Author Response Revision2 - amt-2022-236
Drone-based meteorological observations up to the tropopause
Konrad Benedikt Bärfuss et al., Atmos. Meas. Tech. Discuss.
https://doi.org/10.5194/amt-2022-236, 2022

Dear Editor,

In the following, you will find a list of the referee comments and all associated changes within the manuscript in addition to the diff-file. Please note that line numbers of changes based on reviewer comments are now associated to the plain revised document (not the diff-file).

As a major change, you will now find a well-organized uncertainty discussion starting with details on the correction algorithms. As these techniques are applicable to other drone-based meteorological observations, we think that the value of the additional sections more than compensates the increased length of the article.

We hope that the quite brief notes on what changes were initiated by which of the reviewer's comments are in your interest.

Now let's start:

Referee Report #1

The Authors would like to thank the referee for taking the time to again review our work.

The main criticism of my first review was the lack of a proper sensor description and validation. This includes an error estimation, as the sensors have to give reliable readings under extreme conditions during the flight, including a wide range of temperature, humidity and wind speeds, which is usually not the case in the atmospheric boundary layer. Unfortunately, from my point of view, this is still not adequately and sufficiently addressed in the revised version. However, I expect this for publication in a journal for atmospheric measurement techniques.

We now added a detailed description of postprocessing methods in Sect. 2.5, an explanation to the calibration techniques in Sect. 2.6, and a dedicated error discussion in Sect. 2.7.1.

Furthermore, the new sections implemented to compensate for measurement errors contain equations that are inadequately presented. Moreover, these corrections are not adequately validated.

We intensely overworked and completed the presentation of the measurement error compensation and other post processing techniques which now is a substantial part of the article (see above).

E.g.
• Sect. 2.3 makes assumptions about the inaccuracy of the wind estimation, but there is no systematic analysis or error estimate. Further, in Sect. 2.5.3 a transformation of the aerodynamic coordinate system to the geodetic coordinate system is not appropriately demonstrated, and simplifications and speculations are made that are not comprehensible to me as a reader.

A systematic analysis and a complete description of the wind estimation algorithms including simplifications (also those applied for AMDAR) and their impact on the results are now presented in Sect. 2.5.3 and applied on the drone measurements in Sect. 2.7.1.

• In Sect. 2.5.1. a temperature correction is carried out. This should also contain the dynamic pressure (air speed).
We now recap the standard approach for the correction of the temperature rise caused by compressible effects and discuss the error when incompressibility is assumed (as it is done in this article) in Section 2.5.1.

• Section 2.5.2 introduces an approximation using a power law and specific estimates however, a comprehensible determination of the parameters is also missing here.

The additional Section 2.6 now introduces the calibration/validation method used to estimate the parameters for power law representing the time constant for humidity readings.

Minor remark: In a scientific publication, the official term uncrewed aircraft system (UAS), which legislators and the ICAO use, should be used rather than the still colloquial and military term 'drone'. Using 'uncrewed' also enables the term to be gendered.

Unfortunately, there is no common agreement on the terms to be used for unmanned aerial systems. EASA, ICAO and WMO use the term UAS or even wxUAS for weather sensing UAS (Pinto et. al. 2021). In order to make the article findable by a community searching for other terms, we now at least introduce the term UAS: "Drones (also called unmanned or uncrewed aircraft systems, UAS, or remotely piloted aircraft systems, RPAS)"

Referee Report #2

The authors would like to thank the reviewer for the detailed comments and the interesting discussion. In the following the raised issues are answered point by point. In case line numbers are mentioned by the authors, they refer to the revised version of the manuscript (without track changes).

As I said in my previous review, I am a strong supporter of the utility and benefit of automated aircraft reports to fill the time and space gaps left between other in situ observing systems. Upon rereading the article "Drone-based meteorological observations up to the tropopause" by Barfüss, Schmithüsen and Lampert, I continue feel that it is not yet ready for publication. Although I appreciate that the authors have attempted to address some of my concerns, the text continues to read much more like a "Concept Study" or perhaps "A Report on Preliminary Project Activities" than providing significant scientific study that would support publication on scientific merit as it stands. As such, it needs to be revised and improved before it can be accepted for publication. Some of my reasoning for this decision is summarized below. I apologize for some possible disjointedness in the review, but that may in places reflect the need for more organization in the paper's organization.

We consider the fact that we reached the altitude of 10 km with such a small system and own propulsion a major step towards expanding future radio sounding capabilities based on drones. This provides new opportunities for the branch of meteorological measurements, and therefore we would like to present this new tool to the scientific community. As far as we know, this constitutes even a world record. In addition to this mainly technical achievement, we even included first simple meteorological instrumentation to demonstrate the scientific benefit of this kind of data.

To start, I suggest that the authors retitle the work to include the words "Concept Study" in the title. That will make it absolutely clear to the reader that they should be expecting to see preliminary results, not extensive validation.

As suggested, we changed the title to "Drone-based meteorological observations up to the tropopause - a concept study".

The introduction remains essentially unchanged and includes very little discussion of problems that other authors have documented with automated reports from commercial aircraft that should provide a far more stable platform for collecting data and providing representative measurements or the atmosphere without being affected by artifacts related to aircraft stability. For example, AMDAR wind observations from longer-range aircraft are much more accurate (by a factor of 2) than TAMDAR reports obtained from smaller/lighter regional jets.

We consider that the AMDAR/TAMDAR measurements are not the object of our investigations, and therefore would prefer not to mention too many details in the introduction. The link to aircraft stability is misleading, as artefacts are seen to be caused by the inertial navigation system in literature, see Moninger et al. 2010, which states that “The lower quality of the wind data from TAMDAR is likely due to the less accurate heading information provided to TAMDAR by the Saab-340b avionics system. Accurate heading information is required for the wind calculation, and the Saab heading sensor is magnetic and known to be less accurate than the heading sensors commonly used on large jets.”

Possibly more distracting to is the statement in the first sentence of the Abstract reads “with large data gaps in the atmospheric boundary layer, above the oceans and in polar regions”, with only a brief mention of that “the feasibility of reaching an altitude of 10 km with a small meteorologically equipped drone is shown.” Through the remainder of the text, the utility of the drone observations, however, is judged by the ability to reach 10km in a polar environment. The Abstract should be revised accordingly.

We agree that this is misleading, and revised the first sentences of the abstract. The text in the abstract has been changed to:

“The main in-situ data base for numerical weather prediction currently relies on radiosonde and airliner observations, with large systematic data gaps, horizontally in certain countries, above the oceans and in polar regions, and vertically in the rapidly changing atmospheric boundary layer, but also up to the tropopause in areas with low air traffic.”

Grammar and word choice continue to present problems throughout the paper. For example, the end of the sentence that reads “whereas data from level flight is . . .” in line 49 of the manuscript is confusing to read. Using “with additional observations provided at flight level” would have read more clearly and been more concise.

We changed the text in l. 50 to “with additional observations at flight level provided during cruise”.

This is one of many places where the text should be reviewed grammatically using a tool like the “Editor” included in Microsoft Word.

We would like to thank the reviewer for suggestions for improvement of language, and consequently checked the text using language tools for both spelling and grammar.

This should help improve the conciseness of the text and suggest more appropriate word choices in places. The authors also often incorrectly use the noun “data” as being singular (followed by ‘is’ instead of ‘are”).

We now use data as a plural noun throughout the text as preferred in general scientific writing.

As further point of confusion for me in reading the author’s response to the first review are the errors line numbers noted by the authors. For example, a refence is made early in the response to my review to section line 232 in Section 2.5.3, but that section only extends from lines 314 to 335. The reference about AMDAR/TAMDAR in line 344 can’t be found (except possibly much later in line 446). This line numbering inconsistency is present at other points in the paper as well, making it harder to review.

The line numbers change with the revision of the text, of course. Our line numbering referred to the document with the tracked changes. In the current review we refer to the plain revised manuscript, not the track change document.

I thank the authors for focusing on using “drone” and “LUCA”.

Unfortunately, there is no common agreement on the terms to be used for unmanned aerial systems. EASA, ICAO and WMO use the term UAS or even wxUAS for weather sensing UAS (Pinto et. al. 2021). In order to make the article findable by a community searching for other terms, we now at least introduce the term UAS in l. 86: “Drones (also called unmanned or uncrewed aircraft systems, UAS, or remotely piloted aircraft systems, RPAS)”

Please use metric system measurements throughout the paper, removing conversions to other units (such as kn).

Thank you for insisting on this – we were not aware of the fact that wind barbs are actually also defined in m/s by the WMO (Manual on Codes) and removed the unit “kn” throughout the manuscript.

The term “integral mean wind speed” in line 167 needs to be defined. Also, if the purpose of this restriction is so that “the drone does not have to stay above the launch point, but must be able to return to the base” as stated in line 170, isn’t it more important that the total vector displacement of the drone from its launch site be limited, not just the mean wind speed? Please explain further and restate what this means about limits on wind speeds that will not be observed. This needs to be clearly referenced again at the end of the paper under limitations.

We changed the text in l. 167 to: “As a trade-off for LUCA, the system was designed for operation in the temperature range between -75°C and +30°C, and for a wind speed of less than 28 m/s over the whole vertical profile to limit the maximum vector displacement of the drone from its launch site.”

Line 171 is changed to “Applying wind speed conditions such that the drone does not have to stay above the launch point, but must be able to return to the base, the mean wind speed over the atmospheric profile to be observed by the drone should not exceed the nominal horizontal airspeed component of 28 m/s.”

In the conclusion section, l. 743, we added: “The limitations concerning wind speed are related to the current maximum airspeed of 50 m/s.”

Lines 154-159. Figure 1 should be discussed after the integral speed limit is defined and explained, not before.
Although the minimum takeoff speed and ascent rate of the drone are mentioned, but it is not easy to find the average airspeed of the drone. Please provide this, along with how greatly and rapidly it varies during flight.

Figure 1 was our starting point for designing the aircraft. We used the statistical analyses of environmental conditions for the two sites Neumayer and Lindenberg in order to see what temperature and wind speed range the system has to endure to be competitive with radiosonde launches. We now state at the beginning of the Section 2 (l. 148):

“In the following, the process towards the design of the drone of type LUCA is presented briefly. Requirements for the system are derived from the environmental conditions to be expected, to obtain a high availability of measurements. Based on the environmental conditions for two sites, the design of the mission and of the drone is introduced. The simplistic sensor package that was used for the demonstration flights is described, including uncertainty aspects. The process of obtaining flight permissions for such altitudes is presented. The methods for data post processing are presented, and the obtained data quality is assessed. The section closes with a note on the variability of the atmosphere.”

Concerning airspeed, we included in the text (l. 218): “As the result of a multi-variant optimization for profiling the atmosphere vertically up to 10 km, the design weighs 5-6 kg, depending on the deployed sensor package, has a wingspan less than 2 m, and operates at a constant airspeed of 28 m/s controlled by the autopilot system.”

Table 1 is still not referenced in the text. It should be and it also be supplemented to include information about reporting frequency and representativeness and drone flight speeds etc.

Thank you for mentioning the missing reference. We now include in the text (l. 236): “For the measurements, different sensors were applied. An overview of the sensors and their accuracy according to data sheet is provided in Table 1.”

We state in Sect. 3.1 l. 625 the measurement frequency: “The data are available by remote transfer with a temporal resolution of 1 Hz. The full data set with a resolution of up to 25 Hz can be downloaded after landing and is pre-processed automatically for upload into the GTS.”

As this is independent of airspeed, we do not think that airspeed fits in the table about sensor characteristics.

The text (and Table) should also be clearly stated that wind tunnel tests have not been conducted to assess and document whether the installation approaches would degrade the observations further. The reader can only assume that no such tests were done of the LUCA installation approach.

We changed the text in l. 236 to:

“For the measurements, different sensors were applied. An overview of the sensors and their accuracy according to data sheet is provided in Tab. 1. The placement of the sensors is based on the experience with similar drone-based systems (e.g. Bärfuss et al., 2018, Lampert et al., 2020, Bärfuss et al., 2021), and the sensor behaviour and possibilities of correcting e.g. sensor response time, are well known. The performance of sensors and the data quality is assessed directly from the atmospheric measurements, without conducting simulations or wind tunnel tests for the overall setup.”

Additionally, we clarified the importance of inflight calibration as the most important part within the Simulation – Wind Tunnel – In-flight ecosystem in aerospace science in Sect 2.6.

Lines 201-212. Please define data what is precisely is meant by ‘low frequency’ and how this differs from the practical limits on reporting frequency of independent drone observations in the final meteorological products.

We changed the text (l. 208) to: “During the flight, 1 Hz real time data are available.”

Later in the revised text, it is noted that the time constant for response time of the humidity sensors are “15 s at 0 km, 60 s at 3 km and 2000 s at 10 km”. Because 2000 seconds (33 minutes) is longer than the entire flight, does this mean that the moisture data in the cold half of the sounding may in fact represent only 1 deep layer of the atmosphere layer, even though the profiles shown in Figure 7 show more detail than this? Please explain how the detail can be accurate if the response times are so long. This could explain the differences noted near 550 hPa and above in both profiles in Fig7.

In fact, an immediate response will be measured even with such large time constants, which follows a fundamental law in sensing (and general exchange processes), but the value of the measurement is affected by the time constant. We now recap the control theory basics in Sect. 2.5 and describe the time lag correction of humidity in detail in Sect. 2.2.

The authors then downplay the importance of moisture observations in NWP improvement, sighting results from a global model (and possibly for longer forecast ranges).

We are not sure which section or sentence the reviewer is referring to. Maybe the following part of the manuscript (l. 778): “The quality of drone observations for sampling the complete troposphere is of superior importance, and could possibly contribute to climate applications. Regarding the involved variables, one finds that climate users tend to focus on temperature and humidity data (Ingleby et al., 2022). For NWP applications, wind has arguably more than twice the impact on the quality of short range forecast (Ingleby et al., 2021)."

It is not our intention to downplay the importance of moisture observations. The NWP applications mentioned in Ingleby et al. (2021) refer to data denial tests in the ECMWF data assimilation system, and we changed the phrase in l. 782 to "For NWP applications, wind observations have arguably more than twice the impact on the quality of short range forecast compared to temperature observations (Ingleby et al., 2021)."

Similar data denial tests using other models (global and regional) show notable improvement is the shortest forecast times. That should not be ignored.

It would be great to get the references that are mentioned here, and then we can take them into account in the manuscript.

Section 2.5.3 needs to include much more information about the quality of the measurements going into the wind calculation, such as the frequency/precision of the drone location information (and derived drove velocity vector) along its flight path, the frequency and precision of the true air speed vector relative to the drone, along with a sample how these factors can affect the wind reports.

We now dedicate new sections to the calculation of the wind (Sect. 2.5.3) and error estimation (Sect. 2.7.1).

I was intrigued by the opportunity presented by the flight paths shown in Figure 4. It would be very instructive to study the variability of the meteorological observations during the periods of circling flight. This should be easy to do. The variability of the derived meteorological products during these loops should be quite consistent and seeing the error structures within the sets of drone reports could provide useful insight into the expected data quality.

We totally agree on the idea that measurements shall not correlate with the flight direction during e.g. a circle in constant altitude (neglecting turbulence). This technique is often used for in-flight calibration of sensors (even in geophysics). Upon your comment, we decided to mention/recap inflight calibration/validation techniques in the new Section 2.6.

For the drone measurements presented in the manuscript, comparing ascents with descents is the preferred method for the calibration/validation of temperature and humidity. Wind speed unfortunately is corrupt during flight phases where the engine was used (magnetometer data are distorted by the electrical current), which was also the case during the circles you mention. Therefore, ascent wind data (generally data under "power") are regarded as corrupt and not discussed further, except the illustrative statistics in Table 2.

Errors from factors discussed in lines 244-250 should also be noted in discussion about Table 1.

We now elaborate theoretical errors in Sect. 2.7.1 in detail and decided to keep Table 1 as simple as possible.

The authors now reference work by VAISALA (2021) regarding ventilation effects and instrument response lags as a basis from their temperature and humidity corrections. At no point, however, is the validity of the Vaisala method us affected by the increased ventilation that could be present in a drone moving both vertically and horizontally through the atmosphere at higher speeds than a drifting balloon. This shouldn't be ignored.

We take this comment as indication that we did not properly introduce the control theory essentials for this correction, and dedicate the first part of Sect. 2.5 to the underlying fundamentals of sensing and time response.

I also expected to see a more quantitative discussion of how the wind measurements were derived from the aircraft and how they were affected by turbulence, especially below and near the level of the jet stream, along with a more detailed error analysis.

We now dedicate a whole subsection (Sect. 2.5.3 to the wind calculation, which is based on the vector difference between the airspeed and the ground speed. The wind measurements are not affected by turbulence, instead, turbulence parameters can be calculated from the data.

I also still don't see any evidence that any object that is 2 m in length and 5-6 kg in mass will not be 'tossed around' in areas of large shear and turbulence near the jet stream and in the boundary layer.

"Tossing around" is a quite illustrative term, so let us argue less scientifically: We don't see any evidence why the aircraft should be tossed around. Regarding dynamic soaring with unmanned aerial vehicles (3 m length at 9 kg mass, similar wing loading) in extreme wind shear and heavy turbulence (reaching an airspeed of the UAS over 500 mph at over 50 mph wind speed, which albeit is a significantly lower wind speed than to be expected within a jet stream core), you might be convinced by this video that drones are not "tossed around": https://www.youtube.com/watch?v=GCVK3w5DHbk

Similarly, since the authors state that the temperature reading may be affected by heat generated within the drone itself, I'm still not convinced by the current text that the readings are likely to be contaminated as a function of whether the drone is flying into the wind (in which case the heat generated by the drone should be moved away from the rear of the aircraft and away from the nose mounted sensors) or with a weak tailwind (in which case the heat could remain in the proximity of the drone and possible affect the sensors).

We disagree on this as an aircraft is flying with a certain speed with respect to the air (completely decoupled from the wind which only affects the heading of the aircraft to reach a certain earth-fixed point) and do not know how to resolve the misconception here.

As noted in the previous review, Section 2.7 and Figure 5, though of some interest, do not add to the substance of this paper and should be removed, especially because the data used are at time and space scaled substantially later than those of individual drone flights.

We do not agree on this point, as the figure is not intended to show variability during one flight, but variability within a day – enabling a discussion of the benefit of multiple launches per day. It might be obvious for scientists working on e.g. atmospheric modelling and data assimilation that the variability of the atmosphere decreases above the tropopause (except the diurnal cycle for e.g. temperature is usually well described by models), but not necessarily for others working e.g. in the field of drone- and sensor design.

Also, please remove the ERA5 curves from Fig 7, since they do not add value to the discussion.

The idea of our measurement system is to provide data for the modelling community. As ERA5 has the capability of highly flexible data assimilation, the drone data could be easily integrated here. Therefore we would like to show the comparison. We do consider it interesting that the T profiles agree rather well with radiosonde and drone data, but the altitude and strength of temperature inversions within the boundary layer is not captured properly. Differences in the mixing ratio are more pronounced.

It would be more valuable to compare the differences/similarities between the LUCA profiles obtained during ascent with those obtained during descent.

We agree regarding the parameters T and RH. For wind, the ascent is not regarded as an observation due to (expectable) magnetic interference. We clarify this now in Table 2 and in l. 679 ff.

Finally, summary statistics on the differences between the LUCA profiles and each radiosonde data set should be included on each panel.

We now add the error statistics in the panels of Figure 11 along with the estimated uncertainties.

Also, in Figure 7, winds are plotted about every 300 m in height. Is this the how frequently they are available? If there were more drone wind report available, what was the variability of the drone winds between the reported levels?

We now mention the wind observation frequency (25 Hz) in l. 625. The variability of the wind measurements between the reported levels (1 m resolution) is graphically shown in Figure 11.

Regarding Figure 8 and supporting text, please explain what is meant by "virtual air parcel position was determined by virtually back-shifting the probed air parcel" in the figure caption, especially since the term 'back-shifting' if not used in Section 2.6 as suggested in the text.

We changed the text in line 570ff that this is now accessible.

I commend the authors for adding and a discussing Table 2. However, it is not clear stated that the histograms used all drone reports or only those shown in Figure 7, nor does it explain why there are twice as many data matchup for descents as for ascents, especially since the ascent and descent rates were said to be similar and since they suggested earlier that quality control limits were not imposed on the drone data sets. Also, it would be good to include gross statistics including both ascent and descent matchups. Finally, the more negative (less positive) biases shown during ascent are opposite from what would be expected due to hysteresis effects.

The hysteresis is completely cancelled out by the time lag correction, which is now discussed in detail in Section 2.5. As the matchups follow the launch times of the radiosondes in respect to the drone trajectory and the wind, we omit the explanation of the data matchup numbers and take it as descriptive values. An elaboration of these matchup numbers would be out of the scope of our manuscript.

Finally, the captions should describe the structure of the figures, not an analysis of their content. That should be in the text.

We moved the analysis part of Fig. 7 (previously Fig. 6) into the text section (line 593ff).

In addition, without a substantial greater number of further experiments and tests, the last sentence of the caption for Table 2 stating that "This indicates a possibly higher accuracy of the drone observations than observations within the AMDAR/- TAMDAR programme" is conjecture and must be removed.

We removed the sentence from the caption of Table 2 as suggested.

Though I'd expected to see mention of ascent/descent differences earlier in the text, as my earlier comments said, it is instructive, especially the decreased quality (increased sigmas) of the wind directions and speeds during ascent.

We would prefer to keep the analysis in this section. Differences for ascent and descent as well as the usability of comparing both measurement directions are mentioned in Section 2.6 and in lines 262, 665, 679 and Table 2.

However, the statement in lines 646-448 that "This indicates a possibly higher accuracy of the drone observations compared to operational airliner observations already assimilated into NWP models" must again be modified and weakened, since the limited study presented here used only a subset of altitude ranges and substantially limited the number of very high-speed wind reports that were available. Similarly, the specific humidity statistics are in improved in part because the drone samples were taken primarily at low temperature and therefore low saturation specific limits.

We changed the text (l 681) to:

“For the descent profile, the average differences as well as the standard deviation of the differences between drone observations and radiosonde observations are in the range of (or even below) statistical measures shown in the study of Wagner et al., 2021, which includes a comparison of radiosonde data with observation data from the AMDAR/TAMDAR programme.”

Also, the authors seem focused on moisture in the middle-upper troposphere, saying that it is particularly important. I’ll grant that moisture can have some influence at these levels, but the amount of moisture, its vertical structure and horizontal structure are much more important for forecasting/studying weather events from heavy precipitation and flooding to severe storms.

We are not sure which sentence the reviewer is referring to with the comment; therefore it is difficult to improve the text for making the statement clearer.

We agree that PBL observations are of superior importance for weather events, but think that the limited abilities of recent UAS platforms tended to disregard the advantages of sounding the middle-upper troposphere for weather event studies/forecast.

As a result of inadequacies like these throughout the text, it read still reads more like a preliminary progress report instead of a quantitatively verified scientific contribution to the literature. Although I support concept presented here, the paper is still not ready for publication, though a change in title would provide a major step toward defining the purpose of the effort.

We worked on the points that were raised, and in particular tried to make clearer the main intention, which is to introduce the achievement of a drone system capable of soundings up to 10 km altitude, with the potential to be used complementarily to radiosondes. We did not apply the perfect meteorological instrumentation for this, but this is left for the future.

Again, my intent here is to provide constructive suggestions for improving the paper and look forward to seeing a revised version.

We would like to thank the reviewer again for the detailed comments. We took into account most of them, but could not agree with every point. Including an uncertainty discussion was a huge effort, but also adds substantial value to the manuscript.

References:

Moninger, W. R., Benjamin, S. G., Jamison, B. D., Schlatter, T. W., Smith, T. L., & Szoke, E. J. (2010). Evaluation of Regional Aircraft Observations Using TAMDAR, *Weather and Forecasting*, *25*(2), 627-645. Retrieved Jan 10, 2023, from https://journals.ametsoc.org/view/journals/wefo/25/2/2009waf2222321_1.xml