–2016~~2001–2016) for austral summer (DJF, ~~left column~~(a) and (c)) and austral winter (JJA, ~~right column~~(b) and (d)) on a $0.25° \times 0.25°$ longitude by latitude grid at the southeast Pacific. z_{base} and z_{top} are given above ground level (agl). The red rectangle (~~bottom right~~d) frames the region for which a time series of cloud heights is presented in Fig. 13.

are even lower than over ocean. There, the prevailing maritime stratocumulus clouds form orographic fog as they reach the coastal cliff. Similar spatial and seasonal patterns are apparent for z_{top}. Over ocean, the highest z_{top} is about 2500 m, which is observed during austral summer in the northwest of the region. The lowest z_{top} is about 1000 m, which is observed during winter and closer to the coast of northern Chile.

5.2.2 Cloud height comparison between MISR and ERA-Interim

In order to preliminarily assess how well clouds are represented in common reanalysis, we compare MISR derived z_{base} and z_{top} to cloud heights derived from ERA-Interim (Dee et al., 2011) which is provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). Cloud heights are not a direct output variable of ERA-Interim. Therefore, the cloud liquid water content is used to infer the cloud base height $\tilde{z}_{base}$ and cloud top height $\tilde{z}_{top}$. For each grid point, the vertical column is scanned

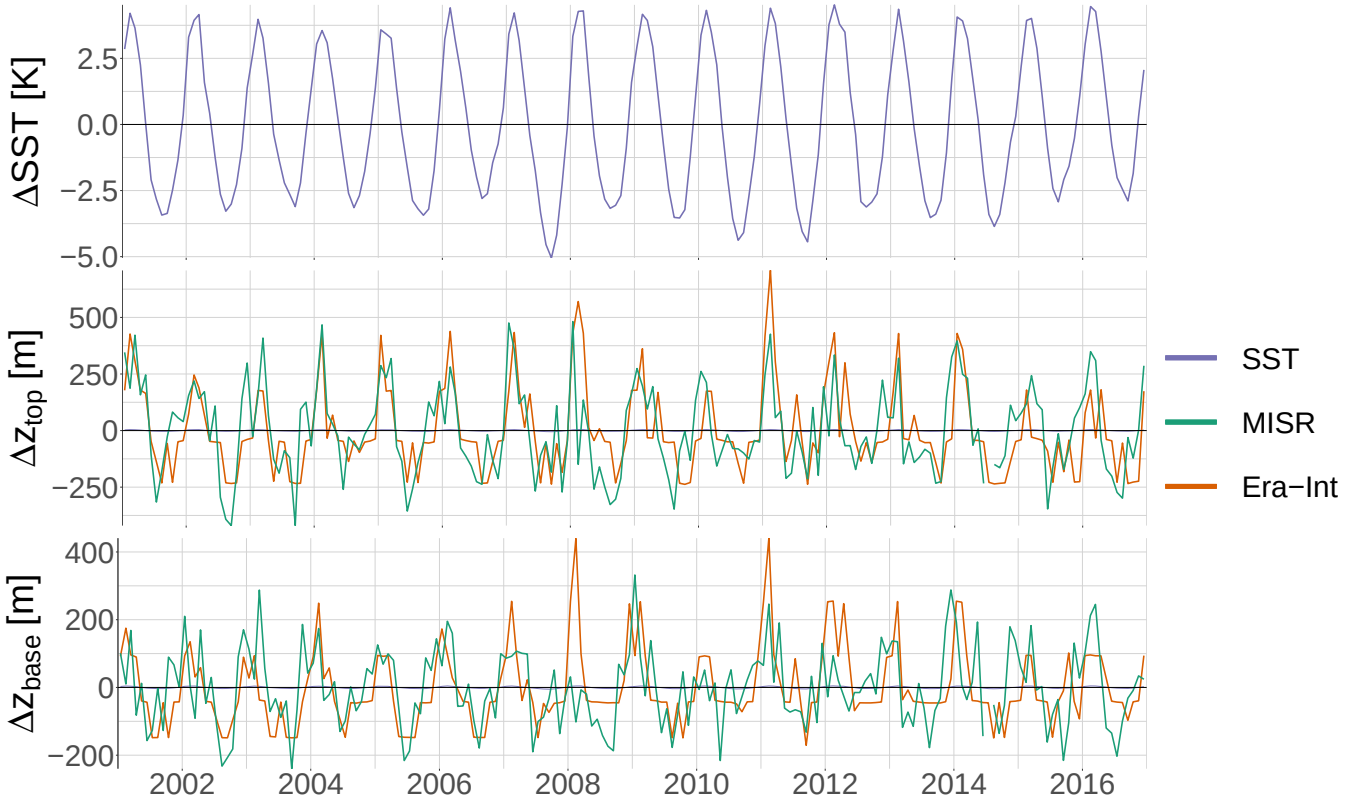

Figure 13. Time series of deviations of sea surface temperature ~~SST~~ΔSST (top), ~~z_{top}~~ cloud top height Δz_{top} (middle), ~~z_{base}~~ cloud base height Δz_{base} (bottom) from the corresponding mean over the entire period from 2001 through 2016. Cloud heights are derived from MISR (green) and ERA-Interim (orange). SST is derived from ERA-Interim. ~~Shown are the deviations from the mean over the entire period from 2001 through 2016.~~

for model levels with a specific cloud liquid water content greater than ~~$10^{-18}\,\frac{\mathrm{kg}}{\mathrm{kg}}$~~ $10^{-18}\,\mathrm{kg\,kg^{-1}}$ (≈ 0). The bottom height of the lowest of such levels is taken as $\tilde{z}_{\mathrm{base}}$. Moving higher in the column, $\tilde{z}_{\mathrm{top}}$ is given by the bottom height of the next higher model level which has a cloud liquid water content equal to zero. We use data with a $0.75° \times 0.75°$ resolution, which is similar to the native grid of ERA-Interim, over a region between 20°S and 23°S and 74°W and 71°W as indicated by the red rectangle in Fig. 12. ERA-Interim data is provided 6-hourly. The comparison is performed using the 18:00 UTC output which corresponds to 14:00 Chile Standard Time (CLT). Note, MISR overpass times range around 10:51 CLT to 11:29 CLT for this particular region.

For each MISR overpass and ERA-Interim 18:00 UTC output, the median cloud heights are used to calculate the median cloud heights of each month over the whole 16-year period. The mean difference of the monthly cloud heights is roughly 500 m for both cloud base height and cloud top height, with ERA-Interim yielding lower cloud heights than MISR. That $\tilde{z}_{\mathrm{base}}$ is lower than z_{base} could be due to the threshold height used to determine the MISR stereo derived cloud mask (Equation 1) which leads

to a cut-off of z_{base} retrievals at h_{min}. At the same time the same bias is found between z_{top} and $\tilde{z}_{\text{top}}$. This could be an indicator that clouds are systematically placed too low by ERA-Interim. Hannay et al. (2009) ~~mention~~ mentioned several studies which conclude that models typically underestimate the height of the planetary boundary layer (PBL) in the southeast Pacific area. This would cause boundary layer clouds to appear lower than observed. Their study compares the PBL height retrieved from in-situ measurements and remote sensing to different models. While the observations show a PBL height of 1100 m, the models produce a PBL height between 400 m and 800 m, hence an underestimation of 700 m to 300 m. This is in accordance with the bias found here.

To reveal the annual cycle of the cloud heights, we look at anomalies from the 16-year mean of each time series (Fig. 13). These anomalies of z_{base} and $\tilde{z}_{\text{base}}$ as well as z_{top} and $\tilde{z}_{\text{top}}$ from their respective mean values agree rather well, thus the amplitude of the annual cycle appears very similar. Figure 13 also shows the anomaly of the ~~Sea Surface Temperature~~ sea surface temperature (SST) from its 16-year mean value. SSTs are taken from ERA-Interim as well. The peaks of the cloud heights correspond to the maxima of the SSTs. While the highest SSTs coincide with the highest cloud heights during austral summer, the lowest SSTs coincide with the lowest cloud heights during austral winter.

6 Conclusions

Here, we present a new method to determine z_{base} over a spatial region from satellite based measurements. The MIBase algorithm derives z_{base} from the high spatial resolution MISR cloud top height product $z$if some preconditions, such as a broken cloud scene, are met. Validation against 1510 ceilometer stations in the continental USA results in a correlation coefficient of 0.66 and a RMSE of 385 m for the validation data set (year 2007). The bias of −59 m even states that MISR sees a slightly lower z_{base} on average. This is possibly due to the larger ~~field of view~~ retrieval cell which is set up for the retrievals from MISR as opposed to the point measurements provided by the ceilometer.

Very few attempts to derive z_{base} from satellite have been performed and evaluated before. Desmons et al. (2013) retrieve Δz from POLDER measurements. The standard deviation of the difference between their Δz retrieval and reference data from CPR and CALIOP is about 964 m. However, their method is hard to compare to the MIBase algorithm, since they retrieve Δz and make a distinction of different types of clouds which is not done in this study. The CBASE algorithm (Mülmenstädt et al., 2018) derives z_{base} from CALIOP measurements even for optically thick clouds. Depending on the circumstances different retrieval uncertainties can be derived. Similar to the study presented here, they compare their z_{base} retrievals with ceilometer data over the continental ~~U. S.A.~~ USA. They obtain RMSEs between 404 m and 720 m depending on the concurrent local conditions of the individual retrievals. The RMSE we obtain for the MIBase algorithm is slightly lower. Even though the two studies make use of a similar reference data base, they measure cloud heights at different times of the day. While CALIOP has an afternoon overpass, MISR has a morning overpass, when more clouds of lesser extent are present. For a more in-depth comparison and validation of the presented algorithm, more cloud height reference observations would be desirable including observations in different climate zones and especially over ocean.

Within Europe, the European Cooperation in Science and Technology (COST) activity is expected to harmonise the networks of the different weather services (e.g., Haeffelin et al., 2016; Illingworth et al., 2018), enabling more intercomparisons in the future.

An important strength of MIBase is the geometric approach which is applied to create the z product from MISR measurements. Neither a calibration nor auxiliary data ~~is~~ are necessary to obtain the z product which is the starting point for the z_{base} retrieval algorithm presented here. In consequence, retrievals are possible over all kinds of terrain even above ice. A disadvantage is the threshold height which MISR requires to create the stereo derived cloud mask. Therefore, depending on the terrain variability in the vicinity of the measurement, this new z_{base} retrieval method is not capable ~~to derive~~ of deriving z_{base} below at least 560 m (flat terrain). The algorithm requires a broken cloud scene. For complete overcast within the chosen MIBase cell, z_{base} cannot be retrieved. Therefore, climatologies derived from this algorithm would be biased towards cloud types for which MISR is able to observe the surface through cloud gaps.

Depending on the application, the MIBase uncertainty and the missing coverage of the diurnal cycle can be a limitation. However, in combination with ceilometer networks, both temporal and spatial patterns can be investigated. The application of MIBase over a three-year period ~~revealed~~ reveals plausible patterns in the global distribution and seasonal variability of z_{base}. A first analysis over the 16-year MISR time series in the southeast Pacific shows the potential to investigate the interannual variability of z_{base}. This makes MIBase a promising tool for the evaluation of climate models on seasonal and interannual time scales in data sparse regions if for example the climate model output is limited to clouds below 5 km and cloud fractions below 1 and if a sufficient amount of MIBase retrievals is provided within the considered region and time period.

Data availability. Multiple archives providing METAR data are available. The data utilized here were downloaded from the Weather Underground archive (https://www.wunderground.com/history/airport/). The MISR Level 2TC Cloud Product data were downloaded from the NASA Langley Research Center Atmospheric Science Data Center (ftp://l5ftl01.larc.nasa.gov/MISR/MIL2TCSP.001/). ERA-Interim data were downloaded from the ECMWF data server via Web-API. The MIBase cloud base dataset (Böhm, 2019) is freely available at the Collaborative Research Centre 1211 database under the DOI https://doi.org/10.5880/CRC1211DB.19. It comprises z_{base} retrievals globally on a $0.25° \times 0.25°$ grid for a three year period (2007–2009). Daily files include z_{base} retrievals derived from the MISR MIL2TCSP product for about 14 respective Terra revolutions around the Earth. Cloud base altitudes are given above the WGS 1984 ellipsoid. Furthermore, the surface altitude is provided to derive the cloud base height above ground level.

Appendix A: Sensitivity to threshold height h_{min}

The distinction between surface and cloud retrieval according to the threshold height described by Equation 1 introduces a constraint to the z_{base} retrieval algorithm. Below a height of 560 m for flat terrain, or higher for more complex terrain, z_{base} retrievals are not possible. As an attempt to lower this threshold height, we adjusted H_{SDCM} in Equation 1, so that:

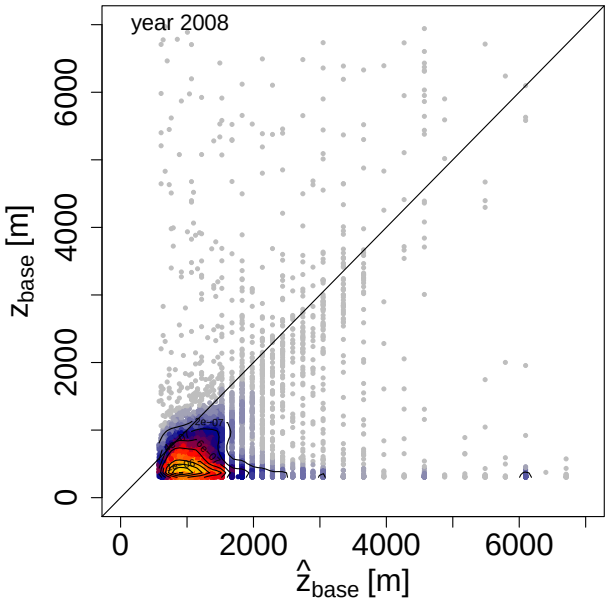

Figure A1. Joint density of z_{base} and $\hat{z}_{\text{base}}$ for the year 2008 applying a lower threshold height ~~$h_{\text{]}} = 300\,\text{m} + H + 2\sigma_{\text{h}}$~~ $h_{\text{min}} = 300\,\text{m} + H + 2\sigma_{\text{h}}$ (Equation A1) for the distinction between surface and cloud pixels in contrast to Equation 1.

$$h_{\text{min}} = 300\,\text{m} + H + 2\sigma_{\text{h}} \tag{A1}$$

This modification results in a bimodal retrieval density clearly showing a mode consisting of surface retrievals (Fig. A1). Therefore, the original threshold height given by MISR has to be applied, in order to ensure that only cloud retrievals are utilized during data processing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. ~~These~~ The MISR Level 2TC Cloud Product data were obtained from the NASA Langley Research Center Atmospheric Science Data Center (Mueller et al., 2013). We gratefully acknowledge financial support by the ~~Collaborative Research Centre "Earth - Evolution at the dry limit" subproject A01 funded by the~~ Deutsche Forschungsgemeinschaft (DFG~~)~~, German Research Foundation) – Projektnummer 268236062 – SFB 1211. We thank the six reviewers for their constructive feedback.

References

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Clouds and Aerosols, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and (eds.), P. M., chap. 7, pp. 571–657, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Bull, M., Matthews, J., McDonald, D., Menzies, A., Moroney, C., Mueller, K., Paradise, S., and Smyth, M.: Data Products Specifications, Tech. Rep. JPL D-13963, Revision S, Jet Propulsion Laboratory, California Institute of Technology, 2011.

Böhm, C.: MIBase cloud base height derived from satellite data, doi:10.5880/CRC1211DB.19, https://doi.org/10.5880/CRC1211DB.19, [Accessed 21. February 2019], 2019.

Costa-Surós, M., Calbó, J., González, J. A., and Long, C. N.: Comparing the cloud vertical structure derived from several methods based on radiosonde profiles and ground-based remote sensing measurements, Atmospheric Measurement Techniques, 7, 2757–2773, doi:10.5194/amt-7-2757-2014, https://www.atmos-meas-tech.net/7/2757/2014/, 2014.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828, 2011.

Desmons, M., Ferlay, N., Parol, F., Mcharek, L., and Vanbauce, C.: Improved information about the vertical location and extent of monolayer clouds from POLDER3 measurements in the oxygen A-band, Atmospheric Measurement Techniques, 6, 2221–2238, doi:10.5194/amt-6-2221-2013, http://www.atmos-meas-tech.net/6/2221/2013/, 2013.

Diner, D.: MISR Level 2 Cloud Heights and Winds HDF-EOS File - Version 1, NASA Langley Atmospheric Science Data Center DAAC, doi:10.5067/Terra/MISR/MIL2TCSP_L2.001, https://doi.org/10.5067/terra/misr/mil2tcsp_l2.001, 2012.

Ferlay, N., Thieuleux, F., Cornet, C., Davis, A. B., Dubuisson, P., Ducos, F., Parol, F., Riédi, J., and Vanbauce, C.: Toward New Inferences about Cloud Structures from Multidirectional Measurements in the Oxygen A Band: Middle-of-Cloud Pressure and Cloud Geometrical Thickness from POLDER-3/PARASOL, J. Appl. Meteorol. Clim., 49, 2492–2507, http://dx.doi.org/10.1175/2010JAMC2550.1, 2010.

Goren, T., Rosenfeld, D., Sourdeval, O., and Quaas, J.: Satellite observations of precipitating marine stratocumulus show greater cloud fraction for decoupled clouds in comparison to coupled clouds, Gephys. Res. Lett., 45, 5126–5134, doi:10.1029/2018GL078122, https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2018GL078122, 2018.

Haeffelin, M., Crewell, S., Illingworth, A. J., Pappalardo, G., Russchenberg, H., Chiriaco, M., Ebell, K., Hogan, R. J., and Madonna, F.: Parallel Developments and Formal Collaboration between European Atmospheric Profiling Observatories and the U.S. ARM Research Program, Meteorological Monographs, 57, 29.1–29.34, doi:10.1175/AMSMONOGRAPHS-D-15-0045.1, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0045.1, 2016.

Hahn, C. J. and Warren, S. G.: A gridded climatology of clouds over land (1971–96) and ocean (1954–97) from surface observations worldwide. Numeric Data Package NDP-026E ORNL/CDIAC-153, Tech. rep., CDIAC, Department of Energy, Oak Ridge, TN, doi:10.3334/CDIAC/cli.ndp026e, 2007.

Hannay, C., Williamson, D. L., Hack, J. J., Kiehl, J. T., Olson, J. G., Klein, S. A., Bretherton, C. S., and Köhler, M.: Evaluation of Forecasted Southeast Pacific Stratocumulus in the NCAR, GFDL, and ECMWF Models, Journal of Climate, 22, 2871–2889, doi:10.1175/2008JCLI2479.1, https://doi.org/10.1175/2008JCLI2479.1, 2009.

Illingworth, A. J., Cimini, D., Haefele, A., Haeffelin, M., Hervo, M., Kotthaus, S., Löhnert, U., Martinet, P., Mattis, I., O'Connor, E. J., and Potthast, R.: How can Existing Ground-Based Profiling Instruments Improve European Weather Forecasts?, doi:10.1175/BAMS-D-17-0231.1, https://doi.org/10.1175/BAMS-D-17-0231.1, in press, 2018.

Lau, M. W., Yung, Y. L., and Wu, D. L.: Determining Cloud Base and Thickness from Spaceborne Stereoscopic Imaging and Lidar Profiling Techniques, Accepted by Caltech Undergraduate Research Journal, Spring Issue, http://web.gps.caltech.edu/~mlau/writeups/Cloud_Base_from_Spaceborne_Sterescopic_Imaging.pdf, 2012.

Marchand, R. T., Ackerman, T. P., and Moroney, C.: An assessment of Multiangle Imaging Spectroradiometer (MISR) stereo-derived cloud top heights and cloud top winds using ground-based radar, lidar, and microwave radiometers, Journal of Geophysical Research: Atmospheres, 112, doi:10.1029/2006JD007091, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD007091, 2007.

Meerkötter, R. and Zinner, T.: Satellite remote sensing of cloud base height for convective cloud fields: A case study, Geophysical Research Letters, 34, n/a–n/a, doi:10.1029/2007GL030347, http://dx.doi.org/10.1029/2007GL030347, l17805, 2007.

Merk, D., Deneke, H., Pospichal, B., and Seifert, P.: Investigation of the adiabatic assumption for estimating cloud micro- and macrophysical properties from satellite and ground observations, Atmospheric Chemistry and Physics, 16, 933–952, doi:10.5194/acp-16-933-2016, http://www.atmos-chem-phys.net/16/933/2016/, 2016.

Moroney, C. and Mueller, K.: Data Product Specification for the MISR Level 2 Cloud Product, Tech. Rep. JPL D-72327, Jet Propulsion Laboratory, California Institute of Technology, 2012.

Moroney, C., Davies, R., and Muller, J. P.: Operational retrieval of cloud-top heights using MISR data, IEEE Transactions on Geoscience and Remote Sensing, 40, 1532–1540, doi:10.1109/TGRS.2002.801150, 2002.

Mueller, K., Moroney, C., Jovanovic, V., Garay, M., Muller, J.-P., Di Girolamo, L., and Davies, R.: MISR Level 2 Cloud Product Algorithm Theoretical Basis, Tech. Rep. JPL D-73327, Jet Propulsion Laboratory, California Institute of Technology, 2013.

Mülmenstädt, J., Sourdeval, O., Henderson, D. S., L'Ecuyer, T. S., Unglaub, C., Jungandreas, L., Böhm, C., Russell, L. M., and Quaas, J.: Using CALIOP to estimate cloud-field base height and its uncertainty: The Cloud Base Altitude Spatial Extrapolator (CBASE) algorithm and dataset, Earth Syst. Sci. Data, 10, 2279–2293, doi:10.5194/essd-10-2279-2018, 2018.

National Oceanic and Atmospheric Administration, Department of Defense, Federal Aviation Administration, and United States Navy: Automated Surface Observing System User's Guide, http://www.nws.noaa.gov/asos/pdfs/aum-toc.pdf, 1998.

Pinto, R., Barría, I., and Marquet, P.: Geographical distribution of Tillandsia lomas in the Atacama Desert, northern Chile, Journal of Arid Environments, 65, 2006.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., and Mitrescu, C. a.: THE CLOUDSAT MISSION AND THE A-TRAIN, Bulletin of the American Meteorological Society, 83, 1771–1790, doi:10.1175/BAMS-83-12-1771, https://doi.org/10.1175/BAMS-83-12-1771, 2002.

Suen, J. Y., Fang, M. T., and Lubin, P. M.: Global Distribution of Water Vapor and Cloud Cover—Sites for High-Performance THz Applications, IEEE Transactions on Terahertz Science and Technology, 4, 86–100, doi:10.1109/TTHZ.2013.2294018, 2014.

Taylor, S., Stier, P., White, B., Finkensieper, S., and Stengel, M.: Evaluating the diurnal cycle in cloud top temperature from SEVIRI, Atmospheric Chemistry and Physics, 17, 7035–7053, doi:10.5194/acp-17-7035-2017, https://www.atmos-chem-phys.net/17/7035/2017/, 2017.

Van Beusekom, A. E., González, G., and Scholl, M. A.: Analyzing cloud base at local and regional scales to understand tropical montane cloud forest vulnerability to climate change, Atmospheric Chemistry and Physics, 17, 7245–7259, doi:10.5194/acp-17-7245-2017, http://www.atmos-chem-phys.net/17/7245/2017/, 2017.

Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant, P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., Mccormick, M. P., Mégie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M. A., and Wielicki, B. A.: The CALIPSO Mission, Bulletin of the American Meteorological Society, 91, 1211–1230, doi:10.1175/2010BAMS3009.1, https://doi.org/10.1175/2010BAMS3009.1, 2010.

Wood, R.: Stratocumulus Clouds, Monthly Weather Review, 140, 2373–2423, doi:10.1175/MWR-D-11-00121.1, https://doi.org/10.1175/MWR-D-11-00121.1, 2012.

World Meteorological Organization: Technical Regulations Volume II: Meteorological service for international air navigation, https://library.wmo.int/pmb_ged/wmo_49-v2_2013_en.pdf, 2013.