*We thank the reviewers for their constructive comments and valuable suggestions, which helped improve our analysis and manuscript.*

*We sincerely thank the handling editor and reviewers for their valuable feed, which we used to improve the quality of the manuscript. Based on the comments and suggestions from RC1 and RC3 (Dr. Bruce), we have carefully and substantially revised the manuscript. We have included a detailed response addressing each of the comments. The comments are laid out below in italicized font, and specific concerns have been numbered (in black). The response of authors is given in normal font, and changes/additions to the manuscript are given in the text (in blue and purple). The line numbers correspond to the revised manuscript without the track changes displayed. For precise details on the modifications made in the latest version, please consult the , authoer comments where track changes are enabled.*

**Response to Referee #1's comments**

***Revision of revised manuscript***

***Synopsis:***

*The quality of the text and the figures got a lot better in the revised manuscript.*
*It is a long but overall quite interesting read. The only remaining major point is a lack of assessment of the statistical significance of the assimilation experiment results. This was pointed out in the first review and was acknowledged by the author's response. Yet this point has not been adequately addressed. Another point that came into my mind is the question whether the drifting balloons could "escape" when drifting away over the ocean and perhaps the communication channel gets lost, and how far they could drift uncontrollably. It may be worth a sentence how this issue is addressed or that this is actually not an issue with this system.*

**Authors: -** We sincerely thank Referee #1 (RC1) for their positive assessment of our manuscript and for acknowledging the improvements in both the text and figures. We have carefully considered all the comments provided and have revised the manuscript accordingly.

Regarding the comment on the statistical significance of the assimilation experiment results—which was also raised in the initial review—we acknowledge that this is an important point. At this stage, the experiment has only been conducted for one month, which limits the ability to draw statistically robust conclusions. We are currently developing a comprehensive optimization plan that includes adjustments to

observation errors and data sparsity schemes. Following this, we plan to carry out a one-year assimilation experiment to more thoroughly evaluate the statistical significance of the results.

In addition, We acknowledge the RC1's concern about communication loss and uncontrolled drift. Please allow us to clarify that this has been considered in our system design. We have strict limits on the altitude and drift range of each balloon. Once the drift exceeds the designated area, the fuse will trigger and cut off the system, preventing the parachute and the payload from drifting uncontrollably. This point, addressed on Page 10, L250-L262 of revised manuscript.

And the primary data reception relies on a network of ground stations spaced about 150 km apart. To mitigate the risk over open waters, we are expanding this network by utilizing offshore oil platforms, which can receive data from balloons within a 200 km range. As a key contingency plan, we will incorporate the BeiDou-3 satellite navigation position and short-message service. This feature will allow balloons to transmit minimal datasets at a lower frequency when other communication links are unavailable, ensuring data retrieval and greatly expanding the operational domain and utility of the ADDRS.

Once again, we appreciate the RC1's constructive feedback, which has helped us improve the quality of this work.

*More specific review points:*

*1)RC1:Abstract: As was pointed out by reviewers, the assimilation experiments with ADDRS data revealed positive but (so far) not statistically significant results. This has been acknowledged by the authors in their response. In the abstract, however, there is still the very detailed description of the improvements (statistically not significant) and of the dates of operational implementation (not so important for the scientific reader.*
*To be better in balance with the text, one should shorten the last paragraph in lines 28-37 to something like:*
*Numerical assimilation and forecasting experiments showed a positive (albeit not yet significant) impact on forecast quality both for general cases and for Tropical cyclone cases. The operational launching and assimilation of more than 100 such radiosondes started on 1 July 2024 and the number of stations is still growing.*
*RC3:L35/36 "an operational experiment" - "a pre-operational experiment"*

*L37 "was operation over one year, which will achieve full operational capability by 2026.""provided data over one year, and will achieve full operational capability by*

*2026."*

**Authors: -** We thank the reviewer for this precise suggestion and have made the recommended changes accordingly. In this article, the assimilation application indeed only has one month of experimental data. The results therefore need to be evaluated more objectively and from multiple perspectives. In future work, we will place greater emphasis on assessing the statistical significance of the assimilation results. We agree with and accept the revision suggestions of RC1 and RC3. We have revised on Page 1, L28-31 "Numerical assimilation and forecasting experiments showed a positive (albeit not yet statistically significant) impact on forecast quality for both general cases and for tropical cyclone cases. The pre-operational launching and assimilation of more than 100 such radiosondes started on 1 July 2024 and provided data over one year, and the number of stations continues to grow."

We agree that the term "a pre-operational experiment" more accurately describes the nature of our study, which was conducted to evaluate the system's performance and potential prior to its full operational deployment. As suggested, we have replaced "an operational experiment" with "a pre-operational experiment" on [Page 1, L 30] of the revised manuscript and throughout the manuscript, including in the abstract, introduction, and conclusion sections. We believe this change enhances the clarity and accuracy of our work's description.

*2) RC1:L93: was organizational developed .. please omit "organizational"*

*RC3:"The ADDRS was organizational developed", delete "organizational"*

**Authors: -** We thank the two reviewers for their careful and thorough reading. We have removed the "organizational" [Page 3, L88] of revised manuscript.

*3) RC1:L115-116: "achieving an approximate vertical stability" sounds a bit strange. Perhaps rephrase to, ..with gravity, so that its height stays approximately constant.*
*RC3:L116/117 "vertical stability. Subsequently," I suggest deleting the sentence stating subsequently and just saying "vertical stability, starting the drift phase."*

**Authors: -** We acknowledge that some phrasing in the original submission may not have been perfectly clear due to our non-native English background.We thank both reviewers for their careful and thorough reading. We have revised "And the buoyancy of the inner balloon attains equilibrium with vertical stability, starting the drift phase."on Page 4, L108-L109 of revised manuscript.

*4) L130: The ground data-receiver ..*

**Authors: -** We thank the RC1 for this careful and thorough reading. We have revised the wording to "the operational ground data-receiver"on Page 5, L122 of revised

manuscript.

*5) L345ff: one should use integer percentages here, e.g. 78% instead of 77,68% since the uncertainty of this number is likely a lot larger than hundredths of a percent and readability is enhanced.*

**Authors: -** We thank the reviewer for this excellent suggestion. We agree that using integer percentages improves readability and is more appropriate given the measurement uncertainties. We have therefore revised the text at Lines 346–347 and 354–360 to report percentages as integers (e.g., 79% instead of 78.65%; 71% instead of 70.91%; etc.)..

*6)Although it has been acknowledged by the authors, there is still no mention in the results section of the revised manuscript that the positive assimilation results are not (yet) statistically significant. This is however an important element of good scientific practice and it is important to allow the reader a balanced assessment of the progress achieved. It is already a big success that the the positive results achieved so far led to an operational adoption of the system, and the authors could state that they are confident that they will achieve scientifically significant improvements in terms of several scores once the data records get longer and the number of case studies gets larger.*

**Authors: -** We sincerely thank the reviewer for this critical and constructive comment. We fully agree that it is essential for good scientific practice to clearly communicate the current lack of statistical significance of the positive assimilation results in the Results section, thereby providing a balanced assessment for the readers. We apologize for this oversight in the previous revision. Following the reviewer's excellent suggestion, we have now made the following two key additions to the manuscript:

In the Results section, specifically in the paragraph discussing the ETS scores, we have added a clear statement regarding statistical significance directly after presenting the ETS improvement rates (e.g., 0.04%, 0.7%, 2.2%), we have inserted the following sentence: "It is important to note that while these ETS improvements are consistent and positive, they have not yet reached conventional levels of statistical significance (e.g., $p < 0.05$) over the one-month test period, primarily due to the limited sample size of case studies. This underscores the preliminary yet promising nature of these findings."on [Page 17, L470-L473].

Furthermore, in the Discussion section, to provide a forward-looking perspective as suggested by the reviewer, we have added a statement This acknowledges the operational potential while framing the results appropriately for the scientific audience: "The positive impact of ADDRS data assimilation demonstrated here, even

if not yet statistically significant over a one-month sample, has shown considerable operational promise. We are confident that statistically significant improvements in precipitation forecast skills will be achieved as the data records lengthen and the number of case studies, particularly for extreme precipitation events, increases in the future."on [Page 22, L594-L598].

We believe these revisions have strengthened the manuscript by enhancing its scientific rigor and transparency, accurately presenting the current findings while outlining a clear and confident outlook for future validation. We are once again grateful to the reviewer for this invaluable guidance.

**Response to Referee #3's comments**

***General***

*1)This is an interesting but time-consuming manuscript to review, partly because of the novel technique described but also due to cultural and language differences.*
*It is improved from the first version but in some cases material in the "response" should have been included in the main manuscript.*

**Authors: -** We sincerely thank the reviewers for their helpful feedback regarding the language expression in our manuscript. We acknowledge that some phrasing in the original submission may not have been perfectly clear due to our non-native English background. To ensure the language meets the high standard required for publication, we have sought professional assistance to polish the entire manuscript. We believe the revised version is now much improved in terms of readability, and we hope it adequately addresses the concerns raised.

*2)I think that the importance of radiosondes in NWP and particularly the value of targeted observations is still overstated - points that I made last time. I have performed observing system experiments with and without dropsondes (mostly dense US dropsondes in the vicinity of tropical storms; which has similarities to the Typhoon targeting of section 5.3.3). I found very noisy results. Sometimes the extra profiles improve forecasts, sometimes they degrade them.*
*Thus I would place little weight on a single case as shown in section 5.3.3, and even a one-month experiment (mentioned in section 5.2) is too short to give robust results (a point also made by another reviewer).*

**Authors: -** First of all, we sincerely thank the reviewer for this critical comment and for sharing their valuable firsthand experience with dropsondes experiments. We agree that the value of targeted observations can indeed be variable, and the original

manuscript may have presented the results with excessive optimism. The reviewer's point about the robustness of results from a single case and a one-month experiment is well taken.

We have thoroughly revised the manuscript, particularly Sections 5.2 and 5.3.3, to adopt a more balanced and scientific tone. We now explicitly state the limitations of the present study, acknowledging that the one-month experiment is a preliminary investigation and that the single typhoon case study is intended for illustrative purposes rather than serving as conclusive evidence. The text now emphasizes that these results, while promising, require validation over longer periods and across more cases to be considered robust.

As the reviewer rightly pointed out based on their own experiments, the impact of additional profiles can be inconsistent. We have expanded our discussion to reflect this complexity. The revised text now includes a paragraph discussing the various factors that influence the effectiveness of targeted observations, which aligns with the points the reviewer has raised. For instance, the positive impact is highly contingent upon factors such as:

The accuracy and rationale behind the identification of sensitive areas for targeted observation.

The quality and represent activeness of the data obtained within these areas.

The effective assimilation of this additional data by the NWP system's algorithms.

Therefore, we have been reframed our conclusions to highlight that the ADDRS system provides a potential tool for enhancing data coverage over data-sparse regions like the ocean. Its operational value lies in this potential, but its actual impact on forecast improvement in any specific event is not guaranteed and depends on a confluence of optimal conditions. Our study should be seen as a step towards understanding how to best utilize such targeted observations, rather than a definitive affirmation of their unconditional success.

We believe these significant revisions have tempered the overstated claims and provided a more objective assessment of our findings, directly addressing the reviewer's core concern. We are grateful for this insight, which has undoubtedly improved the quality of our manuscript.

*3)authors haven't really responded to my questions about the drift phase: the main considerations determining how long it is. There is also the question of whether 'drift' data are good enough to be used in NWP, it seems not for temperature (expected) but*

*to my surprise the wind results (table 5) also appear worse than the ascent/descent winds - it would be very useful to understand this better.*

**Authors: -** We sincerely thank the reviewer for this follow-up comment and for raising these critical points regarding the drift phase. We apologize that our previous response was not sufficiently detailed. We have now provided a comprehensive discussion in the revised manuscript to address these questions, as detailed below.

1. Response to the question: "the main considerations determining how long [the drift phase] is"

1. Main Considerations Determining the Duration of the Drift Phase

The duration of the drift phase is not arbitrary but is the result of a careful balance between scientific objectives and operational constraints. The primary considerations include :

Maximizing Data Coverage vs. Timeliness:A key scientific goal is to sample the data-sparse stratosphere and upper troposphere over as large a horizontal area as possible, particularly over oceans. A longer drift phase (e.g., the designed 4.5 hours in our study) significantly enhances this spatial coverage, providing valuable data on horizontal structures. However, this must be balanced against the timeliness of the data for assimilation into rapid-refresh NWP systems like CMA-MESO. If the drift phase is too long, the data may become obsolete for the intended assimilation cycle.

Synergy with Operational Sounding Schedules:As the reviewer may be aware, conventional GTS balloon soundings are typically conducted at 00 and 12 UTC. Our strategy with ADDRS is to complement this schedule. By releasing a balloon at 23:15 UTC and utilizing a prolonged drift phase, the instrument captures the atmospheric state during the early morning hours (around 04:45-05:15 UTC), culminating in a descent phase profile that is highly timely for initializing the 06 UTC forecast cycle. This "targeted observation" approach aims to fill a specific data gap.

Technical and Safety Limitations:The practical upper limit of the drift phase is constrained by the battery life of the radiosonde and the balloon's buoyancy. The chosen duration ensures complete data transmission throughout the flight. Furthermore, air traffic safety regulations necessitate that the total flight path remains within predictable and manageable boundaries.

2. Suitability of 'Drift' Data for Use in NWP

We agree with the reviewer that this is a crucial point. The seemingly lower quality of wind data during the drift phase, as noted in Table 5, can be attributed to several factors that we have now elaborated upon in the discussion section :

Increased representativeness error: During the ascent and descent phases, the sonde moves primarily vertically, providing a profile that is relatively coherent in the horizontal dimension. In contrast, during the drift phase, the radiosonde is carried significant distances horizontally by the wind. A single "drift profile" is, therefore, a slantwise path through the atmosphere. This introduces a larger representativeness error when the data is compared to a model grid point, which assumes a vertical column. The model's interpolation schemes may not perfectly capture the complex atmospheric structures sampled along this path, leading to larger differences (innovations) and thus statistically worse scores for winds, which are highly sensitive to horizontal displacement.

Sensor Response and Balloon Dynamics:Under the relatively calm conditions of the drift phase, the radiosonde's motion can be more susceptible to fine-scale turbulence and the pendulum motion of the radiosonde suspended under the balloon. This can potentially add high-frequency noise to the wind measurements that is not representative of the synoptic-scale flow the model resolves.

Challenges in Data Assimilation: Standard data assimilation systems, including the one used in this study, are primarily optimized for conventional, quasi-vertical profiles. The assimilation of observations with significant horizontal drift presents a more complex challenge for the observation operator, which typically assumes the observation is located at a single horizontal point. This mismatch can contribute to the increased errors seen in Table 5.

Therefore, we interpret the results not as an indication that the drift data are of "low quality," but rather that its effective assimilation poses different and more complex challenges than ascent/descent data. It highlights an area where future improvements in observation operators (e.g., those that can handle trajectory-based data) could yield significant benefits. Despite these challenges, we believe the drift data provides unique and valuable information on the spatial structure of the atmosphere, particularly over data-sparse regions.

We are grateful to the reviewer for encouraging us to clarify these important aspects, which has significantly strengthened the discussion in our manuscript.

*4)In the response the paragraph starting"The GTH3 radiosonde uses the HC103M2 capacitive thin-film polymer humidity sensor ..." and Table 3 are very useful. They should be included in the main text. Please provide details of the temperature and pressure sensors too.*

**Authors: -** We sincerely thank the reviewer for this positive feedback and the valuable suggestions. We agree that including these technical details in the main text will provide better context for the readers.

As recommended, to ensure consistency and completeness, we have now added the

description of the comparison of CMA's GTS12 and GTH3 radiosonde PTU in L226-236, Page 8

The GTS12 and GTH3 radiosonde use the same PTU sensor, P(MEMS piezo-resistive), T(adopting self-developed NTC bead thermistor (reducing from diameter ⩽1mm to ⩽0.4mm) volume reduction, higher sensitivity), U(capacitive thin-film polymer humidity sensor). The difference lies in that 'Instrumentation and Methods' :

For P [Design the temperature compensation circuit and establish the correction equation by fitting the sensor characteristic surface through multiple temperature segments and multiple pressure points to achieve higher measurement accuracy.]

For T [Optimize the installation structure, included Angle and lead length through simulation analysis, and effectively reduce the influence of solar radiation by improving the installation structure and coating process.]

For U [Optimize the Temperature compensation Correction Algorithm and humidity hysteresis Correction Algorithm for Humidity sensors]

We believe that incorporating these comprehensive sensor specifications directly into the main text greatly enhances the technical rigor and reproducibility of our work. We are grateful to the reviewer for this insightful suggestion.

***Detailed comments***

==========

*1)19 "The balloon-borne radiosonde observations constitutes" - "Balloon-borne radiosonde observations constitute" (delete "The", remove final "s" - better English)*

**Authors: -** We thank the reviewer for this careful correction. We agree that the original phrasing was grammatically incorrect. As suggested, we have deleted "The" and changed "constitutes" to "constitute" in the sentence. The revised text now reads: "Balloon-borne radiosonde observations constitute...". This change has been made on [Page 1, L19] of the revised manuscript.

*2)20 "uppe-air" spelling*

**Authors: -** We thank the reviewer for their careful reading and for identifying this typographical error. The term has been corrected to "upper-air" on [Page 1, L20] of the revised manuscript. We have also taken this opportunity to perform an additional spell-check throughout the manuscript to ensure consistency in terminology.

*3)40 "within comprehensive meteorological measurement framework (Ingleby and Coauthors., 2016)." - "within a comprehensive meteorological measurement framework (Ingleby et al., 2016)."*

**Authors: -** We thank the reviewer for this careful correction. We have revised the sentence as suggested. The text now reads: "... within a comprehensive meteorological measurement framework (Ingleby et al., 2016)." This change has been made on [Page 2, L34] of the revised manuscript.

*4)57/58 "Currently, the radiosonde observations are still very useful for validation/verification satellite observations (Bauer et al.,2015; Bormann et al.,2019; WMO et al.,2024)." The authors have taken three references that I suggested, but not made the main point that I was making. I suggest something more like this: "In recent decades satellite soundings have played an increasing role in Numerical Weather Prediction (NWP) and they are now the observations contributing most to forecast skill (Bauer et al.,2015; Bormann et al.,2019; WMO, 2024). Radiosonde profiles still provide an important contribution to forecast skill and help with the calibration/validation of satellite soundings and the verification of the resulting forecasts." (If the authors do not wish to "own" this statement it should be included as a comment from A reviewer.) (More on satellite cal/val if required can be found in Newman, S., Carminati, F., Lawrence, H., (...), Salonen, K., Bell, W. 2020 Assessment of new satellite missions within the framework of numerical weather prediction Remote Sensing 12(10), 1580 Radio occultation is now more important than radiosonde profiles as a reference for satellite soundings, see Bauer, P., G. Radnóti, S. Healy, and C. Cardinali, 2014: GNSS Radio Occultation Constellation Observing System Experiments. Mon. Wea. Rev., 142, 555–572, https://doi.org/10.1175/MWR-D-13-00130.1.)*

**Authors: -** We sincerely thank the reviewer for this critical guidance and for providing the specific phrasing and key references. We appreciate the reviewer's effort in helping us more accurately frame the evolving roles of satellite and radiosonde observations in NWP. We fully agree that the suggested statement provides a more authoritative and balanced perspective on the current observational landscape.

Following the reviewer's explicit suggestion, we have replaced the original sentence in the main text and incorporated the additional recommended references to strengthen our discussion. The revised text now reads:

"In recent decades, satellite soundings have played an increasing role in Numerical Weather Prediction (NWP) and are now the observations contributing most to forecast skill (Bauer et al., 2015; Bormann et al., 2019; WMO, 2024). While radio occultation data has gained importance as a reference standard (Bauer et al., 2014), radiosonde profiles still provide a crucial contribution to forecast skill and are fundamental for the calibration and validation of satellite soundings (Newman et al., 2020)." on [Page 2, L51-L55] of the revised manuscript. We believe this revision significantly

strengthens the scholarly foundation of our introduction and accurately sets the stage for our study on the value of radiosonde observations.

We are once again grateful to the reviewer for these insightful suggestions, which have undoubtedly improved the quality of our manuscript.

*7)62 "Numerous studies" - "Several studies"*

**Authors: -** We sincerely thank the reviewer for pointing this out. We agree that the term "numerous" can be subjective and "several" provides a more accurate and conservative description of the existing literature on [Page 2, L59] of the revised manuscript. Furthermore, we have taken this opportunity to review the entire manuscript and replaced L64 "numerous institutions" with "several institutions" with more precise terms where appropriate to enhance the overall objectivity of our writing.

*8)62/63 "significantly enhance" - "enhance" (the improvements are fairly minor)*

**Authors: -** We thank the reviewer for this precise comment. We agree that the improvements observed are indeed modest, and using "enhance" without the adverb "significantly" provides a more accurate and conservative description of our results. We have therefore made the change from "significantly enhance" to "enhance" on [Page 2, L60] of the revised manuscript as suggested.

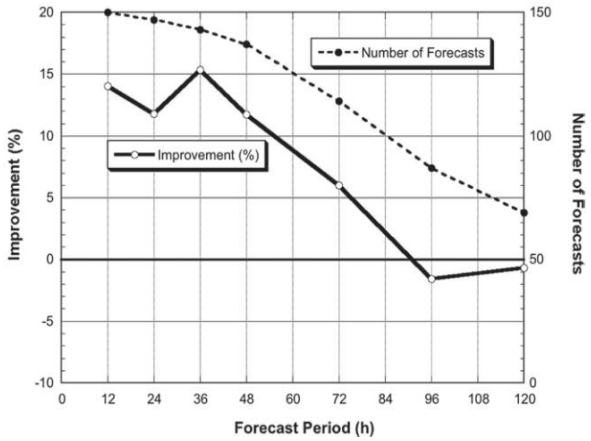
*9)63 "reported about" - "reported"*

*65 "airborne radiosonde" - "dropsonde"*

*65/66 "leading to a 10–15% reduction in errors associated with hurricane trajectory predictions" Overstated.*

**Authors: -**

We thank the reviewer for this precise correction. We agree that "dropsonde" is the more accurate term for the instrument deployed from an aircraft to measure atmospheric parameters during its descent. We also agree that the phrase "leading to a 10-15% reduction in errors" may have been an over generalization in the specific context it was used. (e.g., Wang et al., 2015 Fig.1). However, the actual improvement may vary depending on the observation network, model configuration, and synoptic conditions.




We have replaced the original phrasing with more measured language. Instead of the definitive "leading to a reduction," we now use formulations "associated with an approximate reduction of". This shift in tone more accurately reflects the probabilistic nature of model performance evaluation. And the original sentence (L63-67) was rewritten without using "reported". The changes are as follows: "Dropsonde data assimilation has been shown to hurricane trajectory forecast errors associated with an approximate reduction of 10-15% in specific contexts, such as short-term (0-48h) predictions over ocean basins using the GFS model, as demonstrated in multi-year statistical analyses( (Ingleby, 2021; Cohn et al., 2013; Wang et al., 2015)"This change has been made on [Page 2, L60-L63] of the revised manuscript.

*10)80,81 "Nonetheless, due to high costs and limited ascent rates, they are not viable solutions for long-term operational balloon sounding data collection. Furthermore, these balloons very slow ascent rates " The WindBorne balloons are now quasi-operational (details). The issue of slow ascent rate (mentioned twice, better rewritten) is a problem for temperature and humidity, but not for wind - upper level wind is very important for NWP.*

**Authors: -** We thank the reviewer again for their insightful comments. In the revised manuscript, we have considered:

1.Updated the description regarding the feasibility of WindBorne systems, acknowledging their quasi-operational progress.

2.Clearly differentiated the impact of ascent rate on temperature/humidity observations versus wind observations, emphasizing the importance of wind data for NWP.

3.Refine the text to eliminate repetition, resulting in a more fluent and rigorous discussion.

The part of L72-L78 "The WindBorne emerged in 2019 from the cradle of the

Stanford Space Program. Once launched, the WindBorne balloon regulates its altitude by jettisoning sand ballast and releasing gas, enabling multiple round-trip vertical sounding flights from below 20 km down to near the surface, with an average flight duration of seven days"and L80-L84 "Nonetheless, due to high costs and limited ascent rates, they are not viable solutions for long-term operational balloon sounding data collection. Furthermore, these balloons very slow ascent rates"is rewritten, and has been changed to "While early stratospheric balloon systems faced questions about their long-term feasibility for operational data collection due to cost and ascent rate limitations (WMO-No. 8., 2025). For temperature and humidity, slower ascent can reduce the accuracy of vertical resolution due to issues like sensor response time and radiative errors. However, for wind vectors, especially crucial upper-level winds, GPS-based positioning provides high-precision data less affected by the ascent rate. The development of a new generation of commercial balloon systems has been changing this landscape. " on [Page 2, L72-77] of the revised manuscript.

These revisions will enhance the technical accuracy of the paper and more fully reflect the potential and application value of modern balloon sounding technology.

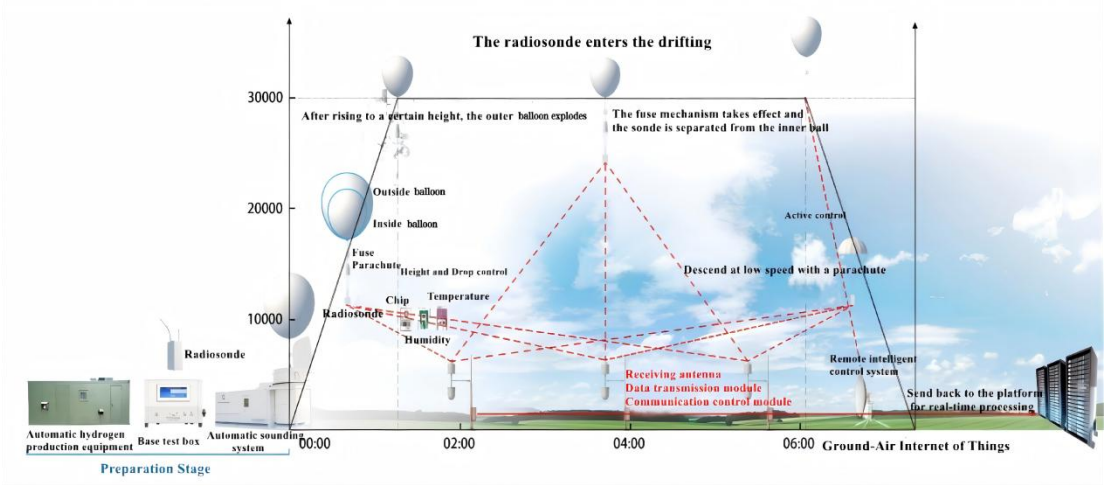
*11)88 "This paper introduce" - "This paper introduces"*
**Authors: -** "This paper introduces" on [Page 3, L82] of the revised manuscript.

*12)91 "within" - "with"*
**Authors: -** "with" on [Page 3, L85] of the revised manuscript.

*13)93 "The ADDRS was organizational developed", delete "organizational"*
**Authors: -** We agree with the reviewers' opinions that the word "organizational" is redundant in this context, and it has been deleted from [Page 3, L87] of the revised manuscript.

*14)94 "domestic units" - "organizations"*
**Authors: -** The expression "domestic units" may be ambiguous and not formal enough. We have changed to using "organizations" on [Page 3, L88] of the revised manuscript

*15)Figure 1. It would be more consistent to refer to "balloon" rather than "ball"*
*349 "outer ball" - "outer balloon"*
**Authors: -** This is a translation mistake, we carefully examined the entire text, redraw Figure 1 on [Page 3, L94] and [Page 13, L358] of the revised manuscript as follows:




*16)Figure 2. '"Ascent-drift-descent" ("add" )subsystem' - '"Ascent-drift-descent" ("ADD") subsystem'*

**Authors: -** Redraw figure 2 on [Page 4, L97] of the revised manuscript.

*17)114 "controlled aeration" - I'm not sure what this means (delete or reword)*

**Authors: -** We thank the reviewer for this comment. As suggested, we have deleted the term "controlled aeration" from the manuscript to avoid potential ambiguity. The relevant sentence has been rephrased as "[The inner balloon enable it to resist bursting. ]". This change can be found on [Page 4, L108] of the revised manuscript.

*18)118 "predefined duration" as I said in my first review I think there should be some discussion (here or elsewhere) about the considerations used to decide the duration. 226-227 "requiring a minimum operational time of six hours." Why does it require at least six hours? Related to my question on line 118.*

**Authors: -** Similar to response to the question: "the main considerations determining how long [the drift phase] is"

1. Main Considerations Determining the Duration of the Drift Phase

The duration of the drift phase is not arbitrary but is the result of a careful balance between scientific objectives and operational constraints. The primary considerations include:

Maximizing Data Coverage vs. Timeliness: A key scientific goal is to sample the data-sparse stratosphere and upper troposphere over as large a horizontal area as possible, particularly over oceans. A longer drift phase (e.g., the designed 4.5 hours in our study) significantly enhances this spatial coverage, providing valuable data on horizontal structures. However, this must be balanced against the timeliness of the

data for assimilation into rapid-refresh NWP systems like CMA-MESO. If the drift phase is too long, the data may become obsolete for the intended assimilation cycle.

Synergy with Operational Sounding Schedules: As the reviewer may be aware, conventional GTS balloon soundings are typically conducted at 00 and 12 UTC. Our strategy with ADDRS is to complement this schedule. By releasing a balloon at 23:15 UTC and utilizing a prolonged drift phase, the instrument captures the atmospheric state during the early morning hours (around 04:45-05:15 UTC), culminating in a descent phase profile that is highly timely for initializing the 06 UTC forecast cycle. This "targeted observation" approach aims to fill a specific data gap.

Technical and Safety Limitations:The practical upper limit of the drift phase is constrained by the battery life of the radiosonde and the balloon's buoyancy. The chosen duration ensures complete data transmission throughout the flight. Furthermore, air traffic safety regulations necessitate that the total flight path remains within predictable and manageable boundaries.

*19)120/121 "Subsequent to this separation .... mission segment." I suggest omitting this sentence, once it has separated the balloon has to come down sooner or later. Calling this a mission segment puzzled me (did this refer to the descent phase? - no).*
**Authors: -** We thank the reviewer for this suggestion. We agree that the sentence was redundant and its description of the inner balloon's separation as a "mission segment" was potentially misleading. We should focus on the radiosonde and its carrier. Accordingly, we have omitted the sentence as suggested. This change can be found on [Page 5, L109] of the revised manuscript.

*20)159-161 "Compared ... observations." I was not familiar with the work of Anand/Vernier and others, I would say that they are addressing a different issue. I see that you have already discussed their work On lines 77-81 and to me the comparison here is unhelpful. I suggest omitting this sentence.*

*161-165 The rest of this paragraph seems to repeat what is said elsewhere and could also be omitted?*
**Authors: -** We thank the reviewer for this insightful comment. We agree that the comparison in this sentence may distract from the main point and is not essential to the argument. Therefore, we have omitted the L159-165 sentence as suggested by the reviewer. The removal of this sentence helps to streamline the discussion and improve the focus of the paragraph. The corresponding change can be found on [Page 5, L151] of the revised manuscript.
*21)171 'the aeration volume' - 'the volume'*
*179 "precise control of aeration" - "precise control of density"? Or "precise control of the mass of air in the balloons"? I am not familiar with "aeration" in this context and*

*would suggest avoiding use of a novel word unless essential. "Aeration" could be omitted from lines 181 and 184 (or replaced by "optimum" on line 184). Please include an indication of how much hydrogen is used in each balloon in the optimal configuration.*

*185-186 "controls the aeration of dual-mode balloon" - "controls the performance of dual-mode balloons"*

**Authors:** We sincerely thank the reviewer for these valuable suggestions to improve the clarity of our manuscript. We agree that using "the optimum volume" is sufficient and clearer, and on [Page 6, L169] of the revised manuscript. Regarding the term "aeration", we agree that it could be ambiguous in this context. We have followed the reviewer's suggestion and replaced "aeration" to "inflation."with the more descriptive phrase "precise control of the mass of air in the balloons" throughout on [Page 6, L164-165] of the revised manuscript to avoid confusion. Furthermore, as helpfully suggested, we have added the specific amount of hydrogen used in each balloon under the optimal configuration.

We add our ADDRS team's research. "And based on the sounding balloon observation data in Guangdong from August 2022 to October 2023, the optimum volume capacities of related factors of the inner and outer balloons of the dual-mode sounding balloon on the success rate of the drift were studied, and a reasonable inflation scheme was established to improve the success rate of the drift. The research shows that the theoretical inflation amount of the inner balloon for drifting at the expected height is 55.6 mol. Affected by day and night conditions, the inflation amount during the day should be controlled at 52.6 ±2mol, and at night, it should be controlled at 57.6±2 mol. The burst height of the outer balloon has a significant impact on the success rate of the drift. When the burst height of the outer balloon is within the expected height, the success rate of the drift can reach 82%. The success rate of the drift during the day is higher than that at night, and the success rate under clear sky conditions is higher than that under cloudy and rainy conditions. Under rainy conditions, the success rate of the drift is only 50.2%. After verification, the success rate of the drift can reach 93.5% by adopting the reasonable inflation scheme (Xu et al., 2025)"on [Page 6, L168-179] of the revised manuscript.

*22)235 "reduce the size, weight, and power consumption." Perhaps add "(The amount of plastic, and toxic materials in the circuit boards, has also been reduced.)" From response to first review.*

**Authors: -** We thank the reviewer for this valuable suggestion. We agree that highlighting the reduction in the usage of plastic and toxic materials further strengthens the environmental benefits of our design. Accordingly, we have added the suggested sentence at the end of the relevant paragraph on [Page 8, L223-L225]. The revised text now reads: "...significantly reducing the overall size, weight, and power consumption of the device. Furthermore, the amount of plastic and toxic materials in the circuit boards has also been substantially reduced (Table 3)".

*23)Table 4. In the viewer I am using (preview) most of the numbers just appear as dots/circles.*

**Authors: -** We sincerely thank the reviewer for pointing out this display issue in Table 4 when using the Preview application. Upon investigation, we found that it might be caused by specific font encoding or rendering compatibility. To ensure the table is displayed correctly across all platforms, we have regenerated the PDF file of the manuscript, paying special attention to embedding all used fonts and using standard character sets. The revised PDF has been uploaded. We kindly request the reviewer to check if the display is normal in the new version.

*24)257 "the gravity or altitude" - "the altitude"?*

**Authors: -** We thank the reviewer for this comment. We agree that the term "altitude" is more precise and direct in this context. As the reviewer suggested, we have revised the phrase from "gravity or altitude" to "altitude" in the manuscript. This change is because the primary environmental factor that the drifting controller's ballast-dropping mechanism actively responds to is the changes in air density and pressure associated with altitude, rather than the gravitational force itself, which remains relatively constant within the operational range. The revised text now reads: "Similar to the Windborne system, it is equipped with a ballast-dropping mechanism to control its weight, allowing the inner balloon to better adapt to changes in altitude." This modification can be found on [Page 10, L255-L256] of the revised manuscript.

*25)285 "the Guide to Operational Upper-Air Meteorological Observation" Does this mean WMO No. 8 (or a CMA document) - please clarify.*

**Authors: -** We thank the reviewer for this question. The document "Guide to Operational Upper-Air Meteorological Observation" is indeed a technical specification issued by the China Meteorological Administration (CMA), which guides the operational practices of high-altitude meteorological observation within China. It is developed with reference to international standards, including the WMO-No. 8 "Guide to Instruments and Methods of Observation", but is tailored to the specific operational context and technical requirements of China's meteorological observation network. To avoid any confusion, we have clarified this point in the revised manuscript by explicitly stating it as a "CMA technical specification" at its first mention (Page 10, L284).

*26)291 "The effect of radiation [in the drift phase] ..." Is it not partly that there is very little ventilation of the sensors in this phase?*

**Authors: -** We thank the reviewer for raising this excellent point. We agree that the limited ventilation during the drift phase could indeed be a contributing factor to the observed sensor effects. The lack of convective cooling in a low-ventilation environment might lead to localized temperature increases or alter the local

atmosphere around the sensors, potentially exacerbating the radiation. Please see author response to comment 28), which refines the simulation-based error correction equations (Yang et al., 2022). There are some innovative methods for temperature observation in the ascending segment of ADDRS. The sources of uncertainty in the drift phase and the descent section are similar. However, the quantification of the temperature measurement uncertainty caused by the CFD model and the long-term accumulation of low ventilation coefficient solar radiation still requires more research. This is a very interesting research direction. Thank you very much for your insights.

*27)294 "Vaisala RS41 radiosonde software often flags such data as invalid" What do you mean by "such data"? Vaisala RS41 does not usually have a "flat drift phase". Either explain better or delete the sentence.*

**Authors: -** We thank the reviewer for this insightful comment. Upon careful consideration, we agree that the statement was ambiguous and could be technically misleading regarding the characteristics of the Vaisala RS41. Therefore, we have followed the reviewer's suggestion and deleted the sentence. This change helps to avoid potential confusion and ensures the accuracy of our description. The removal of the sentence on (Page 11 , L291) has been made in the revised manuscript.

*28)306 " refine the simulation-based error correction equations (Yang et al., 2022)." It would be useful to give an indication of the magnitude of the errors involved.*

**Authors: -** We thank the reviewer for this constructive suggestion. We have refined the discussion of the simulation-based error correction equations and provided explicit indications of the error magnitudes involved, as detailed below.

1. Refinement of Simulation-Based Error Correction Equations

Our study established a framework for error correction based on Computational Fluid Dynamics (CFD) simulations. The core thermal balance equation is described as:

Q_radiation + Q_convection = Q_sensor

where Q_radiation represents solar radiation heat, Q_convection represents convective heat transfer, and Q_sensor represents the heat absorbed by the sensor.

Parametric analysis quantified the impact of key sensor design parameters on measurement error:

Sensor Volume vs. Radiation Error:The simulation showed that at 40 km altitude, the measurement error difference between sensors of different volumes (0.6-1.6 $mm^3$) could be as high as 1.5K. Smaller volumes resulted in smaller radiation errors.

Lead Wire Angle Impact:As the lead wire angle increased from 0° to 180°, the solar radiation error increased from 1.009 K to 1.398 K, a difference of 0.389 K.

2. Indication of Error Magnitude

The major error sources and their magnitudes identified and quantified in this study are summarized below:

| Error Type | Magnitude Range | Main Influencing Factors |
|---|---|---|
| Solar Radiation Error | 1.0 - 3.0K | Altitude, Solar Altitude Angle, Sensor Coating Reflectivity |
| Wet-Bulb Effect Error | 0.5 - 2.0K | Cloud Penetration Process, Humidity Change, Evaporative Cooling |
| Convective Heat Transfer Error | 0.2 - 1.0K | Ascent Speed, Air Density, Sensor Size |

Wet-Bulb Effect: The cooling effect due to water droplet evaporation during cloud penetration causes significant underestimation of temperature.

Ascent Speed Influence: Variations in ascent speed significantly affect convective heat transfer. Data binning (in 2 m/s units) was used to handle this variability effectively.

3. Verification of Error Correction Effectiveness

The proposed multi-model fusion framework (SVR, XGBoost, DNN, LR) demonstrated significant performance improvement. The prediction results of different models are compared below:

| Model | Average Error (K) | Mean Squared Error (K) | Standard Deviation (K) | $R^2$ |
|---|---|---|---|---|
| Before Correction | 0.817 | 0.878 | 0.458 | - |
| LR | 0.074 | 0.573 | 0.364 | 0.80 |
| SVR | 0.126 | 0.315 | 0.351 | 0.81 |
| DNN | 0.104 | 0.215 | 0.336 | 0.83 |
| XGBoost | 0.115 | 0.178 | 0.229 | 0.85 |
| Fusion Model | 0.008 | 0.068 | 0.204 | 0.93 |

The fusion model's predictions most closely matched the true values, as visually confirmed in the comparison figure.

We believe these revisions and clarifications regarding the error magnitudes and correction effectiveness significantly enhance the manuscript's clarity and practical

value. We thank the reviewer for the insightful comments that helped us improve our work.

*29)307 "Vertical wind data extracted by parachute landing" - "Vertical wind data during the parachute descent"*

**Authors: -** We thank the reviewer for this precise comment. We agree that "parachute descent" more accurately describes the phase during which the vertical wind data were collected, as it encompasses the entire downward trajectory rather than just the final landing moment. We have revised the text on [Page 11, L303], which now reads: "Vertical wind data during the parachute descent".

*30)316 "Guo et al, 2018" ("Evaluation of wind performance of domestic Beidou dropsonde of ball-loading") It would be very useful to have a summary of the results of Guo et al in English.*

**Authors: -** Thank you for this suggestion. We agree that a summary of the key findings from Guo et al. (2018) in English will provide valuable context.

"The balloon-launched dropsonde system, utilizing the domestic BD sonde, shows significant promise. The BD sonde demonstrated acceptable accuracy during both ascent and descent phases compared to the RS92 standard. The key advantages identified were the ability to temporally densify observations (effectively obtaining an "ascent" and "descent" profile from one launch) and the potential to spatially expand coverage through the drift of the parachute, allowing for a "one-station launch, multi-station reception" model. This technology aligns with the trend towards automated, quantitative remote sensing in meteorological observation."

We hope this summary provides a clear overview of the relevant results from Guo et al. (2018).

*31)317 "selection of conical parachutes" - please include some of the details of the large conical parachutes mentioned in the 'response' but not in the main text. Please also give the string length used.*

**Authors: -** Thank you for this suggestion. We agree to include some of the details of the large conical parachutes mentioned.

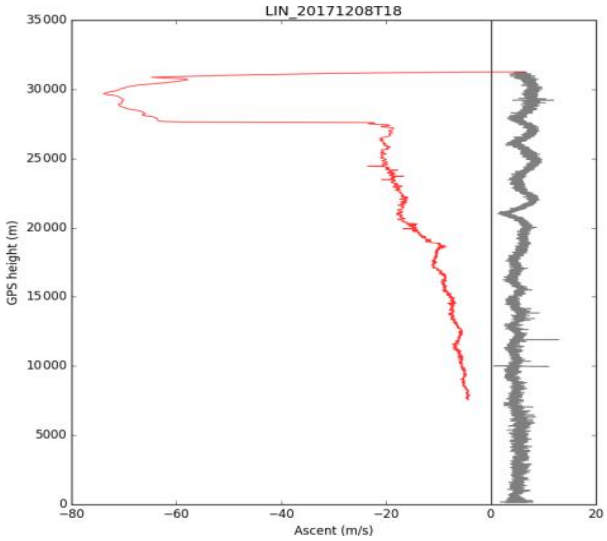
"The string length between the balloon/parachute and the GTH3 radiosonde is standardized at 30 meters in manual launches. This length helps reduce sensor exposure to balloon wake effects but amplifies pendulum motion in wind data. We carried out more than ten comparative tests for the selection of conical parachutes. The oscillation angle of conical parachute is less than 1°, which can ensure the rationality of the Gaussian filtering correction filter window for the horizontal wind in

the descent section. Meanwhile, a larger main parachute size can reduce the descent speed. Therefore, without considering the cost, a ‘Large conical parachute' is recommended for the descent section detection of ADDRS. And make sure the parachute area is tailored to match the radiosonde weight. The descent phase of ADDRS could achieve a low-level descent speed of 6 m/s±1 m/s can be achieved in UTLS (upper troposphere and lower stratosphere (7-17km), and a swing angle below 5°, as well as a vertical wind measurement uncertainty of less than 1 m/s. These findings demonstrate that the model is effective for calculating vertical wind."This modification can be found on [Page 11, L319-L330] of the revised manuscript.

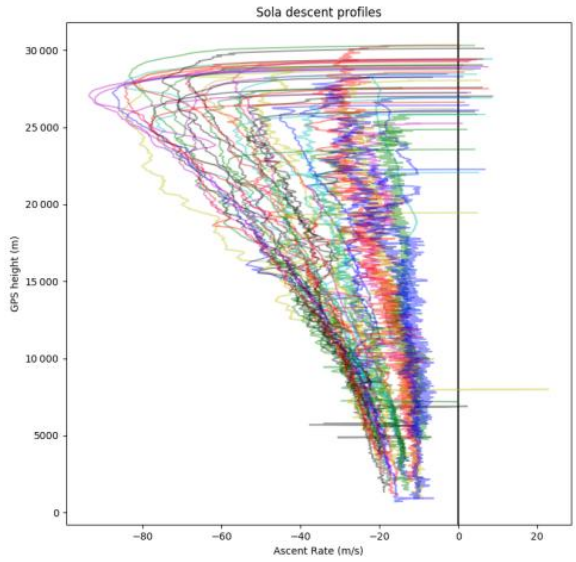
*32)318 "a stable descent speed of approximately 6 m/s" From Figure 11 the initial descent speed is much higher (as seen by Ingleby et al (2022)) this should be mentioned here. Replace 'stable' with 'low-level'? Do you see the large case-to-case variation in descent speed reported by Ingleby et al (2022)?*

**Authors: -** Thank you for this suggestion. The term "stable"has been replaced with "low-level" in L327 to accurately reflect that the ~6 m/s descent rate typically occurs only in the lower troposphere after deceleration. As correctly noted, initial descent speeds are often much higher (frequently exceeding 50 m/s) immediately after balloon burst, especially in the stratosphere.

We explicitly acknowledge the large case-to-case variability in descent rates, as reported by Ingleby et al. (2022). For instance, Figure R4 illustrates a rapid descent (>70 m/s) followed by abrupt deceleration when the parachute deploys:




Similarly, Figure 5 highlights significant variability in descent profiles from Sola, Norway, where no parachutes were used:




This variability is attributed to factors such as parachute deployment effectiveness, entanglement with balloon remnants, and the mass of attached debris (Section 2.4). The revised text now emphasizes that descent speeds are highly unstable initially and stabilize only at lower altitudes.

We can refer to Figure 6 in Ingleby et al. (2022), which shows mean descent rates by country, further underscoring the variability influenced by parachute use and balloon characteristics.

Thank you again for these insightful comments, which have strengthened our manuscript. We are happy to incorporate further suggestions if needed.

*33)325,326 "The ADDRS network spans six stations" - "The ADDRS radiosondes were launched from six stations" (I think this is what you mean.)*
**Authors: -** We agree with the reviewer that the original phrasing could be misinterpreted. The suggested revision more accurately describes the operational context, as the radiosondes were indeed launched from six specific stations rather than implying a continuous spatial coverage. We have modified the text to: "The ADDRS radiosondes were launched from six stations."on Page 12, L335-336 of the revised manuscript.

*34)347,348 "and the effect of daytime was better than nighttime" - "somewhat worse than the daytime performance"*
**Authors: -** We agree with the reviewer that the original phrasing could be interpreted as overly absolute. The suggested revision, "somewhat worse than the daytime performance"provides a more nuanced and scientifically appropriate comparison between nighttime and daytime results. We have adopted this wording on Page 13, L357-L358 of revised manuscript.

*35)351,352 "the drift height should basically meet the 30hPa requirement of GBON (WMO, 2020)." The GBON requirement for some ascents to reach 10hPa should also be mentioned.*

**Authors: -** We sincerely thank the reviewer for this insightful comment. We agree that it is important to present the complete GBON requirements. In the revised manuscript, we explicitly mention both the 30hPa and the 10hPa criteria for a subset of ascents, as stipulated by GBON (WMO, 2020). The revised text now reads:

"The drift height meets the GBON requirement of attaining at least 30hPa for a majority of ascents, with a subset of ascents also reaching the 10hPa level (WMO, 2020)"on Page 13, L3561-363 of revised manuscript.

*36)357 "The simulated and observed trajectories" - do the simulated trajectories include NWP forecasts? More information required.*

*358 "A color gradient based on pressure altitude" The colour gradient is too subtle for me to make out. Use different (larger) markers when a multiple of 200hPa is crossed?*

**Authors:-** Thank you for your valuable suggestions. The simulated trajectory refers to the forecast data generated by the Beidou-based radiosonde trajectory prediction system. We acknowledge that color gradients alone may be difficult to distinguish. Therefore, we have revised the Figure7 to use distinct and larger markers at every 200hPa interval, which is clearly indicated in the legend for better clarity. in L372 of revised manuscript.

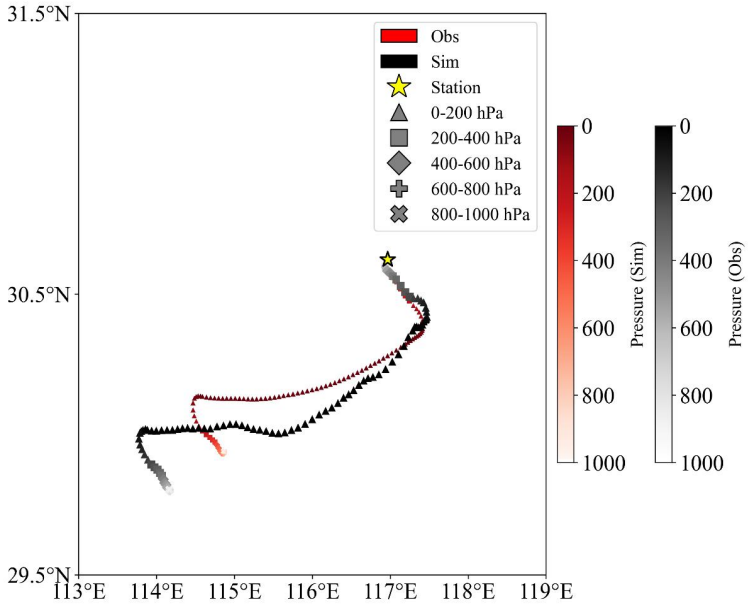



Figure 7. Schematic diagram of the observed (black) and simulated (red) trajectories. The yellow pentagram indicates the radiosonde station. The color of each point represents the

corresponding pressure level. Distinct markers are used every 200hPa to enhance visual differentiation.

*37)Table 5. This is very interesting, but raises some questions.*
*How were the RH comparisons performed? (Which SVP equation was used?) Some details in the response should be included in the main text.*

**Authors: -** We thank the reviewer for these insightful questions. Regarding the Relative Humidity (RH) comparisons and the Saturation Vapour Pressure (SVP) equation, we provide the following clarification. This ensures transparency and reproducibility of the RH analysis. The revised text reads: "But for relative humidity (RH), different radiosonde manufacturers may adopt distinct SVP equations in calibrating humidity sensors. This difference in the choice of SVP equation can lead to discrepancies in relative humidity measurements among different radiosonde types, especially under low-temperature conditions.

The ADDRS radiosonde humidity sensor is calibrated using the Goff (1957) SVP equation recommended in earlier WMO publications (WMO, 2012). Moreover, relative humidity from the ERA5 dataset is calculated using the Buck SVP equation. To ensure a more consistent comparison, relative humidity was recalculated from temperature and specific humidity in the ERA5 dataset using the Goff (1957) SVP equation, with the SVP assumed over liquid water only. The results show that the O–A bias and RMSE of relative humidity under low-temperature conditions are significantly reduced and become much more physically reasonable."on Page 15, L390-L399 of revised manuscript.

*38)For the "drift" results I understand the temperature results being worse than the ascent/descent results, but I don't understand it for wind. This should be mentioned in the text and an explanation attempted if possible. (How strong a function of pressure are the wind results in the stratosphere?)*

**Authors: -** Regarding the wind "drift" results and their pressure dependence in the stratosphere, the reviewer's observation about the wind results is astute. We agree that the degradation in wind agreement during the drift phase, unlike temperature, is less intuitive. Our analysis, supported by recent studies, suggests this is primarily due to the increasing uncertainty and potential biases in ERA5's representation of wind fields

with height in the stratosphere, coupled with the unique conditions during balloon drift. While ERA5 generally correlates well with radiosonde winds in the troposphere, its accuracy can diminish in the stratosphere due to a combination of factors: (i) sparser conventional observations for data assimilation at these levels, (ii) model-specific dynamics and damping schemes near the model top, and (iii) the challenge of representing small-scale processes and gravity waves that significantly influence stratospheric winds. During the ascent/descent phase, the balloon's trajectory is relatively vertical, and the measurements are closely tied to the model's vertical column structure. However, during the drift phase at high altitudes, the balloon travels horizontally over large distances. The model's ability to accurately represent the horizontal wind shear and fine-scale vertical structure of the jet streams in the stratosphere under these conditions is limited. This can lead to larger discrepancies when comparing a point measurement (radiosonde) from a specific location and time against the model's grid-box average (ERA5).

Pressure Dependence: The wind differences indeed show a function of pressure. Studies indicate that ERA5 often exhibits a systematic low bias in wind speed compared to radiosondes in the lower stratosphere (above ~15 km / 200hPa), which becomes more pronounced with increasing altitude. This bias is a known characteristic identified in intercomparison studies over China, where radiosonde winds tend to be stronger than ERA5 winds in the upper levels. The "drift" phase samples these higher, more uncertain regions of the ERA5 model for a prolonged period, amplifying the observed differences.

Supporting References from Search Results:

1.Banyard, T. P., Wright, C. J., Osprey, S. M., Hindley, N. P., Halloran, G., Coy, L., … Alexander, M. J. (2024). Aeolus wind lidar observations of the 2019/2020 quasi-biennial oscillation disruption with comparison to radiosondes and reanalysis. Atmospheric Chemistry and Physics, 24(4), 2465–2490. https://doi.org/10.5194/acp-24-2465-2024. Banyard, T. P., et al. (2024). Aeolus wind lidar observations of the 2019/2020 quasi-biennial oscillation disruption with comparison to radiosondes and reanalysis. Atmospheric Chemistry and Physics. This study directly compared the performance of Aeolus satellite, radiosonde and ERA5

reanalysis data in stratospheric wind fields (covering the quasi-biennial oscillation interruption events of 2019/2020). The results show that the small-scale fluctuations observed by Aeolus are significantly smoosed in the ERA5 reanalysis data, and the intensity of the high-level westerlies jet stream detected by Aeolus is higher than that of ERA5, providing a direct observational basis for explaining the differences in stratospheric wind field characteristics among different data sources.

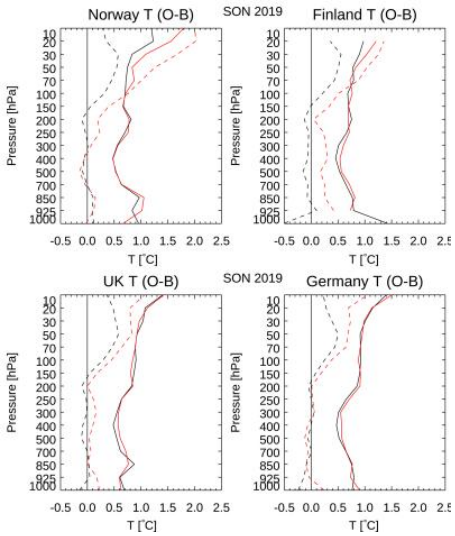
2.Liu, B., Guo, J., Gong, W., Zhang, Y., Shi, L., Ma, Y., Li, J., Guo, X., Stoffelen, A., de Leeuw, G., and Xu, X.: Intercomparison of wind observations from ESA's satellite mission Aeolus, ERA5 reanalysis and radiosonde over China, Atmos. Chem. Phys. Discuss. [preprint], https://doi.org/10.5194/acp-2021-41, 2021. Liu, B., et al. (2022). Intercomparison of wind observations from ESA's satellite mission Aeolus, ERA5 reanalysis and radiosonde over China. Atmospheric Measurement Techniques Discussions. This study focused on the Chinese region and systematically evaluated the wind field consistency of Aeolus, ERA5 and radiosonde. The research found that in high-altitude areas such as the Qinghai-Xizang Plateau, the zonal wind measured by radiosondes is generally higher than the reanalysis results of ERA5. This deviation trend is consistent with the previous conclusion that the systematic difference in wind speed between radiosondes and ERA5 in the stratosphere is greater than that in the troposphere, further supporting the influence of regional topography and atmospheric paired wind field observation consistency.

3.Future Research Direction: From the analysis of the significant differences between the upper-level sounding data and ERA5, it can be seen that the re-analysis data, including numerical forecast products, have relatively large errors above the stratosphere. The continuous observation of the drift phase data may supplement the scarcity of observations above 20 km traditionally limits validation. It can be applied to the improvement of reanalysis data and numerical forecasting products. We fully hope that the future application of high-resolution data from systems like ADDRS could further evaluate and constrain model representations in the stratosphere. The writing of this article is also in the hope that researchers in the forecasting model can pay attention to using data above the 20km stratospheric atmosphere, thereby strengthening research and interaction in this area.

The revised text reads:"We also note that the discrepancies between the measured wind and temperature data during the drift phase and the ERA5 reanalysis data are significantly larger, a phenomenon consistent with most radiosonde data comparisons at high altitudes (typically above 300hPa) against ERA5 (Liu, B et al., 2021). With the continued accumulation of continuous drift phase of ADDRS data in the future, it holds the potential to support research and improvements in reanalysis datasets like ERA5. "on Page 15, L400-L404 of revised manuscript.

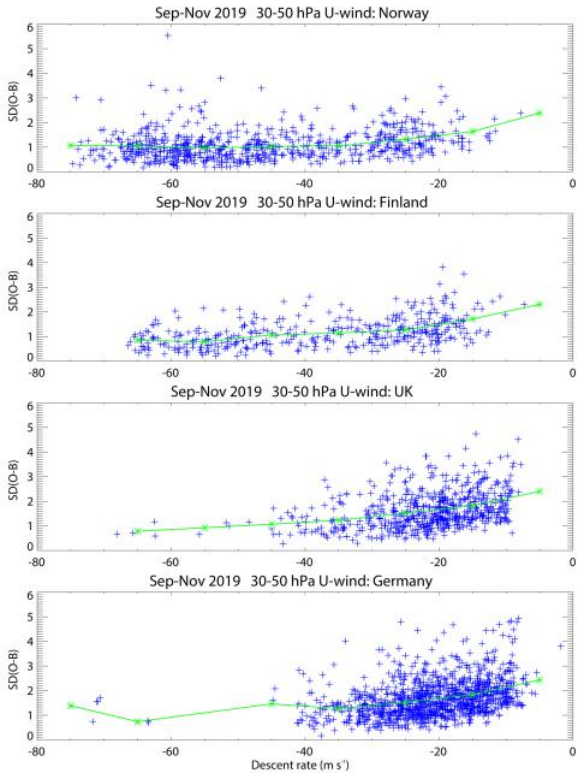
*39)For the descent temperatures there is some sign of a warm bias, especially at upper levels, which was discussed by Ingleby et al (2022) - the European biases may be larger in magnitude due to larger fall-speeds but to see this more clearly further results (as a function of height or fall-speed) would be needed. If the upper-level temperatures are biased there is a question of whether they should be assimilated or not.*

**Authors: -** We agree with the reviewer that a warm bias is evident in descent temperatures, particularly at higher altitudes, and that this bias is closely linked to descent rate. Ingleby et al. (2022) systematically investigated this issue and provided clear evidence of the relationship. For instance, as shown in their Figure 13, descent temperatures from European stations (e.g., Norway and Finland) exhibit a pronounced warm bias in the stratosphere compared to ascent data, which correlates with higher descent speeds in these regions.




Moreover, the dependence of temperature bias on fall speed is explicitly demonstrated in Ingleby et al., 2022 's Figure 14, where a quadratic relationship is observed

between descent rate and O-B temperature differences. Stations with higher fall speeds (e.g., Norwegian sites without parachutes) show larger biases, supporting the reviewer's suggestion that European biases may be magnified by faster descent.




Further quantifying this relationship, reveals a consistent trend across multiple stations and confirming that the bias is well-characterized by the equation $\Delta T = A \cdot DR^2$, where DR is the descent rate. Regarding assimilation, we fully acknowledge the concerns about using biased upper-level temperatures. Ingleby et al. 2022 noted that ECMWF's operational system excludes descent temperatures above 150hPa during assimilation due to these biases. This conservative approach ensures that only reliable data are used, while further research continues to refine bias corrections. We have emphasized this point in our manuscript to clarify our methodology.

The revised text reads:"And (Ingleby et al., 2022) noted that ECMWF's operational system excludes descent temperatures above 150hPa during assimilation due to these biases. This conservative approach ensures that only reliable data are used, while further research continues to refine bias corrections.the temperature 'descent' warm results represents a noteworthy issue for future research."on Page 15, L404-L407 of revised manuscript.

We thank the reviewer for highlighting this issue and have revised the text to more explicitly reference the evidence from Ingleby., et al. 2022. Please let us know if any additional details or adjustments would be helpful.

In a separate paper focusing on the "ascent – drift – descent" method for temperature and wind observation, we will conduct a more detailed analysis of this issue. Specifically, we will incorporate multiple influencing factors, such as the variation in ascent/descent speeds (with the ascent phase being balloon-borne and the descent phase parachute-borne), string length, and ventilation coefficients. This approach aims to develop a more robust and applicable Quality Assurance/Quality Control (QA/QC) method specifically tailored for temperature observations during the descent segment.

*40)382 "The applications in weather analysis" - "The applications to weather analysis" (more grammatical) This subsection doesn't actually show NWP analyses. "A case study" might be more accurate.*

**Authors: -** We agree with the reviewer's grammatical suggestion and we also agree that the original subsection title was somewhat misleading, as it did not present NWP analyses. We have changed "The applications in weather analysis"to "A case study of weather analysis"(Page 15, L410).

*41)394 "trajectory data" - "trajectories" (for ADDRS is it just the ascent trajectory?) What time and how far downstream was the ADDRS descent?*

**Authors:** - We thank the reviewer for pointing this out. We have replaced 'trajectory data' with 'trajectories' for clarity (Page 15, Line 415). Specifically, the ADDRS ascent and descent phase of trajectory is complemented (Fig. R1).Although the intended drift phase at the preset altitude was not successfully maintained due to strong convective conditions, the observed descent path coincidentally captured meaningful environmental information. Notably, under the influence of intense convection propagating from west to east, the ADDRS sonde also followed a west-to-east trajectory during its descent, which occurred in the upper troposphere and lower stratosphere (around 15km). In contrast, other soundings (Fig. R2). that successfully achieved stable drifting at higher stratospheric levels (around 25km) exhibited an opposite, east-to-west trajectory over a four-hour drifting period. This divergence in

drift direction between different atmospheric layers highlights the strong shear and contrasting flow regimes associated with the convective system. For certain large-scale, long-lived convective systems, it is possible that after drifting eastward in the stratosphere, they may be detected during the descent phase moving westward in the troposphere. This would allow monitoring of the same convective system with a time lag of approximately 3 – 4 hours. Although a well-documented case of such a detection pattern has not yet been reported, this novel observation approach may hold potential for development as climate extremes become more frequent in the future.

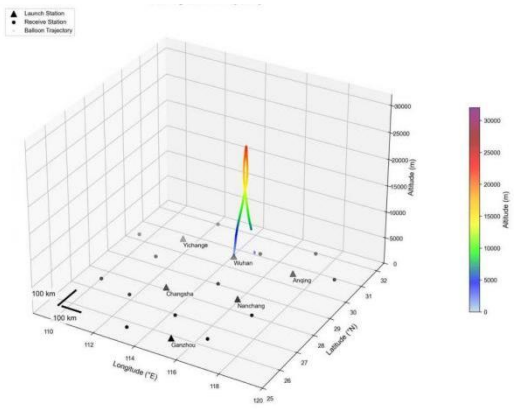



Figure R1. Trajectories of GTH3 radiosonde at the Wuhan station at 12:00 UTC on 11 July 2021.

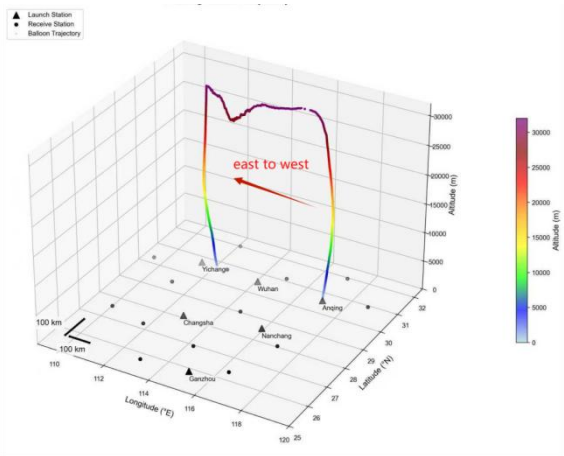



Figure R2. Trajectories of GTH3 radiosonde at the Wuhan station at 12:00 UTC on 12 July 2021.

*42)395 "At 19:15 on the night of the 8th" - is the time in UTC? Also the other times in this section.*

**Authors: -** All times mentioned in this section are UTC+8 means BTC. We have clarified this in the text (Page 16, L422-L431) to avoid any ambiguity.

*43)428 "vertical sparring methods" should "sparring methods" be "spacing"*

**Authors: -** We thank the reviewer for catching this typo. "Sparring methods"has been corrected to "spacing methods"(Page 17, L455).

*44)462 "an important method" - "a possible method"*

**Authors: -** We thank the reviewer for these suggestions on improve the precision of our language. We have revised the manuscript as follows: Changed "an important method"to "a possible method"on (Page 18, L492).

*45)463 "they significantly enhance" - "they sometimes enhance" (as it stands the text overstates the Importance of targeted observations in my opinion)*

**Authors: -** We agree that the original statement was too strong and have modified it to "they sometimes enhance"on (Page 18, L493).

*46)464 ", which is crucial for predicting extreme weather disasters" - delete (overstatement)*

**Authors: -** We thank the reviewer for this technical correction. We have deleted it (Page 18, L494) as it more accurately describes the system.

*47)467 "power system" - "steering system"?*

**Authors: -** We thank the reviewer for this technical correction. We have changed "power system" to"steering system" (Page 18, L496) as it more accurately describes the system.

*48)532,533 "Therefore, the demand for vertical profile data of the internal interface of Typhoon 'SAOLA' became imperative." Overstatement, perhaps "Therefore forecasters wanted profiles of the interior of Typhoon 'SAOLA'." Please add "Note also the importance of the steering flow and use of satellite data to analyse the wider environment of tropical storms (Magnusson et al, 2025)." If the authors do not wish to 'Own' this statement it should be included as a comment from a reviewer.*

**Authors: -** Targeted observation is an important approach to achieving more precise and efficient meteorological observation. Its effectiveness is influenced by multiple factors, including the understanding of weather evolution mechanisms, the reasonable determination of the targeted observation area, and the feasibility of the observation methods. This paper focuses on proposing an easily implementable targeted observation technology solution, namely the realization method based on the dropsonde from a drifting balloon. The analysis of Typhoon "Sara" as a case study verified the effectiveness of this method in specific circumstances. However, due to the complexity of influencing factors, it cannot be concluded that all targeted observations will achieve significant results. Therefore, in the discussion process, this paper will adhere to a scientific and rigorous attitude, striving for objective and fair expression. We agree that the original statement was too strong and have modified it to "Therefore forecasters sought profiles of the inner structure of Typhoon 'SAOLA'. Note also the importance of the steering flow and use of satellite data to analyse the wider environment of tropical storms (Magnusson et al, 2025)" (Page 21, L561-L563).

*49)540 "required air capacity" - "hydrogen" or "gas" rather than "air"*

**Authors: -** We thank the reviewer for this technical correction. We have changed "required air capacity" to"hydrogen" (Page 21, L571) as it more accurately describes the system.

*50)544 "Late test results" - "Test results"*

**Authors: -** We agree. "Late"has been deleted, and the text now reads "Test results"(Page 21, L575).

*51)561 rewrite*

**Authors: -** We agree that the original statement was too strong and have omited it (Page 22, L591)..

*52)563 "As well as ADDRS represents a next-generation approach" - "ADDRS represents a possible next-generation approach"*

**Authors: -** We thank the reviewer for these suggestions to improve the precision of our language. We have revised the manuscript as follows: Changed "As well as ADDRS represents a next-generation approach" to "ADDRS represents a possible next-generation approach"' (Page 22, L591).

*53)579 "<or ethical concerns>" - delete (or explain)*

**Authors: -** We agree. "or ethical concerns"has been deleted (Page 22, L612).

*54)Data availability:*

*"https://www.geodata.cn/collect/site/#/" When I tried this I had an error message "502 Bad Gateway"*

*I managed to connect to "https://www.geodata.cn/" - all in Chinese except for an 'English version'*

*Link which redirects to "http://wdcrre.data.ac.cn". In this website a keyword search for radiosonde gave no results.*

Authors: - Apologies, "https://www.geodata.cn/collect/site/#/" this website is no longer providing data sharing due to policy considerations. "The ADDRS research team is currently building a new data-sharing platform((https://www.r7tec.com/html/report/20040875-1.htm), which will be available for scientific sharing to registered users starting in December this year. More data acquisition requires formal requestto the corresponding author < gqyaoc@cma.gov.cn> forrestricted availability. " (Page 22, L612-615).