

We really appreciate the time and effort that the two reviewers have dedicated to revising our manuscript. The comments were constructive and have improved the quality of the manuscript tremendously. In particular, three sections have been modified based on the reviewers' comments. In Section 3, the description of the instruments highlights the technical details of the instruments, followed by a description. A better context for the rainfall detection analysis has been added in Section 4.2. As the two reviewers remarked in their observations, we describe the scientific interest in analyzing the two study cases, the doldrums and the gust front. In the following lines, we address the comments of each reviewer. Reviewers' comments are in black color and the answer is in blue.

Reviewer 1

General Comments

This manuscript describes the atmospheric and oceanic measurements taken during the 40-day ship cruise of FS Meteor in the moist tropical Atlantic as part of the BOWTIE campaign in August and September 2024. After providing a brief introduction into the historic context of the measurements and their coordination within the larger-scale ORCESTRA campaign, large parts of the manuscript are dedicated to the description of the comprehensive suite of in-situ and remote-sensing instruments measuring atmospheric and oceanic variables (section 3). The authors also present instrument comparison analyses for different measured parameters (section 4) and close with two case studies of atmospheric phenomena within the main focus of the project (section 5). All observational data described in the manuscript are well documented and publicly available via a dedicated data browser for the ORCESTRA campaign using a modern IPFS infrastructure, which eases the access to and usability of the data.

Overall, the manuscript describes a unique and comprehensive data set that can be of great relevance for both the meteorology and oceanography community. Given the large number of deployed instruments and studied processes, the rather technical and lengthy description in section 3 is justified and will be referenced by future data users. My main criticism concerns the following sections 4 and 5, where the authors should better motivate the presented analyses and provide more scientific context. I acknowledge the fact that this data description paper does not aim for deep scientific analyses, however, I believe that the authors could do a better job in characterizing the uniqueness of their data set. The instrument comparison of sampled rainfall in section 4 is certainly of help when interpreting the data, but it is unclear if these results are specific to the BOWTIE deployment and observation period/region or if the same results could have been found with existing data of these suite of instruments. Also the choice of case studies in section 5 appears arbitrary and the reader is left on their own to judge the significance of these events and what new aspects the data set could add to our current knowledge of process understanding. Nevertheless, my overall assessment of this study is positive and after addressing these general concerns as well as the specific comments listed below, I would be happy to recommend the manuscript for publication in ESSD.

We acknowledge the effort of the reviewer in providing constructive comments that have improved the quality of the paper. As the reviewer stated, the purpose of the manuscript is not to give scientific insights, but rather to show how useful the data collected during BOWTIE is for research purposes. With this line of thinking, the comment concerning Section 4 and Section 5 made us provide better context regarding the rainfall detection and the two case studies. This is translated into a revision of the literature and modifications in Sections 4 and 5. In the following lines, we respond to each reviewer comment.

Specific Comments

1. L34: Please define "small-scale". Does this scale differ between atmosphere and ocean?

Modified in lines 32-35:

“In this sense, BOWTIE sampled the high-frequency variability (seconds to hours) in processes at play in the free troposphere, the atmospheric boundary layer, and the upper ocean. This is meant to be used in studies analysing the meso- (2-2000 km) and micro-scale (< 1 km) processes in the atmosphere and the sub-mesoscale (1-10 km) and microscale (<1 km).”

2. L41-42: "It is in this context that the measurements of BOWTIE are important.": Please rephrase to be more specific here.

Modified in lines 42-45:

“One way to constrain such uncertainties is the use of observational datasets measuring the high-frequency variability of microscale processes that such models cannot capture, but such observations are scarce over oceans. This is addressed by the measurements taken during BOWTIE.”

3. L164-180: This should be a table.

The sensors are already explained in Table 2. Thus, we reduce the text and highlight which variables are measured by each sensor. See lines 175-179:

“The weather station provides measurements of wind direction and speed with a resolution of 0.1°; 0.1 m s⁻¹ (2D Ultrasonic Anemometer), respectively; air temperature with a resolution of 0.1 K (EE33); relative humidity with a resolution of 0.1%; air pressure with a resolution of 0.1 hPa (PTB220), and precipitation (SRM450, Thies, and PWD22). Shortwave (Pyranometer CM11) and longwave radiation (Pyrgeometer CG4) are measured, as well as water temperature (Pt 100 type 2020).”

4. L181-186: I suggest to remove or shorten this paragraph as it does not appear to be relevant for the average data user.

The paragraph was shortened. See lines 180-182:

“The meteorological measurement data are processed on the meteorological computer located in the ship's meteorological office. The collection, analysis, quality control, and storage of measurements are conducted on this system, which is then fed into the DSHIP on-board data network.”

5. L229: Define "lower" elevation angle

Lower elevation angle is replaced by (<90°). See lines 234-235:

“These off-zenith observations (<90°) improved the accuracy of temperature retrievals within the lowest approximately 1.5 km.”

6. L266-270: I am skeptical if this level of detail is needed here.

This was moved to the Cloudnet processing section (Section 3.3)

7. L312-313: Specify what "perpendicular" means. I would guess that it refers to the orientation of the optical sampling paths. Why would you expect the measurements to depend on the orientation of the instrument?

Yes, it refers to their optical sampling path. This has been added in the manuscript, and an explanation of this alignment is also included in lines 310-317:

“Two PARSIVEL optical disdrometers (Löffler-Mang and Joss, 2000) were installed on the observation deck with their optical sampling paths oriented perpendicular to each other. Disdrometers are labeled DSD_METEOR1 and DSD_METEOR2 in the BOWTIE dataset. The sampling path of DSD_METEOR1 has a bow-stern direction, and the sampling path of DSD_METEOR2 has a port-starboard direction. This alignment of the two disdrometers follows the recommendations of the Marine ARM GPCI Investigations of Clouds (Kollias and Bartholomew, 2016). Measurements of rainfall intensities, particle size distribution, and particle velocities can be affected by wind direction, wind speed, and ship motion (Yuter and Parker 2001). This is because of turbulence generated by the ship structure (Yuter and Parker 2001) and instrument (Friedrich et al., 2013) structure in the wind field.

8. L345: What does "scrutiny is advised" mean? Please specify or remove.

Modified. See lines 347-348

“The motion-stabilization platform is not critical for DBS wind profiles when averaged over 10 min. However, it's recommended to take into account that unrealistic values could arise.”

9. L479: What is the P2 process study region? Please define or remove.

Modified to the second process study area. The process study areas are now explicitly defined in Section 2.1.

10. L564: I guess you are implicitly saying that the irregular measurement potentially leads to biased sampling. If so, please state this explicitly.

Thanks. This has been modified. See line 619:

“This irregularity could lead to biased sampling.”

11. L568: This explanation is unclear to me.

This explanation was added in lines 621-624:

“Measurements using microtops were pointing directly at the sun when the line of sight between the microtops and the sun was free of clouds. This means that, under partly cloudy conditions, measurements were taken only if no clouds were in the line of sight to the sun. This explains microtop measurements in the cloudy category.”

12. L575: ... to be constant *with IWV* -> Calling this a bias could be maybe more precise.

Modified to: “to be constant with IWV”

However, calling it a bias will mean that one is the true value. Of course, radiosonde measurements are more accurate, but they could also have errors. For this, we decide to change "difference" to "discrepancies".

13. L577-578: See previous comment.

This has been addressed in a previous comment.

14. L584: exact quantification of *the occurrence and amount of* surface rain

Done

15. L586: Why can "one expect that the detection of rain is more likely to be similar among the instruments"? Given the different sampling methods and volumes, I would not agree with this statement.

Thanks. An explanation is added in lines 647-650:

“These factors could also impact the timing at which instruments effectively measure precipitation. The latter could be translated into the following question: How different is the occurrence of rain detected among the instruments during BOWTIE? This has implications for whether it is enough to use the rain flag from a single instrument or not.”

16. L622-624: Same as earlier comment: Why would you expect differences in the rain detection due to the alignment of sensors? Given the last statement, I would rather suspect the sensor sensitivities (or other technical issues) than their alignment causing the differences.

To address the point raised by the reviewer, the rain rates measured by the two disdrometers are compared only in precipitation events ($P > 0$). Figure R1.1 shows the scatter plot of rain rates for DSD_METEOR1 (P_{D1}) and DSD_METEOR2 2 (P_{D2}). P_{D2} shows greater values of rain rate compared to P_{D1} , in particular when rain rates are higher than 10 mm h⁻¹. Over the whole campaign, the mean difference in rain rate between disdrometer 2 and disdrometer 1 is 2 mm h⁻¹ for precipitation events. Figure R1.2 displays the scatter plot between the wind direction measured on the ship and the difference in rain rates between DSD_METEOR2 and DSD_METEOR1 ($P_{D2} - P_{D1}$) for precipitation events and is colored by the wind speed. The difference in rain rate is greater when the direction of the wind is between 200 and 300 degrees. The wind speed seems not to have a direct impact on the difference in rain rate.

Based on the analysis, we added the following text in lines 685-692:

“The differences in the frequency of rain detection between the disdrometers can be explained by the ship motion, the turbulence induced by the ship's structure (Yuter and Parker 2001), and the turbulence induced by the instrument (Friedrich et al., 2013). Induced turbulence by the ship's structure and instrument is tied to the changing nature of wind speed and direction, and ship motion. The difference in rain rate measured by the two disdrometers is related to the direction of the wind and their orientation with respect to the steaming direction of the vessel rather than the speed (not shown). In this sense, two disdrometers perpendicular to each other with respect to their sampling path are affected differently by rain from different wind directions and can differ in their rain detection measurements, in terms of the occurrence, as shown in this analysis.”

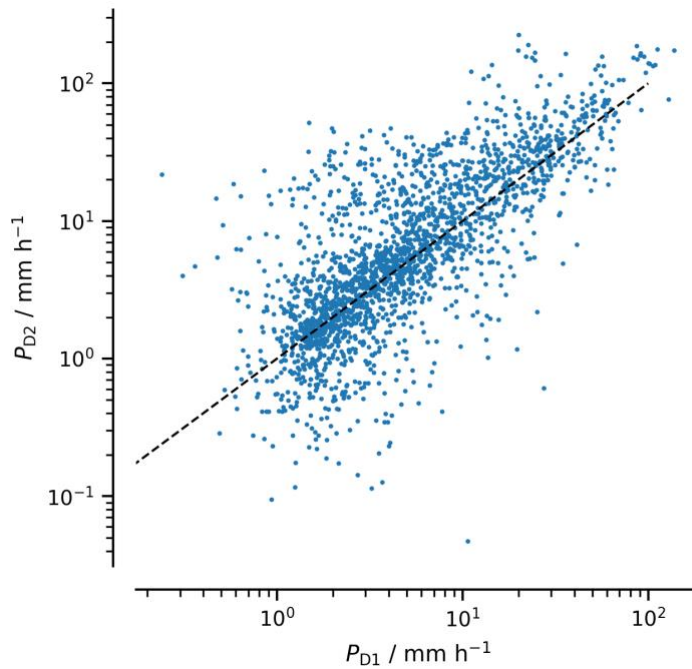


Figure R1.1. Scatter plot of the rain rates of two disdrometers: DSD_METEOR1 (P_{D1}) and DSD_METEOR2 (P_{D2}) for precipitation events ($P > 0$).

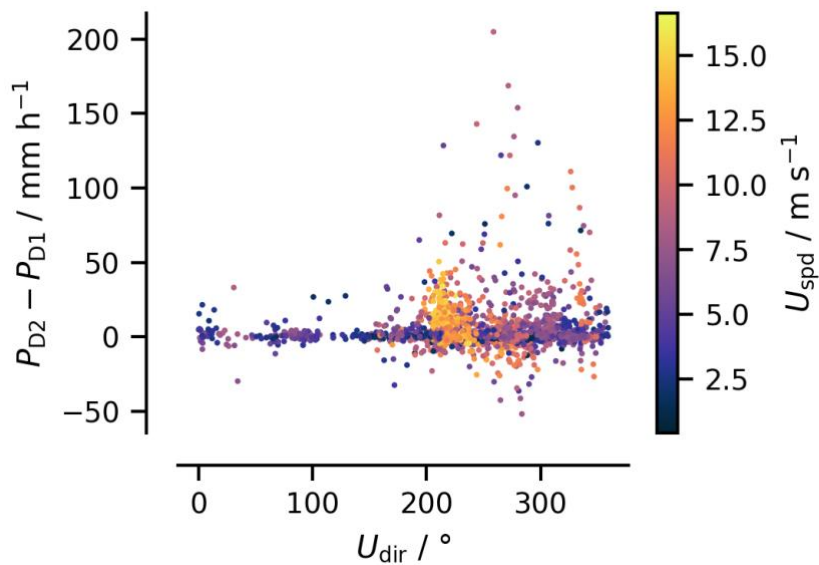


Figure R1.2. Scatter plot of the wind direction measured on the ship and the difference in rain rate measured in disdrometer 2 and disdrometer 1 ($P_{D2} - P_{D1}$) for precipitation events. Colors indicate the measured wind speed measured on the ship.

17. L638-640: Is this result specific to the BOWTIE deployment or instruments? Would you expect different results, if the same suite of instruments would be set up somewhere else?

As explained in the previous comment, the differences in the rainfall amount between instruments installed on a moving platform, such as a ship, have been shown in (Yuter and Parker 2001) and land-based disdrometers (Friedrich et al., 2013). Dry surface layers

can also explain differences between rainfall amounts of remote-sensing and in-situ instruments (Kalesse-Los et al., 2023). Thus, our result, which is related to the occurrence of rain detection, is expected to be found in other sites with the same setup of instruments as in BOWTIE. However, a detailed analysis should be made using other field campaign measurements. For the case of BOWTIE, the interest of this analysis is to address the issue of which data set should be taken as a rain flag. To stress this, the following text is added in lines 708-713:

“The differences in rain amount between instruments installed on ships have been addressed in (Yuter and Parker 2001). Dry near-surface layers can also explain differences between rainfall amounts of remote-sensing and in-situ instruments (Kalesse-Los et al., 2023). Our analysis shows that the rain occurrence between the instruments can also differ, even if measurements are averaged for 10 minutes. Among the instruments, the microrain radar and the rain gauge are the ones with a higher frequency of rain occurrence on the ship. Thus, using rain flags from only one instrument is not enough for the BOWTIE campaign.”

18. L714-715: Can you speculate on possible explanations? Has this been observed before? Is this specific to the measurement location or the sensor types?

We looked for past campaign reports to verify if the too-cold skin temperature was observed. To our knowledge, the coldest skin temperature is around 1K. We want to avoid speculating about this difference without a clear physical argument. Because of this, we decide to introduce the following text in line 781:

“To the knowledge of the authors, no previous campaigns have shown such a cold difference in $\Delta SST_{\text{skin}-5\text{cm}}$.”

19. L732-733: "should be taken with caution": What does this mean? Can you give more specific guidance for the data user?

We add the following argument in lines 800-801:

“Until a physical explanation is provided for such too cold skin-layers events, it is recommended to avoid the use of those events for scientific analysis“

20. L738-749: This paragraph requires much more context, e.g.: What is the significance of doldrums? How many events have been observed during the cruise? Was this a event special in any sense? Are there any new aspects of doldrums that cannot be studied with existing observations? Is the observed vertical structure consistent with expectations?

We have added more context to the case studies by adding more information at the beginning of Section 5 in lines 803-808.

“In this section, we present two scientific case studies that exemplify the use of the BOWTIE dataset and address key scientific motivations of the campaign. The first case study focuses on light-wind conditions, while the second examines a strong-wind regime. Both cases highlight the wealth of observations enabling the study of these atmospheric phenomena regarding processes in the atmospheric boundary layer, including near-surface variables in the atmosphere and SST in the case of the doldrums. While not studied here, the BOWTIE data set also allows the analysis of such atmospheric phenomena and their interaction with the upper ocean.

Moreover, a better context for studying the doldrums is added in lines 810-825:

“One atmospheric mesoscale phenomenon targeted by BOWTIE is the doldrums,

characterized by weak surface winds. Their role in shaping local convection within the ITCZ remains an open question (Windmiller 2024). Global storm-resolving simulations further suggest that the doldrums can influence tropical climate more broadly Segura et al., (2025). The impact of doldrums on local convection and the large-scale circulation is challenging to determine, because of the limited information on the scales controlling the air-sea energy and momentum transfer when winds slacken. This calls for a comprehensive characterization of the structure of doldrums regarding their energetics and momentum, i.e., the evolution of the atmospheric boundary layer and upper ocean. Two questions arise here: Were doldrums observed during BOWTIE? And what are the thermodynamic and dynamic structures of doldrums?

Using the DShip data, we identified instances in which the 10-min running mean of the wind speed measured at 30 m was below 3 m s^{-1} continuously for at least 10 minutes. There were 90 such instances of these light wind events. Prior studies have used various definitions of doldrums, which typically require the light wind events to be present for a sufficiently long period of time (hours) to indicate a persistent event that is also likely widespread spatially. During BOWTIE, we observed 28 light wind events that lasted for at least an hour, 18 for at least 2 hours, 10 for at least 3 hours, 8 for at least 4 hours, 6 for at least 5 hours, and 2 that lasted at least 6 hours. Thus, taking one hour as the threshold, 28 doldrums were observed. The longest-duration doldrum event was observed from August 26th at 11:10 UTC to August 27th at 2:25 UTC, having a duration of 15 h and 15 min. This is the event whose evolution is examined in Figure 11.”

This is also added in lines 833-839:

“Light winds occurred throughout the first kilometer of the atmosphere, and it is sustained throughout the event. This means that doldrums are not just wind events occurring at the surface but also in the whole boundary layer. The homogeneous structure of winds along the boundary layer is also observed in the structure of humidity and temperature. This means that despite the fact that winds are light, at least this doldrums event shows a homogeneous structure of the dynamics and thermodynamics in the boundary layer. Thus, one could ask which type of scales are driving this homogenization of the structure of the boundary layer and what the implications are for the surface heat fluxes during doldrums.”

21. L751-767: Same as for the previous paragraph: Please add more context on the significance of this case study. Since such convective gust fronts have been extensively observed and studied before (also over the tropical Atlantic), it would be important to clarify which new aspects this data set can add to the literature.

More context has been added in the first paragraph of Section 5 (see previous comment). Better context has also been added to the description of this case in lines 841-845:

“Convective precipitation results from interactions across scales, from cloud microphysics to the large-scale circulation, coupling thermodynamic, dynamic, and radiative processes. Characterizing convective systems therefore requires observations of their energetics and momentum, a key objective of BOWTIE, PICCOLO, and STRINQS. During BOWTIE, precipitation exceeding 1 mm was recorded by the rain gauge station on 21 days. On 8 September, ~8 mm of rain was measured. This amount of rain is due to a squall line crossing the ship. A gust front ahead of the squall line was also observed.”

The following text is also added in lines 863-866:

“The characterization in this study is only a first step towards a deeper analysis regarding the interaction between cloud particles and radiation, turbulence in the boundary layer, and the impact of convective systems in the air-sea exchange of momentum and energy. This can be done by combining measurements from the Sea-Pol and cloud radars, aerial drones, radiosondes, and radiation measurements on the ship.”

L798: "this does not depend on the velocity of the ship": To my understanding, this was not stated before (or only implicitly)

Thanks. We added the following information in Section 4.3, in line 740:

“This is observed when classifying the data with respect to the ship velocity (not shown).”

22. L799: same as previous comment

The following argument is added in lines 897-898:

“It is observed that the difference in the zonal component of the current between X-band and the ADCP is bigger than in the meridional component.”

This comment is already mentioned in the results section (see Section 4.3, lines 737-740).

“The greater bias metric is observed in the zonal currents than in the meridional currents. The root mean square error is also higher in the zonal currents than in the meridional direction. These results indicate a stronger impact of the wind drag in the zonal component than in the meridional one at 17-m depth.”

23. L803-804: Same as earlier comments. Please try to give more specific guidance to the data user.

We added the following arguments in lines 901-903:

“and it is recommended to avoid the use of skin temperature data in events when the skin layer is 1 K colder than the sub-skin layer.”

24. L808: "which is a first step for a deeper understanding": I find this statement way too broad and unspecific.

This comment is erased.

Technical Corrections

1. L7: Please use consistent notation for "2-D" vs "3D" (see L11 and elsewhere)

Done

2. L10, 43: upper ocean (without hyphen)

Done

3. L28: Define ocean Conductivity-Temperature-Depth-Oxygen (measurements?)

Done: ocean Conductivity-Temperature-Depth-Oxygen sensor

4. L45: Please define "doldrums" when first used

Doldrums are now defined in lines 48-51:

“calm winds ($< 3\text{-}4\text{ m s}^{-1}$) events in the equatorial regions, a.k.a. the doldrums in the formation of clouds and precipitation. Doldrums were highlighted in early wind atlases (Maury 1864) and re-emphasized in recent studies (Klocke et al., 2017, Windmiller 2024), and there are still unknowns in their impact on convection, momentum, and heat transfer across the air-sea interface.”

5. L52: I suggest to remove the repetition "a 40-day journey of FS Meteor" here

Done

6. L83: PIRATA

Done

7. L83-84: Please use a consistent notation for coordinates ("8°N, 38°W" vs "8°N and 38°W")

Done

8. L96: for atmospheric measurements

Done

9. Fig. 3, Caption: Please explain the meaning of the grey periods in panel a)

Done

10. L100: CTDO stations (without hyphen)

Done

11. L114: Please define the acronyms "PICCOLO" and "STRINQS"

Done

12. L116: Please define "drones" already on first use

Done

13. L117: lowest 1.5 km

Done

14. L123: EarthCARE

Done

15. L140-141: I suggest to merge these two sentence to avoid repetition.

Done

16. L150-153: Tables 2 and 3 are not referenced.

Done. See lines 162-164:

“Table 2 shows instruments taking continuous measurements, while Table 3 shows instruments deployed in selected stations (Section 2.2). Table 2 and Table 3 also display specific details of sensors and specific comments.”

17. L161: I suggest to make it a sub-sub-section (3.1.1 and following)

Done

18. L165: resolution of ... *for wind direction and wind speed*, respectively

Done

19. L187: Who is Prof. Hasse?

Thanks, the reference is eliminated.

20. L208: in Mindelo *before departure*

Thanks

21. L215: in off-wavelengths (?)

Done.

In off-wavelength Raman frequencies

22. L232: 90° elevation angle

Done

23. L253: Is this the same stabilisation platform already mentioned in L239-240? If so, I suggest to move the previous sentence here.

Done

24. L293: I suggest to move any support statements to the acknowledgments.

Done

25. L310 (and elsewhere): Please move full links to a footnote (as done earlier in the manuscript) or to the literature section

Done.

26. L321: The term "wind lidar" is rather general, given that the following paragraph also describes a wind lidar. I suggest a more specific paragraph heading.

Changed to “Wind lidar Abacus”

27. L398: "another foot or so": Please avoid slang and non-metric units.

It is 30 cm. This is corrected in the manuscript.

28. L402: recording averages with 1-s resolution/every second

Changed to: with a 1-s resolution

29. L403: sidetracked -> impacted/biased (or similar)

Changed to impacted

30. L442-444: Use lower case

Done

31. L518: Define "HF"

It was changed by "high frequency"

32. L519: Define "LOV"

This has been removed.

33. L535: latest *access* (?)

Modified to latest access: 24-04-2026

34. L536: from the ECMWF-IFS model

Done

35. L551: I suggest to make it sub-section 4.1

Done

36. L561: decrease *of* 0.37 kg/m²

Done

37. L571: Please be more specific than "interestingly"

More argument has been provided in lines 626-629

"IWV derived from microtops is substantially greater than radiosondes in clear-sky conditions ($\Delta \sim 2.5 \text{ kg m}^{-2}$), but the discrepancy is near 0 kg m^{-2} in cloudy conditions. Thus, microtops is the instrument retrieving greater values of IWV compared to the other three sensors in cloudy and clear-sky conditions."

38. L571: IWV is "higher", not "moister"

Changed to greater

39. L573: trustworthy -> reliable/accurate

Changed to accurate

40. L577+579: gap -> difference

Done

41. L591: There is no Section 3.1.2.

Modified to Section 3.2.1

42. L603: detected by a different number of instruments

Done

43. L662-668: Remove (redundant information that belongs to the figure caption)

Done

44. L671-674: Move to the figure caption.

Done

45. L677: between the ocean surface and the lower atmosphere

Done

46. L689: section 3.2.3

Done

47. L691-693: Move to the figure caption.

Done

48. L704: Please use K for temperature differences (here and elsewhere in text and figures)

Done

49. L710: Delete "during the night" (repetition)

Done

50. L711: is *observed for* 4.7 % of ...

Done

51. L726-727: can be -> is

Done

52. L727: Delete "during the day" (repetition)

Done

53. L728: Only one *measurement*

Done

54. L730: Delete "in general"

Done

55. Fig. 10a: Please decrease the widths of the blue and grey lines to increase readability.

Done

56. Fig. 10, caption: Remove "y axis" and "x axis"

Done

57. Fig. 10, caption: Similar to -> same as

Done

58. Figs. 11 and 12, caption: Doldrums event on August 26, 2024: ...

Done

59. L769: Ocean

Done

60. L770: moist (?) tropical Atlantic

Changed to: moistest region of the tropical Atlantic

61. L781: FS Meteor (lower case)

Done

62. L791: by *up to/on average* (?) 2 kg/m²

Done

63. Data set/DOI table: What is RAPSODI?

The name is in the footnote.

References:

Kollias, P., & Bartholomew, M. J. (2016). Parsivel Disdrometer Support for MAGIC Field Campaign Report. DOE ARM Climate Research Facility, Pacific Northwest National Laboratory, Richland, WA. <https://doi.org/10.2172/1255440>

Friedrich, K., Higgins, S., Masters, F. J., & Lopez, C. R. (2013). Articulating and stationary PARSIVEL disdrometer measurements in conditions with strong winds and heavy rainfall.

Journal of Atmospheric and Oceanic Technology, 30(9), 2063–2080.
<https://doi.org/10.1175/JTECH-D-12-00254.1>

Yuter, S. E., & Parker, W. S. (2001). Rainfall Measurement on Ship Revisited: The 1997 PACS TEPPS Cruise. *Journal of Applied Meteorology*, 40(6), 1003–1018.
[https://doi.org/https://doi.org/10.1175/1520-0450\(2001\)040<1003:RMOSRT>2.0.CO;2](https://doi.org/https://doi.org/10.1175/1520-0450(2001)040<1003:RMOSRT>2.0.CO;2)

Referee 2

< General comments >

This article provides basic information about observations made during the BOWTIE cruise that would act as a reference for future use. The authors also demonstrate the characteristics of ocean and atmosphere during the two cases; calm conditions and convectively active conditions as indications for future studies. The authors' efforts producing quality-controlled data and providing those data to the broader community would be highly appreciated. However, several descriptions are unclear. In addition, while I understand this is an introduction paper of BOWTIE datasets, some "depths" are needed to ensure their data reliability. I put some questions and suggestions as specific comments below. Please revise them accordingly.

We thank the reviewer for the constructive comments of our manuscript. By addressing the reviewer's comments and incorporate in the manuscript, it has undergone an improvement. In the new version of the manuscript, we give more information on the calibration of the instruments. Section 3.1 has also undergone a reorganization, highlighting key details of the instruments. Furthermore, a better context for the two cases study has been added in the new version of the manuscript. In the following lines, we addressed each comment posed by the reviewer.

< Specific comments >

1) L. 67-69; It is helpful to know the locations of two ports, so that the readers can obtain UTC from their local times and also can find rough distance between them. (I know we can see it in Fig. 2, but it is better to describe it in the text.)

Done. The port of Mindelo is located at 16.8896N and 24.9993W, and the port of Bridgetown is located at 13.106N and 59.632W. This information has been added in lines 69-71.

2) L. 97; It might be better to add "aerial" before "drone", as recently uncrewed ocean surface vehicles such as saildrone and wave-glider are also referred as marine (oceanic) drones.

It was modified to aerial drone

3) L. 111-143; It is one idea that section 2.3 can be summarized into one table, which lists the simultaneous observations among ORCESTRA's sub-campaigns instead of the current description style. I mean, it is not necessary to divide into two sub-sections. Instead, one section with one table might be enough. Or, a paragraph that explains the following two sub-sections can be inserted at the beginning of this section.

We have provided a paragraph introducing both subsections. See lines 115-123:

"BOWTIE was closely coordinated with four other sub-campaigns of ORCESTRA, both through measurements conducted from the same platform and through coordinated observations across multiple platforms. PICCOLO (Process Investigation of Clouds and Convective Organization over the Atlantic Ocean; Wing et al. 2026, manuscript submitted to Bull. Amer. Meteor. Soc.) and STRINQS (Soundings and TuRbulent eddy measurements in the ITCZ with a Network of QuadcopterS; Stevens et al., 2026) operated alongside BOWTIE aboard *FS Meteor*, complementing the ship-based observing system with additional radar and aerial drone measurements. PERCUSION (Persistent EarthCARE Underflight Studies of the ITCZ and Organized Convection) and CELLO (Cloud and EarthCARE caL/vaL Observations) operated

from independent airborne platforms, with their observations coordinated with *FS Meteor* and EarthCARE overpasses. These two forms of collaboration are described separately below.”

4) L. 114; If there is no published work on STRINQS, at least, relevant paper (Stevens et al. 2026 Tellus paper) should be cited, as the authors did for PERCUSION and CELLO in the next subsection. In addition, the readers cannot imagine anything from its abbreviation alone. Please spell them out.

The meaning of STRINQS is written in the manuscript, and Stevens et al. (2026) is cited.

5) L. 131; These numbers (17 overpass and 7 additional ...) might be the ones only during the period corresponding to the BOWTIE campaign. I found the different numbers (23 flights during Aug. 11 - Sept. 28) in Table 1 of Stevens et al. (2026, Tellus). If so, it is better to describe such facts to avoid any misleading on PERCUSION and/or HALO campaign period.

The differences between Stevens et al. (2026) and this manuscript are in how the flights are counted. In Stevens et al. (2026), HALO, the aircraft used in PERCUSION, had 23 flights. But not all HALO flights coordinated with BOWTIE, and of those that did, a flight can be composed of several overpasses to *FS Meteor* or non-overpass close approaches. In total, there were 17 overpasses, and in other 7 times, *FS Meteor* was close to HALO. To specify the difference between days and overpasses, lines 142-143 are modified:

“In total, 17 overpasses and 7 additional close approaches to *FS Meteor* were achieved in coordination with BOWTIE in 10 days between 16th August and 23rd September.”

Moreover, we add the following text to direct the reader to more information about the coordination between BOWTIE and both PERCUSION and CELLO, see lines 149-150:

“For more information about the coordination between BOWTIE and both PERCUSION and CELLO, refer to Klocke et al. (2026).”

6) L. 156-205; Was any calibration for each sensor performed before/after the cruise? Such calibration is useful to detect temporal drifts. If only annual maintenance by the instrument manufacturers is done, it should be noted. In any case, please provide calibration information (not calibration results).

Yes, calibration of instruments was performed before, during, and after the cruise, and the period of calibration depends on each instrument. The calibration information is now detailed within the description of each instrument in Section 3.

7) L. 206-530; It is better to provide the same basic information for all instruments with notification (name, manufacturer, type, sampling intervals, etc.) first. It is hard to understand the current descriptions which is which. Or, simply refer to Tables 2 and 3, which tell above. In addition, it is possible to omit descriptions of sensor types in the text, as we can see them in the tables.

We have revisited the description of the instruments to take into account the reviewer’s comment. See section 3.2.

8) L. 251-270; After all, final quality-controlled datasets of W-band radar have not been produced yet. Right? If so, it should be mentioned. Also, if any definitions of QC level (or version) exist, it should be noted for future updates.

The data set provided by Cloudnet, which has been used in this manuscript, has passed an internal quality control procedure. This quality control indicates a strong signal attenuation, and it has been labeled as level 2a data. However, this quality control procedure does not indicate a correction of the data; it is just a flag. Further flagging regarding the ship's motion will be added in data level 2b.

This discussion is added in lines 593-598:

“Using the quality-controlled and heave-corrected radar data (level 1 data), Cloudnet creates a rain detection flag for the W-band cloud radar dataset. Subsequent data analysis revealed that the rain flag is true when the reflectivity in the fourth range gate (at 165m) exceeds 3.1 dBZ and the Mean Doppler Velocity (MDV) is negative. The dataset resulting from the Cloudnet processing gives a level 2 product. Adding a flag for ghost echoes and strong radar signal attenuation manually results in level 2a data. An additional variable of radar beam misalignment angle from zenith will be added in the level 2b product for a better assessment of the bias induced to MDV observations by horizontal wind speeds.”

9) L. 367-370; Any special reason to set the timing of reaching top of troposphere at the nominal 00/03/./21 UTC? As the authors may know, WMO requires to launch the balloon between 45-min to 15-min prior to the nominal time as much as possible. This makes it possible to re-launch in case of failure within the same nominal time slot. (So, I guess re-launch was not performed during the BOWTIE as the same slot.) Also, I wonder if those data were sent to global communities through the GTS or not. If yes, it should be noted.

The selected timing was to follow the DWD protocol. Moreover, following this protocol allowed data from launches reaching the top of the troposphere at 00/06/12/18 UTC to be sent to the Global Telecommunication System. This text is added in the manuscript in lines 374-378:

“ Launches were timed approximately 1 h 20 min before the target observation times of 00, 03, 06, 09, 12, 15, 18, and 21 UTC, such that the balloons reached the upper troposphere near these times, following the DWD protocol. Soundings targeting 00, 06, 12, and 18 UTC were performed using the DWD ground station and transmitted to the Global Telecommunication System (GTS) for operational weather forecasting and data assimilation. One of the eight routine daily launches was performed by DWD personnel”

The following lines are also included in lines 388-389:

“Further information on quality control, five cases of oscillating balloons, and transmission to the GTS is provided by Winkler et al. (2026).”

10) L. 375-376; I wonder if any quality check on near surface data was performed or not. As long as I read Schulz et al. (2024) package, which is cited here for postprocessing, I cannot find any explicit description on that. In the past big campaigns in the tropics such as TOGA-COARE in the western Pacific and DYNAMO in the Indian Ocean, special care was taken for radiosonde near-surface data obtained onboard especially in the calm sea conditions (e.g., Wang et al., 2002; Ciesielski et al., 2014), which might affect boundary layer research. If apparent errors are not found, it is fine. If check was not done, at least, it should be noted. Please note that in old days (like TOGA-COARE era) many errors were attributed to radiosonde itself or operation. However, I mention here that even accurate data were taken, data itself around the ship might be distorted due to the ship structure. For example, I'm afraid of any influence from a funnel and deck, if a schematic in Fig. 4a demonstrates actual operation. So, quality checks are required to confirm it.

References.

Ciesielski, P. E., H. Yu, R. H. Johnson, K. Yoneyama, M. Katsumata, C. N. Long, J. Wang, S. M. Loehrer, K. Young, S. F. Williams, W. Brown, J. Braun, T. Van Hove, 2014: Quality-controlled upper-air sounding dataset for DYNAMO/CINDY/AMIE: Development and corrections. *J. Atmos. Oceanic Technol.*, 31, 741-764. <https://doi.org/10.1175/JTECH-D-13-00165.1>

Wang, J., H. L. Cole, D. J. Carlson, E. R. Miller, K. Beierle, A. Paukkunen, and T. K. Laine, 2002: Corrections of humidity measurement errors from the Vaisala RS80 Radiosonde-application to TOGA COARE data. *J. Atmos. Oceanic Technol.*, 19, 981-1002.

[https://doi.org/10.1175/1520-0426\(2002\)019%3C0981:COHMEF%3E2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019%3C0981:COHMEF%3E2.0.CO;2)

Radiosondes in BOWTIE followed the quality control described in Winkler et al. (2026). This means that no special quality control was done in the first 40 meters. As it was mentioned by Yoneyama et al. (2002) and Ciesielski et al. (2014), distortions in the radiosonde measurements of air temperature and humidity can appear in the first 50 meters of the lower atmosphere due to the ship structure. Because of the reviewer comment, a short analysis was done in the first levels of the atmosphere in Level 2 data (Figure R2.1a). From 10 to 20 m height, a temperature drop is observed; then there is an increase of temperature with height, forming a V shape in the temperature profile. This V shape is also observed in the Level 1 data. Large temperature drops are related to warm and dry conditions, as observed in the measurements on FS Meteor (Figure R2.1b). The reason for this behaviour is not fully understood. The temperature drop could be related to the effect of the ship structure, as pointed out by Yoneyama et al. (2002). The increase in temperature with height suggests an inversion layer, but a better understanding of this should be gained. Then, the recommendation is not to use data below 20 m in height.

In the manuscript, the following text is added in lines 384-388:

“No specific quality control was performed on the first 50 meters of the measurements. In the Level-2 data, temperature frequently decreases by more than the dry adiabatic lapse rate between 10 and 20 m before increasing over the subsequent one to three levels (not shown). The cause is not fully understood but may involve effects of the ship's structure (Yoneyama et al., 2002). We therefore recommend excluding radiosonde measurements below 30 m in the Level-2 data and the first three height levels in the Level-1 data.”

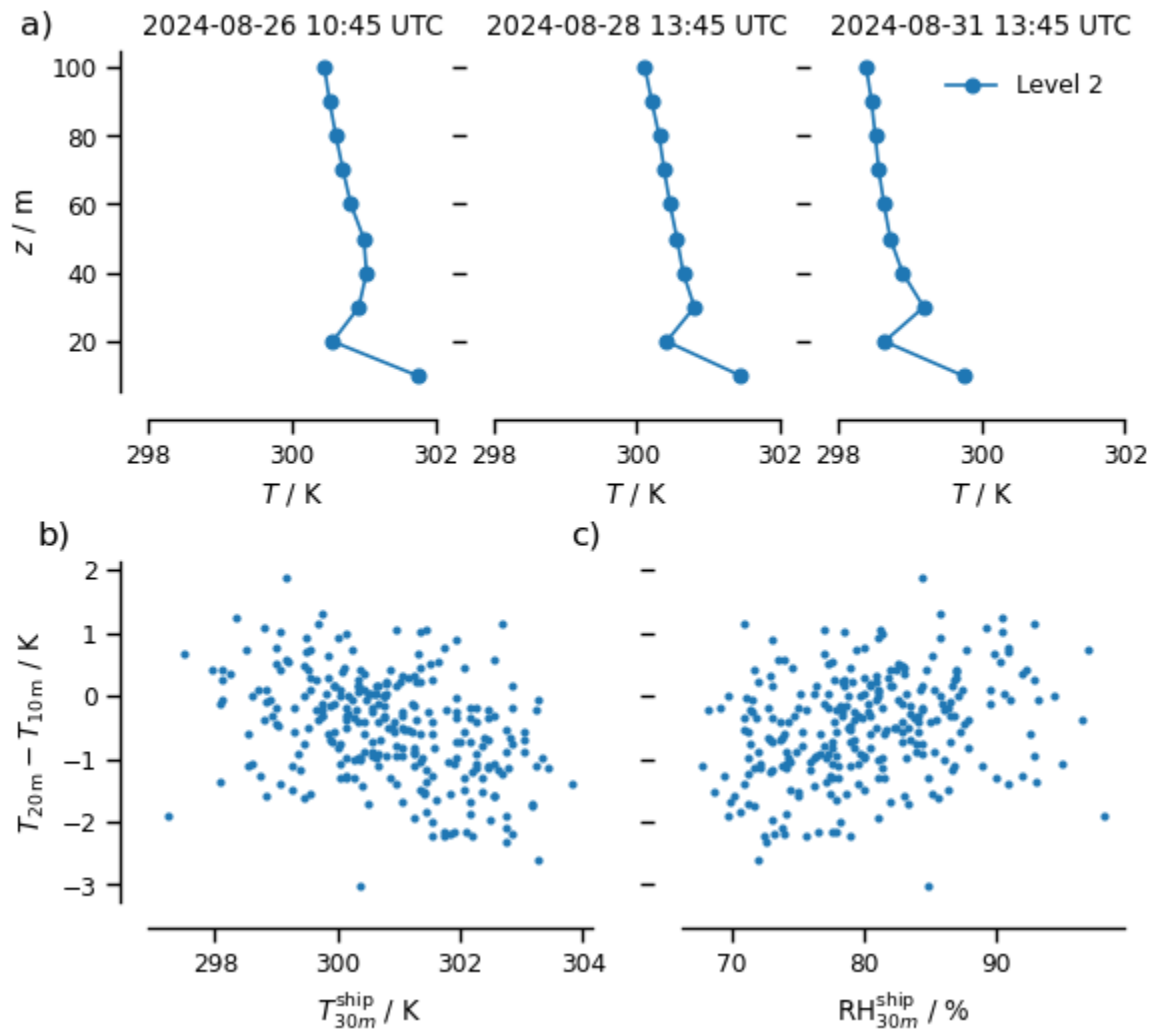


Figure R2.1. a) Temperature profile from the Level 2 of radiosonde data for three different times. b) Scatter plot of the difference in temperature at 20 m and 10 m ($T_{20m} - T_{10m}$) and the air temperature measured on ship (T_{30m}^{ship}). c) Similar to b) but with the relative humidity measured on ship (RH_{30m}^{ship}).

11) L. 391-396; Since ISAR cannot be operated under rain, this fact should be noted.

This text was incorporated in line 418:

“The storm shutter prevents measurements during rainy conditions.”

12) L. 399-402; I cannot understand this description, because usually a sea snake is deployed near the bow or far side avoiding the ship’s wake. Fairall et al. (1997), which is cited here, also mentioned they deployed a sea snake from an outrigger boom about 4 m from the ship’s hull. Similar discussions can be found in many papers (e.g. de Szoeke, 2021). Could you confirm this point (sea snake was deployed within the influence of wake intentionally)?

Reference.

de Szoeke, S. P., 2021: Fast floating temperature sensor measures SST, not wet-bulb temperature. J. Atmos. Oceanic Technol., 38, 995-1000. <https://doi.org/10.1175/JTECH-D-20-0193.1>

The sea snake was deployed intentionally in the ship's wake. The objective of the thermistor in the sea snake is to give a representative measurement of the sub-skin layer. The near-surface mixing in the ship's wake is providing these representative measurements because the sub-skin layer is being renewed, and foam and bubbles are proof of this mixing. Moreover, the vertical motions of the sea snake within the wake help in providing a representative measurement of the sub-skin layer. One disadvantage of colocating on the wake is that the thermistor can jump out of the water, but the thermistor used in BOWTIE is slow-responding. This means that it does not quickly adjust to the air temperature or the temperature of bubbles. This is not the case for the RBR sensor, used in Szoeké (2021). Then, installing the sea snake in the ship's wake is a good option. This has already been explained in the manuscript.

One important thing to take into account is to avoid the thermistor measuring warm water being drained from the ship's plumbing, or heated by the engine below the ship. For this reason, the sea snake was colocated as described in the manuscript. It would have been perfect to install it near the bow, but for structural issues, this was not possible. However, the measurements of the calibrated thermosalinograph and the sea snake were very close during late night. This is expected because during nighttime, mixing in the upper ocean homogenizes temperature. Then, sea snake measurements are taken as valid. To avoid confusion in what Fairall et al. (1997) did and what BOWTIE did with respect to the installation of the sea snake, the following paragraph is modified in lines 423-428:

“During BOWTIE, the sea snake was towed from the port side of the midship deck using an outrigger boom, approximately 3 m from the ship's hull. The sensor consists of an exposed slow-response thermistor attached near the end of a floating, air-filled rubber tube, with approximately 30 cm of tubing extending beyond the sensor to keep it submerged. The deployment minimizes contamination by warm water discharged from the ship and heating associated with the hull and engines. Although the installation differs from that described by Fairall et al. (1997), the measurement principle is the same.”

13) L. 493-500; It might be better to cite Pinkel et al. (2011) paper on Wire Walker explanation.

Reference.

Pinkel, R., M. A. Goldin, J. A. Smith, O. M. Sun, A. A. Aja, M. N. Bui, and T. Hughen, 2011: The Wirewalker: A vertically profiling instrument carrier powered by ocean waves. *J. Atmos. Oceanic Technol.*, 28, 426–435. <https://doi.org/10.1175/2010JTECHO805.1>

Done. The reference has been added in lines 534-535:

“The WireWalker is a freely drifting, wave-powered platform that repeatedly profiles the water column (Pinkel et al., 2011).”

14) L. 551-582; Integrated water vapor. It is not necessary to add any figures, but I'm wondering if the authors compared those data from diurnal cycle viewpoint. This is useful to confirm dry bias of radiosonde due to solar radiation. Besides, as asked in my previous comments on radiosonde measurement, previous studies indicated careful treatment on near-surface data was needed for ship-based data, which may cause drier/wetter in daytime/nighttime. So, I believe it is worth checking.

The integrated water vapor using radiosonde data is calculated taking into account heights above 30 m. Following the suggestion of the reviewer, it is checked whether or not radiosondes show a dry bias during daytime compared to nighttime. Because there is no other source of integrated water vapor measurement with small uncertainties, the differences in the integrated water vapor

between radiosondes and the microwave radiometer are used. If the differences between radiosondes and the microwave radiometer decrease during daytime, this could suggest a dry bias in radiosonde measurements during daytime. Figure R2.2 shows the probability density function of integrated water vapor calculated using radiosonde data and the microwave radiometer for clear-sky conditions. Using local hours, daytime (10 LT to 18 LT) and nighttime (22 LT to 5 LT) measurements are selected. The difference in integrated water vapor between the radiosondes and the microwave radiometer is similar during daytime and nighttime ($\sim 1.7 \text{ kg m}^{-2}$). Because the difference is constant, no apparent systematic dry bias in the radiosondes is evident. The analysis also shows that both radiosondes and the microwave radiometer display greater values of integrated water vapor during nighttime compared to daytime ($\sim 1.4 \text{ kg m}^{-2}$). This day-night difference has been analyzed in previous campaigns (Ruppert 2016, Ruppert and Johnson 2015), but the main mechanism is still an open question.

Going back to the dry bias pointed out by the reviewer, the following text is added in lines 634-636:

“Moreover, the discrepancies between the integrated water vapor measured by the radiosondes and the microwave radar are independent of daytime and nighttime (not shown). This indicates no dry bias in radiosondes during daytime.”

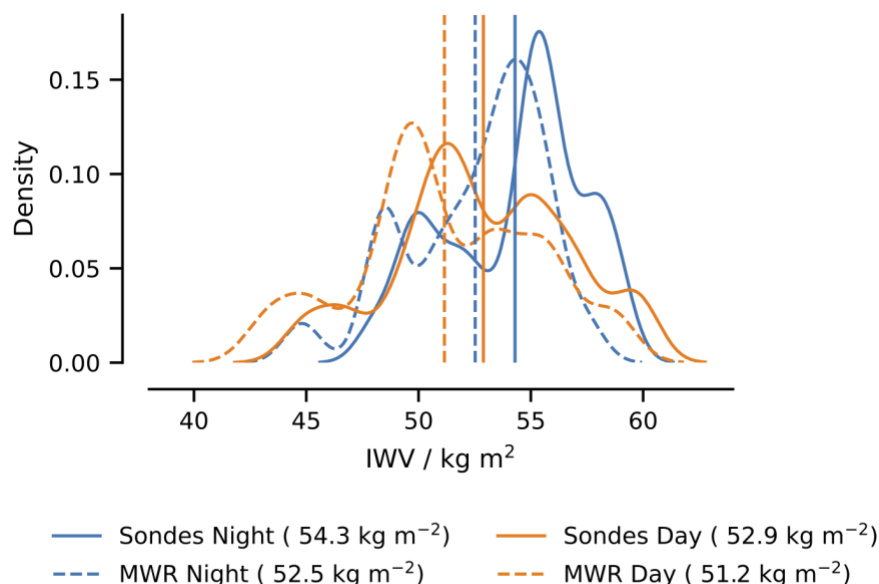


Figure R2.2. Probability density function of the integrated water vapor using radiosonde measurements (solid lines) and microwave radiometer (dashed lines). Nighttime measurements (blue), considered from 22 LT to 5 LT, and daytime measurements (orange), considered from 10 LT to 18 LT. In parentheses, the mean value is indicated. Only measurements during clear sky conditions are used.

15) L. 607-637; In L. 271-279, the authors mentioned MRR had problems with deployment. Doesn't it affect anything on detection technically?

Rain flag values of the MRR are not affected by the fact that it had to be tilted by 7° and was not stabilized.

This has been explicitly stated in lines 269-270:

“This type of measurement is not affected by the MRR tilting or by not having a stabilizing platform.”

16) L. 734-767; I cannot see the “philosophy” of these selections as examples. What kind of actions do the authors expect from the readers as a next step? Or, the authors simply demonstrate the usefulness of data? If the authors intend to provide the readers with guidance for future studies, it should be explicitly mentioned using these examples.

Reviewer 1 also pointed out this issue. Due to comments from Reviewer 1 and Reviewer 2, a better context have been provided in the new version of the manuscript in Section 5.

17) L. 748; “11 UTC” might be better to be expressed as local time.

We think it’s better to keep UTC because it’s easier for a reader to look into the data. However, we add the local time in parentheses next to the UTC time. This is to give a more local perspective of the time, as the reviewer is suggesting.

< Technical corrections >

1) Fig. 2, caption; typo (pepth -> depth)

Done

2) L. 147; STRIN”G”S --> STRIN”Q”S

Done

3) L. 278-279 and L. 627; Which is correct, “> 3 dBz” or “greater than 3.1 dBz”?

Corrected. It is greater than 3 dBZ

4) L. 591; Not Section “3.1.2” but “3.2.1”?

Corrected

5) L. 929; Exact citation is required. This paper is from “Reviews of Geophysics, 17, 1762-1772. <https://doi.org/10.1029/RG017i007p01762>”

Done

6) L. 967; Published. <https://doi.org/10.1175/BAMS-D-25-0166.1>

Done

Reference:

Klocke, D., Wing, A. A., Segura, H., Dengler, M., Bell, M. M., Ruppert, J. H., George, G., Kalesse-Los, H., Nuijens, L., Möller, K. O., Kiko, R., Mohr, W., Kidane, A. T., Trosits, A., Mertz, C., Imker, C., Begler, C., Blandfort, D., Colón-Burgos, D., ... Windmiller, J. (2026). 40

days and nights in the Rains: The inner life of the Atlantic ITCZ during BOWTIE. *Bulletin of the American Meteorological Society*. <https://doi.org/10.1175/bams-d-25-0166.1>

YONEYAMA, K., HANYU, M., SUEYOSHI, S., YOSHIURA, F., & Katsumata, M. (2002). Radiosonde observation from ships in the tropical region. In JAMSTECR (Vol. 45). https://www.godac.jamstec.go.jp/doc_catalog/external/metadata/shiken45_04/file/shiken45_04.pdf

Ruppert, J. H. (2016). Diurnal timescale feedbacks in the tropical cumulus regime. *Journal of Advances in Modeling Earth Systems*, 8(3), 1483–1500. <https://doi.org/10.1002/2016MS000713>

Ruppert, J. H., & Johnson, R. H. (2015). Diurnally Modulated Cumulus Moistening in the Preonset Stage of the Madden–Julian Oscillation during DYNAMO. *Journal of the Atmospheric Sciences*, 72(4), 1622–1647. <https://doi.org/https://doi.org/10.1175/JAS-D-14-0218.1>