The authors thank the reviewer for the well thought out and constructive comments on the manuscript. All replies to these comments are inserted below (in yellow) and information is provided, how the manuscript has been changed accordingly (in green). In the text the new parts are highlighted in grey.

General Comments

The paper describes a new dropsonde system – the KITsonde, which has novelties such as (1) multiple sounding profiles from a single launch, (2) a modular payload design thanks to a release-container concept and (3) reception from up to 30 channels simultaneously. The paper is well-written and coherent. The structure makes sense, starting with a background on dropsonde development, detailed description of the system itself, the different possible configurations and the tests done in the field. Overall, the measurement techniques described herein will improve the type of atmospheric soundings being taken currently as well as the aircraft strategies being employed. The KITsonde holds exciting potential for advancing atmospheric observations especially with the three aforementioned novelties.

However, I do have a few comments, which I wish the authors would clarify.

1. The current design is compatible with dispensers of RD-series dropsondes, which is great for integrating with existing aircraft systems. However, these dispensers are being phased out as NCAR and Vaisala transition to the NRD series (e.g., NRD-41). The newer, smaller sondes require research aircraft teams (incl. HALO) to modify or replace dispensers, often adopting the automated launcher systems. This is a crucial point because a smaller release container for the KITsonde would impact the first two novelties I list in the general comments. As the older dispensing systems are becoming obsolete, the KITsonde's compatibility with them is no longer an advantage. Could the authors address the KITsonde's compatibility with the new dispensers, either for HALO specifically or more broadly?

We agree with the reviewer's remarks. The consequences are, however, only, that the KITsonde cannot be used with aircraft solely offering the smaller dispenser. Present aircraft have the necessary opening for the larger dispenser and – as confirmed for HALO - a kind of adapter will be used to reduce the open diameter (the hole) within the aircraft skin. Both the old and new dispenser can be alternatively used even in the future. New aircraft may offer dispensers only for the new form factor. For those cases, single EL18 dropsondes are small enough to be dropped.

The raised issue is addressed in the future outlook section as follows (lines 537-540):

With a gradual transition to smaller dispensers for the NRD series (e.g., NRD-41) by NCAR and Vaisala, only single EL18 dropsondes are small enough to be used. Since present dropsonde-releasing aircraft have the necessary opening for the larger dispenser, a kind of adapter, as confirmed for HALO, will be used to reduce the open diameter (the hole) within the aircraft skin. In this case, both the old and new dispenser can be used in the future.

2. The variability in the parachute design is unclear to me and the paper will benefit from better description. For instance, phrases like "individually sized parachutes" (L97) and "different sizes of the parachutes" (L469) tell me that parachutes may vary in size to maintain separation in vertical space. If so, what are the possible

sizes and configuration options? This information is critical for planning the launches because sounding strategies depend on the descent rate, which affects how closely the sounding trajectory approximates a vertical profile as well as how closely the sounding approximates to an “instantaneous” profile. For the KITsonde, in a 4-sonde release container, what parachute sizes are used and what are the expected descent rates for the 4 sondes, and is this configuration standard or could the user choose? Similarly, for a single-sonde release container (with or without other payloads), what is expected descent rate? H

We agree with the comment. The description of the parachutes is now extended (including some wording of the reviewer) and an analysis of the fall speeds for the different sizes has been performed, and a related table of results is added to the manuscript as a new subsection 2.3.2:

2.3.2 Parachutes

The parachutes are chosen to determine the descent rates, which affects how closely the sounding trajectory approximates a vertical profile as well as how closely the sounding approximates to an “instantaneous” profile. The parachutes of 64 cm in diameter being used are manufactured by BBL Elektronik & Aeromet GmbH (https://www.meteorologyshop.eu/en/balloons/radiosonde-balloons/368/meteorological-parachute-pc-055). To allow for different fall speeds, the effective areas of the parachutes were reduced by cutting away outer trapezoidal sectors of fabric between the holding lines. Tab. 2 shows the effective diameters, corresponding to the effective area of the parachute. Fall speeds versus parachute size were analysed for the SouthTRAC campaign (see Sect. 3), where 60 meteorological sondes EL-18 without any additional sensor or communication electronics were used.

Table 2: Mean fall speed and standard deviations in m s^{-1} of meteorological sondes dropped with parachutes of different effective diameters. The fall speeds are given for different height ranges during the SouthTRAC campaign.

Height range (km)	Parachute effective diameter (cm) in line 2 and area (cm^2) in line 3 for the meteorological sonde EL-18 of 73 g weight			
	64	40	32	24
	3217	1257	804	452
8 - 12	4,5±1,2	7,0±2,1	8,1±2,1	11,2±1,6
4 - 8	3,3±1,0	4,9±1,5	5,9±1,4	8,7±1,0
0 - 4	2,6±0,6	3,8±1,2	4,8±1,1	7,3±0,5

The parachutes for the release container, for the coupled radioactivity/meteorological measurements, and for the particle/meteorological measurements are more robust and of a x-pentamine shape. They are manufactured by Spekon Co.

(https://spekon.de/seilschirme.html) and consist of 5 quadratic sections, each 15 x 15 cm. The total area is 1125 cm^2.

3. This third point relates slightly to the previous one in terms of descent rate. I am not completely convinced by the motivation for the satellite modem configuration. Currently (e.g. Ehrlich et al 2024, https://doi.org/10.5194/essd-2024-281) HALO dropsondes are sent to GTS in near real-time, i.e. as soon as the dropsonde makes a landing and therefore can also be sent to ground-support. Generally, the difference in real-time and near real-time is around 12-15 minutes and should not affect data assimilation too much. Of course this changes if the descent time is close to 45-60 minutes, which means that the aircraft telemetry could be out of range. But this advantage of the satellite telemetry then comes at the cost of a multi-sounding launch, which provides the novel spatio-temporal density I would argue is the best feature of the KITsonde. It does work as an example of different payload capabilities, but I struggle to find a practical use-case for it where it is advantageous over the conventional sondes, such as what the UCASS and radioactive payloads demonstrate. I would appreciate the authors' clarification here.

The arguments are generally correct, as far as that the option of satellite communication is not highly necessary or a big advantage to available systems, because it comes at the cost of a multi-sounding launch. The satellite link was tested only once and is not in central focus. It was developed to allow for small descend speeds, when a fast aircraft like HALO may lose the communication link. It also maintains future options of dropping sondes from normal, non-research aircraft without a signal receiving unit. NRT assimilation of very high spatio-temporal multiple dropsonde data like KITsonde are anyhow questionable for operational DA. A „one sonde per container" is preferable for that case, and KITsonde would provide comparable information to the NCAR/VAISALA system.

We modify the sentence in line 297-301:
"Using the satcom module is only recommended for cases when the descent of the sonde is slow, causing a descent time of 45-60 minutes, and the aircraft would need circling to keep telemetry contact. The satellite communication allows only one meteorological sonde to be dropped with the release container (Fig. 1) so that it comes at the cost of a multi-sounding launch. The satcom module would also allow for dropping sondes from normal, non-research aircraft without a signal receiving unit."

Minor comments

Title: Why HALO specifically when the system has already been demonstrated with 2 other aircraft too?

The use on HALO is the final goal, and the development was performed and mostly funded in the context of the HALO consortium.

But we agree, that the KITsonde is currently used from other aircraft and may be even used more in the future, and drop HALO from the title.

L45-47: Acronyms are not defined

Acronyms are checked and modified as follows (new lines 45-49):

Early attempts to use dropsondes with the capability to measure wind based on the very low frequency radio navigation system OMEGA were made in the 1970s, continued by also using the navigation system LORAN-C in the 1980s. In 1994, the American National Center for Atmospheric Research (NCAR) and National Oceanic and Atmospheric Administration (NOAA) as well as the German Aerospace Center (DLR) agreed to develop a GPS dropsonde (RD93) based on Vaisala radiosonde technology.

L49 : Unclear why the horizontal spacing of 100 km between drops? Is there a reference for explanation?

The cited spacing of 100 km is from grey literature, which – when checking again – could not be found for proper citation.

We remove the part of the sentence “with the telemetry system allowing for a horizontal spacing of 100 km between drops.”

L52 : NWS here is particularly the US National Weather Service as opposed to different countries’ weather services when defined in L44.

Checked and modified by replacing NWS by US National Weather Service

L81: What type of flexibility in operations?

Checked and modified by deleting “in operations” because it is clear that the flexibility refers to aircraft or platforms.

L99-100: This is an excellent and very pragmatic advantage in favour of having the release container concept. :)

L157 : Is there a reference for 250 m/s? I believe 200 m/s might be closer to cruising speed during flight operations, but I am not familiar with all HALO payload configurations.

We had a view into the SouthTRAC data. For flight sections ST23a und b the average „Speed over Ground“ was 193 and 180 m/s, and „True Air Speed“ was 242 and 236 m/s.

We replace 250 m/s by 240 m/s.

L194: Please mention the weights and CoG for the standard 4-sonde configuration?

We added: The container with 4 EL-18 sondes has it’s centre of gravity at 195 mm from the bottom. The weight of the release container with 4 sondes and parachutes is 761 g.

L207: How many minimum satellite connections are needed for this?

We assume the comment refers to "The cold-start time (time to first fix) is 26 s under good conditions (open sky)."

Time to first fix is primarily dependent on acquisition time and the time the receiver needs to obtain enough of the almanac and ephimeris to be able to provide a valid navigation solution. The latter is also dependent on the number of satellites received. The manufacturer does not specify which number of satellites needs to be in view for the 26 s figure in the datasheet to be valid. However, experience shows that a time of less than 30 s can usually achieved with 10-15 satellites in view. The navigation engine itself has a 72 channel receiver.

Slightly extended text (lines 220-225):
“For wind and position retrieval, the GNSS receiver u-blox MAX-M8C with 72 channels is used. It can receive various combinations of the L1 signals from GPS, GLONASS, and Beidou.” ….. “Experience shows that a time of less than 30 s can be usually achieved with 10-15 satellites in view.”

L210: PTU 1.12 s (and wind 1 s)… Discrepancy with the abstract, where 1.2 s is mentioned for both.

1.12 s and 1 s are correct and values are corrected in the abstract.

L230: Humidity above 100% RH stated, but table shows

We cannot claim to measure supersaturation quantitatively, so we correct the wording.

Replaced in Table 1 „Range of suitable measurements” | “0 – 100 %RH

L232-233: It is unclear why the heated sensor is not used and why it is less relevant than to radiosondes. Could riming not be a case for when heated sensors prove to be useful?

The dropsonde measures and reports the temperature of the humidity sensor. As the temperature sensor cannot be collocated on the humidity sensor, there is a small space between the two sensors, which leads to a larger uncertainty component than with the heated sensor. Nevertheless, the temperature data can be used to correct the humidity measurement for the temperature of the humidity sensor. We have rephrased the sentence to highlight this better. This setup is also able to detect situations, where the humidity sensor might experience more wet-bulbing than the ambient temperature sensor. The main difference between the two sensors for cloud-exits is the faster response time of the heated sensor. However, in the dropsonde application, response time decreases during flight progression, inverse to radiosonde operation.

Modified text (in lines 240-247):

The humidity sensor uses the same type of polymer but is unheated and has its accompanying temperature sensor located a few mm apart (Fig. 5b). The heated humidity sensor of the latest DFM-17 generation was not yet available at the time when the EL-18 was designed. The close temperature sensor can be used to correct the humidity measurement for the humidity sensor temperature. Still, this slightly reduces dynamic performance and response time, especially after the sensor has experienced precipitation. However, as the sensor response time decreases during the sounding, inverse to radiosonde operation, this problem is not as severe as with radiosoundings. As for most radiosonde algorithms (Dirksen et al. 2024), an active clipping of measured relative humidity values > 100 %RH is performed. Humidity values before clipping of up to 120% could be observed, which is a typical value.

L248: Are the interruptions often enough to justify an SD card buffer (thereby increasing waste and expense per sonde)?

During the tests the interruptions were frequent enough to justify a SD buffer card, which costs only 4 €.

We added in lines 296-297: This buffering system is necessary to avoid data loss when the connection to the satellite is interrupted. The data losses happened during the tests.

L310-312: Have these comparisons been documented somewhere? Could references be provided?

We wanted to mention that numerous tests and checks have been done in all stages of development. They were partly just for functionality and mostly only qualitatively. The results took influence on modifications and improvements. No systematic documentation has been done and no citable references are available.

L351: anomaly with respect to what?

As already written in the figure caption we add on that and clarify the sentence as (line 415): "…. we show the potential temperature anomaly (differences from the mean of all shown profiles) and ….."

L368: I believe "clear differences" is questionable phrasing except in the case of humidity.

Comment is accepted.

As we discuss the differences in the subsequence sentences in detail, the sentence is modified (in lines 434-435):
"The differences in the profiles of potential temperature, relative humidity, wind speed, and wind direction measured by the four sondes of container 1 (Figs. 10 a-d), which was released at 16:19 UTC, are as follows. "

L382: Shouldn't it be sonde 1c instead of 1b? For me, 1b looks like it went through the deepest cloud layer.

That is right and corrected as proposed.

L403: meso-gamma… Is it per Orlanski (1975)? Please define spatial extent or provide suitable reference.

Yes, we referred to Orlanski.

We add the scale and Orlanski in the text (lines 468-469): "…… captured the spatial heterogeneity of dynamic and thermodynamic conditions on the meso-gamma scale (2-20 km, Orlanski, 1975)."

Orlanski, I.: A rational subdivision of scales for atmospheric processes. BAMS, 56, 527-530, 1975.

L423: Where is the “independent modelling” part in Figs 13 and 14?

The “independent modelling” included the MACC, NAAPS and BSC-DREAM8b models, among others. Unfortunately, some of the model data appears to be ephemeral and in the intervening decade has ceased to become publicly available. Two exceptions are the Barcelona Dust Regional Center and NAAPS archival datasets – see e.g. the optical depth data for 3rd August 2013:
https://www.nrlmry.navy.mil/aerosol/globaer/ops_01/europe/201308/2013080306_globaer_ops_europe.gif.

We have inserted the following explanatory text (485-487):
“including the MACC, BSC-DREAM8b and NAAPS models, the latter two archives available at the Barcelona Dust Regional Center (2025) and Naval Research Laboratory websites (2025), respectively.”

and the references:
Barcelona Dust Regional Center, Products: https://dust.aemet.es/products/, last access: 20 January 2025.
Naval Research Laboratory, NAAPS (Navy Aerosol Analysis and Prediction System): https://www.nrlmry.navy.mil/aerosol_web/, last access: 20 January 2025.

L444: Is there a suitable reference for such coarse aerosol origins locally at 2 km altitude?

The authors are not aware of any such published descriptions, the remark in question is based on some visual observations made during the recovery of the dropsondes, such as during sunset, when such a layer seemed to partly shade the sun disk.

Figure 14: Why is there a 2-hour difference between the compared measurements? And why was the 3-5 km altitude window chosen?

AERONET size distribution retrievals can only be made when full almucantar measurements can be carried out, which can only happen when the sky is sufficiently clear of clouds. Hence there can be wide temporal gaps between retrievals. Fortunately, in this case a successful measurement at the Leipzig AERONET site was made only about two hours after the sounding. The altitude window was chosen for the comparison of sonde U3 with the AERONET retrieval to include just the purported Saharan air layer, as indicated by discontinuities in both number concentration (not shown for sonde U3 but also visible for sonde U4, Figs. 11 and 12) and humidity profiles, especially to the exclusion of the boundary layer dust that was assumed to be of local origin specific to the drop area, as stated in lines 440-444.

No related text modification. An error was spotted in the text, line 452: "sonde U4 on 03 August" should be "sonde U3 on 03 August". The caption to Fig. 14 that the text refers to correctly gives "sonde U3".

L466: Typos in the first few words

Corrected: The KITsonde system meets all expectations

At multiple places, figures are not numbered in the same sequence as their appearance in the text.

The text was checked accordingly and figures are now numbered in the sequence of their appearance.