Reviewer #1

General comments

This manuscript introduced an analogue-based attribution framework, TurboMeter, which had been developed to assess relative contributions of anthropogenic climate change and natural variability on aviation turbulence in changing climate. The key procedures to the framework are, first, select turbulence-prone days in an independent time frame (a one year period of 2024 in this study) using automated in-situ turbulence observations from commercial aircraft; second, characterize the synoptic-scale patterns of the selected turbulence days using the 500-hPa geopotential height in ERA5 as a single metric; third, for each of the selected turbulence days, identify analogues in both past (1950–1986) and present (1987–2023) climates that have similar 500-hPa geopotential height patterns with the selected turbulence event; and then last, compute and compare a single turbulence index (TI1) and its components (vertical wind shear (VWS) and deformation (DEF)) to assess if there are any changes between the past- and present-climate analogues despite the similar 500-hPa geopotential height patterns.

I agree to the necessity of assessing the attribution of anthropogenic climate change to aviation turbulence, separately from the attribution of natural variability. However, I have several major concerns about the algorithms on which the TurboMeter framework is built on, i.e., the characterization of the synoptic-scale patterns and the case selection, in particular. In brief, I doubt if using the 500-hPa geopotential height alone is sufficient to characterize synoptic-scale patterns and identify analogues, especially characterizing the synoptic conditions that are critical to generation, maintenance, and intensification of aviation turbulence; in addition, the case selection was made based on the EDR threshold only, without separating sources of turbulence (e.g., clear-air turbulence, mountain-wave turbulence, convectively induced turbulence), which seems to make it even harder to justify using the 500-hPa geopotential height alone as a single metric to identify synoptic-scale patterns responsible for the selected turbulent days.

Besides my concerns on the TurboMeter framework, the analysis of the TurboMeter output needs to be improved to provide more in-depth explanation of physical mechanisms supporting the output. The result plots in the manuscript focused on 500-hPa geopotential height, TI1, VWS and DEF. However, the main texts mentioned and relied heavily on changes in and patterns of the jet stream in explaining the TurboMeter output, while the jet stream and other synoptic-scale patterns were never shown or compared in the manuscript. I think the analysis needs to be improved by explicitly showing the differences and similarities in jets, troughs, and other possible sources of turbulence between the past and current climate analogues.

While I agree to the motivation, objective, and necessity of this study, I have major concerns about the methodology that would need significant revisions and could lead to significant changes in the results, as well as lack of in-depth analysis on physical mechanisms explaining the TurboMeter output. Therefore, my recommendation is to reject this manuscript to allow the authors sufficient time to work on their revisions but strongly encourage to re-submit the manuscript.

We sincerely thank the Reviewer for the thorough, thoughtful, and highly constructive evaluation of our manuscript, as well as for recognizing the importance of developing attribution

frameworks for aviation turbulence. We acknowledge that the original version did not sufficiently explain several key aspects of the TurboMeter framework, particularly the philosophy underlying the analogue-selection procedure, the role of the 500-hPa geopotential height field, the interpretation of circulation-conditioned attribution, and the distinction between changes in turbulence-favourable atmospheric environments and changes in the individual physical mechanisms responsible for turbulence generation. We also agree that the physical interpretation of the results could be considerably strengthened through additional atmospheric diagnostics and a more comprehensive discussion of the underlying dynamical and thermodynamical processes.

In response to these comments, we have substantially revised the manuscript while preserving the original methodological framework. Rather than modifying the analogue methodology itself, which follows the circulation-conditioned attribution philosophy previously developed within the ClimaMeter framework and other attribution studies, we have clarified its scientific scope and considerably strengthened its physical interpretation.

The revised manuscript includes several major improvements:

- **The philosophy of the analogue-selection procedure has been clarified throughout the manuscript. We now explicitly explain that Z500 is used exclusively to constrain the large-scale synoptic circulation and not to reproduce the complete atmospheric state associated with each turbulence event.**
- **The methodology has been expanded by introducing additional atmospheric diagnostics that are analysed after analogue selection, including upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN).**
- **The Results section has been substantially reorganized. In addition to the turbulence diagnostics, each regional case study now includes surface and upper-tropospheric diagnostics that reveal how thermodynamic and dynamical environments evolve under comparable circulation conditions.**
- **The Discussion has been completely revised to distinguish the physical pathways emerging over oceanic and continental regions, highlighting both the common features and the regional differences in the climate signal.**
- **The Conclusions have been substantially expanded to clarify the methodological philosophy of TurboMeter, explicitly discuss the rationale for using Z500 as the circulation-conditioning variable, and emphasize the limitations of ERA5 in resolving mesoscale and microscale turbulence-generation mechanisms, including mountain-wave and convectively induced turbulence.**

We believe that these revisions substantially strengthen the manuscript and directly address the Reviewer's concerns while preserving the scientific foundations and the original philosophy of the proposed analogue-based attribution framework.

Major comments

1. This study used a single metric of 500-hPa geopotential height to characterize synoptic-scale patterns of turbulence-prone days in 2024 and then to identify analogues (i.e. days with similar synoptic-scale patterns) in past (1950–1986) and present (1987–2023) climates. Then, based on the comparison between the past and present climate analogues, the contribution of anthropogenic climate change was determined; that is, if the characteristics of TI1 are found differently between the past- and present-climate analogues despite similar synoptic-scale patterns, their differences can be attributed to anthropogenic climate change.

For this analogue-based attribution framework works, it should be validated that the 500-hPa geopotential height is a good metric to characterize synoptic-scale patterns that generate, maintain, and intensity turbulence in the selected turbulent days. However, this study does not provide that critical validation of the metric (500-hPa geopotential height). This validation can be done by comparing the weather patterns between the analogues for each case. Do the past and climate analogues selected based on the 500-hPa geopotential height have similar synoptic weather patterns (i.e., jet, trough, baroclinicity, convection, etc.)? Or do they just happen to have similar 500-hPa geopotential height patterns but show very different weather conditions among the analogues? Without such a validation of the use of 500-hPa geopotential height as a metric to characterize synoptic-scale patterns, it is hard to justify the outcomes of the attribution method based on the metric.

We sincerely thank the Reviewer for this thoughtful comment, which addresses the central methodological aspect of the TurboMeter framework. We fully agree that the original manuscript did not sufficiently explain the rationale behind the use of the 500-hPa geopotential height (Z500) field for analogue selection, nor did it provide enough evidence that the selected analogue populations are characterized by physically consistent atmospheric environments beyond the similarity of the Z500 field itself. We have therefore substantially revised the manuscript to clarify both the methodological philosophy of TurboMeter and the physical interpretation of the selected analogues.

The purpose of the analogue-selection procedure is not to identify atmospheric states that reproduce all the physical processes responsible for a particular turbulence event, nor to demonstrate that Z500 alone can characterize the complete multiscale pathway leading to aircraft-scale turbulence. Rather, following the circulation-conditioned attribution previously developed within the ClimaMeter framework (Faranda et al., 2024; Faranda and Alberti, 2026), the objective is to constrain the large-scale synoptic circulation, thereby allowing the contribution of climate change to be investigated under comparable dynamical conditions.

Within this framework, the complete Z500 field is used because it provides a robust and physically meaningful description of the synoptic-scale circulation governing the evolution of weather systems. Unlike reduced-order representations based on weather regimes or empirical orthogonal functions, the analogue search is performed using the full geopotential-height field, preserving the complete spatial organization of troughs, ridges, Rossby-wave patterns, and planetary-scale circulation anomalies. The analogue populations are therefore selected to represent comparable large-scale circulation states rather than identical atmospheric realizations.

To address this point directly, the revised manuscript now explicitly presents the composite Z500 fields for the target event together with the corresponding past- and present-climate analogue populations for each case study. These new figures provide a direct visual assessment of the analogue-selection procedure and demonstrate that the selected analogue populations successfully reproduce the large-scale synoptic configuration of the target event, including the position and amplitude of the main troughs, ridges, and planetary-wave structures. The close agreement between the event and both analogue composites provides independent evidence that the analogue search robustly constrains the large-scale circulation before any turbulence-related diagnostics are analysed. Building upon this validation, one of the principal improvements of the revised manuscript is the introduction of several independent atmospheric diagnostics that were not used during the analogue-selection procedure. In addition to the original turbulence diagnostics (TI1, VWS and DEF), the revised manuscript now includes analyses of upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). These additional diagnostics provide an independent physical assessment of the analogue populations. The new analyses consistently show that the selected past and present

analogues reproduce remarkably similar large-scale circulation structures, including comparable trough-ridge configurations and jet-stream locations, while simultaneously revealing systematic differences in the thermodynamic and dynamical environments associated with present-day climate conditions. In particular, the present-day analogue populations generally exhibit enhanced upper-level wind speeds, stronger vertical wind shear, warmer surface conditions, increased atmospheric moisture, and, in some regions, enhanced convective instability. Because these variables are completely independent of the analogue-selection procedure, they provide additional confidence that the selected analogue populations are physically coherent and that the observed differences reflect climate-related modifications of the atmospheric environment rather than artefacts of the analogue search itself.

Accordingly, the revised manuscript explicitly describes Z500 as the variable used to condition the large-scale circulation, while the additional atmospheric diagnostics are used to independently evaluate how the associated atmospheric environment changes between the historical and present-day analogue populations. We believe that this clarification considerably strengthens both the methodological presentation and the physical interpretation of the framework.

Finally, we note that the objective of TurboMeter is fundamentally different from validating Z500 as a universal predictor of aviation turbulence. Rather, the framework is designed to quantify how turbulence-favourable atmospheric environments evolve under comparable synoptic-scale circulation conditions. We have revised the manuscript throughout to make this distinction explicit and to consistently describe TurboMeter as a circulation-conditioned attribution framework, thereby avoiding possible overinterpretation of the role of the analogue-selection variable.

2. In this study, turbulence-prone days were selected based on automated in-flight turbulence reports using an EDR threshold (EDR $\geq$ 0.2), without distinguishing the primary mechanisms and/or sources of turbulence. Turbulence can be generated by diverse sources, which include but are not limited to clear-air turbulence in the vicinity of jet, trough, tropopause etc. in the absence of clouds; mountain-wave turbulence due to breaking waves; convectively induced turbulence associated with moist instabilities, intense updrafts/downdrafts, outflows etc. This can be especially problematic as relative contributions of turbulence sources may differ among the four flight corridors (i.e., North Atlantic, North Pacific, East Asia, and CONUS), therefore may impact the differences in the TurboMeter outcomes among them. I think it is needed to validate that the selected turbulence days in 2024 and their past and present analogues are associated with the same source of turbulence; for example, if a selected case in 2024 is clear air turbulence, their analogues should be clear-air turbulence days as well. Also, the turbulence index TI1 and its components seem more appropriate for clear-air turbulence than for mountain-wave turbulence or convectively induced turbulence.

Such complexity in the possible source of turbulence is another reason why I don't think the 500-hPa geopotential height alone is a good metric to identify similar synoptic patterns and analogues. Do the analogies have similar synoptic patterns in terms of jets and troughs, background atmospheric conditions for wave breaking, and initiation and development of convection, which all are synoptic conditions for turbulence generation?

We sincerely thank the Reviewer for raising this important point. We fully agree that aviation turbulence can arise from a variety of physical mechanisms, including clear-air turbulence (CAT), mountain-wave turbulence, convectively induced turbulence (CIT), near-cloud turbulence, and gravity-wave breaking. These mechanisms frequently coexist and interact, and their relative importance can vary considerably among the four aviation corridors analysed in

this study. We also agree that this aspect was not sufficiently discussed in the original manuscript.

We would like to clarify, however, that distinguishing the physical mechanism responsible for each individual turbulence encounter is not the objective of the TurboMeter framework. Rather, TurboMeter follows an event-based, circulation-conditioned attribution philosophy, in which objectively observed turbulence events are first identified from EDR measurements and subsequently analysed under comparable large-scale circulation conditions. In this context, EDR provides an objective measure of the observed turbulence intensity, independently of the physical mechanism that generated the encounter.

Accordingly, the analogue-selection procedure is not intended to ensure that the target event and all its historical analogues originate from the same turbulence-generation mechanism. Instead, it constrains the large-scale synoptic circulation and subsequently investigates how the associated atmospheric environment differs between historical and present-day climate conditions. The attribution therefore concerns the evolution of turbulence-favourable atmospheric environments under comparable circulation states rather than the attribution of individual turbulence-generation processes.

We recognize that this conceptual distinction was not sufficiently emphasized in the original manuscript. For this reason, the revised manuscript now explicitly states throughout the Introduction, Methods, Discussion and Conclusions that TurboMeter should not be interpreted as a turbulence-classification framework. Instead, it provides circulation-conditioned attribution of observed turbulence events, irrespective of the specific physical mechanism responsible for each encounter.

To further strengthen the physical interpretation, the revised manuscript also introduces several independent atmospheric diagnostics that help characterize the environments associated with the analogue populations. In addition to the original turbulence diagnostics (TI1, VWS and DEF), we now analyse upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). These additional variables reveal how the thermodynamic and dynamical environments evolve under comparable synoptic circulation conditions and allow us to discuss the possible influence of different physical processes without attempting to classify the dominant turbulence mechanism for each event.

We also agree with the Reviewer that TI1 is primarily designed to diagnose environments favourable to clear-air turbulence and does not explicitly represent all turbulence-generation mechanisms. We now discuss this point explicitly in the revised Discussion and Conclusions, where we identify the extension of the framework toward process-specific attribution—including convectively induced turbulence, mountain-wave turbulence, and other turbulence types—as one of the main directions for future research.

Finally, we emphasize that the objective of the present work is intentionally broader than process-specific turbulence attribution. By conditioning on comparable large-scale circulation states and analysing independent atmospheric diagnostics, TurboMeter quantifies how the environments in which aviation turbulence develops evolve under anthropogenic climate change, regardless of the individual mechanism ultimately responsible for the observed turbulence encounter. We believe that this clarification better reflects both the scope and the limitations of the proposed framework and substantially strengthens the interpretation of the results.

3. Overall, I think the result section needs revisions in terms of plots and texts both.

The plots presented focused on the TurboMeter outputs, including 500-hPa geopotential height, TI1, VWS, and DEF. However, neither the main text nor the figure caption mentioned what vertical levels the TI1, VWS, and DEF plots are for. If those turbulence parameters are presented for a 500-hPa level, I think it needs to be updated to an upper level, for example 300 hPa, which is in typical cursing altitude range and is of great interest in aviation turbulence.

For the main texts, the differences in turbulence between the past and present climates, as well as their difference among the four flight corridors, were mainly explained in terms of the jet dynamics. However, this manuscript never showed how the actual synoptic-scale patterns look like including jets, temperature patterns, etc., to support their arguments. I think the analysis needs to be improved by explicitly showing the differences and similarities in jets, troughs, and other possible sources of turbulence between the past and current climate analogues.

We sincerely thank the Reviewer for this valuable comment. We fully agree that the original manuscript did not provide sufficient physical interpretation of the analogue-based results and that the connection between the turbulence diagnostics and the associated atmospheric environment was not sufficiently demonstrated. We also acknowledge that the description of the turbulence diagnostics could have been clearer, particularly regarding the vertical layer over which TI1, vertical wind shear (VWS), and deformation (DEF) are computed. These points have now been comprehensively addressed in the revised manuscript.

First, we have clarified throughout the manuscript that TI1, VWS and DEF are computed over the upper-tropospheric cruise-altitude layer (300–200 hPa), rather than at the 500-hPa level. The previous wording may have unintentionally created confusion because the analogue-selection procedure is based on Z500, whereas the turbulence diagnostics are evaluated at typical commercial aircraft cruising altitudes. This distinction is now explicitly stated in both the Data and Methods sections.

Second, and more importantly, we have substantially expanded the physical interpretation of the analogue populations by introducing several new diagnostics that were not included in the original manuscript. For each regional case study, the revised manuscript now presents the target event together with the corresponding past- and present-climate analogue composites of the Z500 field. These additional figures provide a direct assessment of the robustness of the analogue-selection procedure, demonstrating that the selected analogue populations successfully reproduce the large-scale circulation of the target event, including the principal trough-ridge configuration and planetary-wave structure.

Building upon this circulation-based validation, we further introduce independent diagnostics that were not used during the analogue-selection procedure, including upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). These variables provide an independent physical characterization of the analogue populations and allow the dynamical and thermodynamical environments to be compared between the historical and present-day climates.

The newly added upper-level wind fields show that the position and large-scale organization of the jet streams remain remarkably consistent between the target events and the corresponding analogue populations, while revealing systematic strengthening of the upper-level winds and vertical wind shear under present-day climate conditions. Likewise, the additional surface and thermodynamic diagnostics demonstrate coherent increases in atmospheric moisture and surface temperature across all regions, together with region-dependent changes in convective

instability. These independent diagnostics therefore provide a much more comprehensive physical interpretation of the analogue populations than was possible in the original manuscript.

The Results section has consequently been substantially reorganized. Rather than interpreting the changes exclusively through TI1, VWS and DEF, each regional case study now combines circulation, turbulence, surface and thermodynamic diagnostics, allowing the physical pathways associated with climate-related changes in turbulence-favourable environments to be discussed in considerably greater detail. This expanded interpretation is further synthesized in the revised Discussion, where the common and region-specific mechanisms emerging across the four aviation corridors are explicitly compared.

We believe that these additions directly address the Reviewer's concerns. The revised manuscript now demonstrates not only that the analogue populations are dynamically consistent with the target events, but also that independent atmospheric diagnostics provide a coherent physical interpretation of the climate-related differences identified by the circulation-conditioned attribution framework.

Reviewer #2

We sincerely thank the Reviewer for the careful, thorough and constructive evaluation of our manuscript, as well as for the many valuable comments and literature suggestions. We appreciate the recognition that the topic addressed by this study is timely and scientifically important. We also recognize that many of the concerns raised by the Reviewer relate to the interpretation and presentation of the methodology rather than to the analogue-based framework itself.

We acknowledge that the original manuscript did not always clearly distinguish between circulation-conditioned attribution and attribution of individual turbulence-generation mechanisms, which may have led to an overinterpretation of the scope of the proposed framework. We have therefore substantially revised the manuscript to clarify the methodological philosophy of TurboMeter, strengthen the physical interpretation of the results, and explicitly discuss both the capabilities and limitations of the approach.

In particular, the revised manuscript:

- **adopts more precise terminology throughout, avoiding wording that could imply attribution of individual turbulence-generation mechanisms;**
- **clarifies that the analogue-selection procedure conditions only the large-scale circulation through the Z500 field, while additional atmospheric diagnostics are analysed independently to characterize the analogue populations;**
- **introduces several new dynamical and thermodynamical diagnostics, including upper-level wind speed, SST, T2m, TCWV, CAPE and CIN;**
- **substantially expands the Discussion and Conclusions to explicitly address the limitations associated with ERA5, the coexistence of different turbulence-generation mechanisms, and possible future developments of the framework.**

We believe that these revisions considerably strengthen the manuscript while preserving the original scientific philosophy of TurboMeter as a circulation-conditioned attribution framework.

1. Misuse of EDR ≥ 0.2 as “extreme turbulence”

The manuscript defines the event sample using automated in situ aircraft reports with **EDR ≥ 0.2 $m^{2/3}$ s^{-1}**, describing these as moderate-or-greater events and using them to frame “extreme aviation turbulence” in the title and throughout the paper. However, **EDR values higher than 0.2 from in situ AMDAR/ACARS aircraft data cannot automatically be linked to, or labeled as, “extreme” turbulence**. The authors should revise the title and terminology and should more carefully follow previous studies on EDR reporting and intensity classification. The current use of EDR ≥ 0.2 as the basis for “extreme turbulence” is misleading and needs substantial correction.

We sincerely thank the Reviewer for raising this important point. We fully agree that the terminology adopted in the original manuscript could inadvertently suggest that EDR ≥ 0.2 m^2 s^{-3} represents "extreme turbulence" in the operational aviation sense. This was not our intention. Our objective was simply to identify objectively observed turbulence events belonging to the moderate-or-greater (MOG) category using automated EDR measurements, following the operational thresholds commonly adopted in the aviation literature.

To avoid any possible ambiguity, we have carefully revised the terminology throughout the manuscript. In particular, we have removed the expression "extreme turbulence" wherever it could be interpreted as an operational intensity classification and consistently replaced it with

"moderate-or-greater (MOG) aviation turbulence", "aviation turbulence", or "turbulence events", depending on the context. The manuscript now consistently refers to EDR ≥ 0.2 $m^2 s^{-3}$ as the threshold used to define the analysed event sample rather than as a definition of extreme turbulence.

We have also expanded the description of the EDR dataset in the Data section by introducing the operational interpretation of the Eddy Dissipation Rate (EDR), its advantages as an objective aircraft-independent turbulence metric, and the rationale for selecting the MOG threshold adopted in this study. In addition, we incorporated the literature suggested by the Reviewer regarding EDR reporting and turbulence-intensity classification.

We believe that these revisions remove the ambiguity identified by the Reviewer while preserving the scientific objective of the study, namely the attribution of observed moderate-or-greater aviation turbulence events under comparable large-scale circulation conditions.

2. Synoptic analogues cannot guarantee the same turbulence-generation processes

A second major concern is the assumption that **similar 500 hPa geopotential-height analogues can meaningfully represent the full downscaling pathway from synoptic flow to small-scale turbulence affecting cruising aircraft**. The method identifies historical analogues using Z500 patterns and compares turbulence diagnostics between past and present periods.

This approach may be useful for broad circulation context, but it does not guarantee the same mesoscale or microscale processes responsible for aircraft-scale turbulence. Numerous high-resolution modeling studies have shown that severe aviation turbulence depends on localized mechanisms such as gravity-wave breaking, jet-front interactions, near-cloud processes, and mesoscale convective influences. The authors should substantially weaken or reframe their attribution claims and more fully consider previous high-resolution modeling studies from both U.S. and Korean scientists.

We sincerely thank the Reviewer for this important comment. We fully agree that similarity in the large-scale circulation does not imply identical mesoscale or microscale processes responsible for individual turbulence encounters. Indeed, we emphasize that reproducing the complete dynamical pathway from synoptic-scale circulation to aircraft-scale turbulence is not the objective of the TurboMeter framework.

TurboMeter follows a circulation-conditioned analogue philosophy, analogous to that adopted in ClimaMeter and related analogue-based attribution studies. Within this framework, the 500-hPa geopotential height field is used exclusively to constrain the large-scale atmospheric circulation, thereby allowing atmospheric environments evolving under comparable synoptic conditions to be compared across different climate periods. Consequently, the attribution concerns changes in the atmospheric environments favourable to aviation turbulence rather than the attribution of individual turbulence-generation mechanisms.

We recognize that this conceptual distinction was not sufficiently emphasized in the original manuscript. Accordingly, the revised manuscript now explicitly clarifies throughout the Introduction, Methods, Discussion, and Conclusions that the analogue-selection procedure is not intended to reproduce identical turbulence-generation processes, but rather to isolate the influence of anthropogenic climate change under comparable large-scale circulation conditions.

To further strengthen this interpretation, the revised manuscript now includes composite Z500 fields for the target events together with the corresponding past- and present-climate analogue populations. These new analyses demonstrate that the analogue populations successfully reproduce the large-scale synoptic configuration of the target events, providing direct evidence of the robustness of the analogue-selection procedure.

In addition, we have substantially expanded the physical characterization of the analogue populations by introducing several atmospheric diagnostics that are completely independent of the analogue-selection procedure, including upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). These additional diagnostics reveal physically consistent differences between the historical and present-day analogue populations while preserving comparable synoptic-scale circulation patterns.

We have also revised the Discussion and Conclusions to explicitly acknowledge that ERA5 cannot resolve many of the mesoscale and microscale processes responsible for aircraft-scale turbulence, including gravity-wave breaking, convectively induced turbulence, and mountain-wave turbulence. Consequently, TurboMeter should not be interpreted as attributing individual turbulence-generation mechanisms, but rather as quantifying how turbulence-favourable atmospheric environments evolve under comparable synoptic-scale circulation conditions.

We believe that these revisions substantially clarify both the scope and the limitations of the methodology, while preserving the original scientific philosophy of the analogue-based attribution framework.

3. Convective systems are not sufficiently considered

Third, the manuscript treats many events primarily as clear-air/jet-related turbulence events, but **many extreme or severe turbulence encounters are directly or indirectly associated with convective systems**, and the separation between natural variability and anthropogenic forcing is not straightforward in such cases. The manuscript itself acknowledges that its framework could be extended in the future to convectively induced turbulence, mountain-wave turbulence, and other turbulence types, which underscores a key limitation of the current analysis. The authors should not present broad attribution conclusions without explicitly addressing convective contamination, near-cloud turbulence, convective outflow, and convectively induced gravity-wave processes.

We sincerely thank the Reviewer for raising this important point. We fully agree that aviation turbulence can originate from a broad range of physical mechanisms and that convective processes—including convectively induced turbulence, near-cloud turbulence, convective outflows, and convectively generated gravity waves—may contribute significantly to turbulence encountered by commercial aircraft. We also agree that this aspect deserved a more explicit discussion in the original manuscript.

Our intention, however, was not to interpret the analysed events exclusively as clear-air turbulence events, nor to attribute the observed turbulence to a single physical mechanism. Rather, TurboMeter was intentionally developed as a circulation-conditioned attribution framework in which objectively observed turbulence events are analysed irrespective of the specific process responsible for the aircraft encounter. Consequently, the framework does not require the selected events to be exclusively associated with clear-air turbulence, nor does it attempt to separate convective, mountain-wave, or other turbulence-generation mechanisms.

To make this distinction clearer, the revised manuscript now explicitly states that the EDR observations used for event selection provide an objective measure of turbulence intensity without identifying its physical origin. The attribution therefore concerns the evolution of the atmospheric environments in which turbulence develops under comparable large-scale circulation conditions, rather than the attribution of individual turbulence-generation mechanisms.

In addition, one of the principal improvements introduced in the revised manuscript is the inclusion of independent thermodynamic diagnostics, including total column water vapour

(TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). Although these variables are not intended to classify convective turbulence, they provide valuable information on the thermodynamic environments associated with the analogue populations and allow the possible influence of convectively favourable conditions to be discussed where appropriate. In particular, the revised Results and Discussion now show that the thermodynamic response varies considerably among the analysed regions, with continental regions exhibiting stronger changes in atmospheric instability than the oceanic flight corridors.

We have also substantially revised the Discussion and Conclusions to explicitly acknowledge that convective processes represent one of several possible contributors to aviation turbulence and that the current ERA5-based framework cannot explicitly resolve many of the mesoscale and microscale processes associated with deep convection. Accordingly, the revised manuscript avoids attributing the observed climate signal to individual turbulence-generation mechanisms and instead consistently interprets the results in terms of changes in turbulence-favourable atmospheric environments.

We believe that these revisions substantially clarify both the scope and the limitations of the proposed methodology while providing a more balanced discussion of the potential role of convective processes in aviation turbulence.

4. Mountain-wave turbulence is insufficiently addressed

Fourth, the treatment of **mountain-wave turbulence** is inadequate. Mountain-wave turbulence is an important source of upper-tropospheric and lower-stratospheric turbulence, especially over the western United States and other mountainous regions. Such processes cannot be fully captured by ERA5 fields and synoptic analogue matching alone. The CONUS case, for example, includes many reports over North America, but the manuscript does not sufficiently distinguish jet-related CAT from mountain-wave-related turbulence or other terrain-induced processes . The authors should explicitly discuss the limits of ERA5 and Z500-based analogue methods for diagnosing mountain-wave turbulence and should incorporate prior modeling and forecasting studies on this process.

We sincerely thank the Reviewer for highlighting the important role of mountain-wave turbulence, particularly over regions characterized by complex orography such as western North America. We fully agree that mountain-wave turbulence represents an important contributor to aviation turbulence and that the original manuscript did not sufficiently discuss the limitations associated with diagnosing this process using an ERA5-based analogue framework.

We would like to emphasize, however, that the objective of TurboMeter is fundamentally different from identifying the physical mechanism responsible for individual turbulence encounters. The framework conditions the large-scale atmospheric circulation using Z500 analogues and subsequently evaluates how turbulence-favourable atmospheric environments differ between historical and present-day climate conditions. Consequently, it is not intended to distinguish whether an individual turbulence report originated from clear-air turbulence, mountain-wave turbulence, convectively induced turbulence, or other interacting processes.

We fully agree that mountain-wave turbulence involves mesoscale and microscale dynamical processes that cannot be explicitly resolved by ERA5. Likewise, analogue selection based on large-scale circulation cannot reproduce the detailed wave-breaking dynamics associated with individual mountain-wave events. We have therefore substantially expanded the Discussion and Conclusions to explicitly acknowledge this limitation and to clarify the scope of the methodology.

At the same time, it is worth noting that large-scale circulation remains an important controlling factor for mountain-wave activity, since synoptic-scale flow determines the strength and orientation of the airflow impinging on major mountain ranges. In this sense, the analogue framework captures the large-scale atmospheric environments within which mountain-wave turbulence may develop, while recognizing that the detailed wave dynamics themselves require much higher spatial resolution than that provided by ERA5.

For this reason, the revised manuscript now consistently interprets the attribution results as describing changes in turbulence-favourable atmospheric environments rather than changes in individual turbulence-generation mechanisms. We also identify the integration of higher-resolution numerical simulations and process-specific turbulence diagnostics as one of the principal future developments of the TurboMeter framework.

We believe that these revisions provide a more balanced discussion of the capabilities and limitations of the proposed methodology while explicitly recognizing the important role of mountain-wave turbulence in aviation meteorology.

5. East Asia and Pacific turbulence mechanisms are oversimplified

Fifth, the interpretation of East Asia and the Pacific is too simplified. The manuscript links East Asian turbulence mainly to jet-stream intensification and deformation/shear along the East Asian jet . However, in East Asia and the western Pacific, high turbulence-index values in ERA5 are not simply related to the jet stream alone. They often result from complicated combinations of typhoons, tropical cyclone outflow, upper-level frontogenesis, vertical shear, deformation, and convective or near-cloud processes. Therefore, the manuscript should avoid attributing high TI, VWS, and DEF values only to jet-stream changes and should consider previous regional studies on East Asian and western Pacific turbulence mechanisms.

We sincerely thank the Reviewer for this insightful comment. We fully agree that aviation turbulence over East Asia and the western Pacific results from a complex interplay of dynamical and thermodynamical processes and that the original manuscript placed excessive emphasis on jet-stream variability alone. Our intention was not to imply that the observed turbulence environments are controlled exclusively by jet dynamics, but rather to describe the dominant large-scale circulation features emerging from the analogue analysis. We recognize that this distinction was not sufficiently clear in the original version.

In response to the Reviewer's suggestion, we have substantially revised both the Results and the Discussion for the East Asia case study. Rather than interpreting the observed changes exclusively in terms of jet intensification, the revised manuscript now presents a broader characterization of the atmospheric environment by combining circulation, upper-level dynamics, surface conditions, atmospheric moisture, and thermodynamic diagnostics. In particular, the inclusion of upper-level wind speed, total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN) demonstrates that the climate signal over East Asia cannot be interpreted solely as a strengthening of the jet stream. Instead, it reflects concurrent modifications of the dynamical and thermodynamical environment in which aviation turbulence develops.

Accordingly, the revised Discussion now explicitly acknowledges that aviation turbulence over East Asia and the western Pacific may arise from multiple interacting processes, including upper-level jet dynamics, frontal systems, tropical-cyclone outflow, convective activity, gravity waves, and near-cloud processes. The revised manuscript therefore avoids attributing the observed changes exclusively to jet-stream variability and instead consistently interprets the results as describing circulation-conditioned atmospheric environments that may favour different turbulence-generation mechanisms.

We have also incorporated the additional literature suggested by the Reviewer and revised the discussion to better reflect the complexity of the East Asian upper-tropospheric environment. These modifications provide a more balanced interpretation of the regional results while remaining consistent with the scope of the analogue-based attribution framework.

We believe that these revisions substantially strengthen the physical interpretation of the East Asia case study and avoid possible overinterpretation of the role of individual dynamical processes.

6. Main methodological limitation: Z500 cannot represent the full turbulence pathway

The central methodological issue is that **500 hPa geopotential-height fields cannot represent the full synoptic-to-mesoscale story linking atmospheric circulation to turbulence reports**. The paper's framework may identify broad circulation analogues, but it cannot by itself establish the physical pathway to aircraft-scale turbulence. This limitation is especially important because the paper makes strong claims that approximately 68% of events were mainly strengthened by anthropogenic forcing . These claims require much stronger validation against observed EDR intensity, turbulence type, convective proximity, cloud environment, mountain-wave environments, and high-resolution model diagnostics.

We sincerely thank the Reviewer for this thoughtful comment, which highlights what is arguably the most important conceptual aspect of the proposed methodology. We agree that the 500-hPa geopotential height field alone cannot represent the complete multiscale dynamical pathway linking the large-scale atmospheric circulation to aircraft-scale turbulence. However, we believe that this observation also reflects an important distinction between the objective of TurboMeter and that of process-based turbulence modelling studies.

TurboMeter was not developed to reconstruct the complete chain of physical processes leading from synoptic-scale circulation to individual turbulence encounters. Rather, following the circulation-conditioned attribution philosophy established in ClimaMeter and related analogue-based attribution studies, the objective is to constrain the large-scale atmospheric circulation and subsequently investigate how the associated atmospheric environments evolve under different climate conditions. Within this framework, the Z500 field serves as the conditioning variable that defines comparable synoptic-scale circulation states, while the physical interpretation of the analogue populations is intentionally based on additional atmospheric diagnostics that are completely independent of the analogue-selection procedure.

One of the principal improvements introduced in the revised manuscript is precisely the substantial expansion of these independent diagnostics. In addition to the original turbulence indices, the revised manuscript now includes composite Z500 fields for the target events and the corresponding analogue populations, upper-level wind speed, sea surface temperature (SST), near-surface air temperature (T2m), total column water vapour (TCWV), Convective Available Potential Energy (CAPE), and Convective Inhibition (CIN). Together, these analyses demonstrate that the analogue populations reproduce remarkably similar synoptic-scale circulation patterns while simultaneously revealing systematic climate-related modifications of the associated atmospheric environments.

We have also substantially revised the Discussion and Conclusions to explicitly acknowledge that ERA5 cannot resolve many of the mesoscale and microscale processes responsible for aircraft-scale turbulence, including gravity-wave breaking, mountain-wave turbulence, convectively induced turbulence, and near-cloud processes. Consequently, the revised manuscript consistently avoids interpreting the attribution results as identifying individual turbulence-generation mechanisms. Instead, TurboMeter is explicitly presented as a circulation-

conditioned attribution framework that quantifies how atmospheric environments favourable to aviation turbulence evolve under comparable synoptic-scale circulation conditions.

We therefore respectfully emphasize that the capability of Z500 to represent the complete turbulence pathway is not an assumption of the methodology. On the contrary, the revised manuscript explicitly separates the two components of the framework: (i) the conditioning of the large-scale circulation through analogue selection using Z500, and (ii) the independent physical characterization of the resulting analogue populations through complementary dynamical and thermodynamical diagnostics. We believe that this clarification significantly strengthens both the methodological presentation and the interpretation of the results while remaining fully consistent with the scope of the analogue-based attribution framework.