

EDITORIAL

Reviews and Responses for **Vertical Efficiency Analysis of Continuous Climb and Descent Operations in Canadian Terminal Areas**

Authors: Guilherme Trindade Tolentino Bernardo, Marcelo Xavier Guterres, and Alexandre de Barros

Reviewers: Tatiana Polishchuk, Raúl Sáez, and Ryota Mori

Editor: Xavier Olive

1. Original paper

The DOI for the original paper is <https://doi.org/10.59490/joas.2026.8470>

2. Review - round 1

2.1 Reviewer 1 – Tatiana Polishchuk

Page 1: Abstract. I suggest removing the airport ICAO codes from the abstract to improve readability.

Page 2: Introduction. Line 40: The authors may want to cite more recent works on evaluation of vertical flight performance using open data.

Line 45: How are optimal vertical profiles modeled in this work? If not applicable, why is it claimed in the end of the Introduction, where contributions are typically highlighted?

Methodology section. Line 56: Use full OSN name, introduce abbreviation if you want to use it further on, and provide corresponding references.

Line 68: “The trajectories of flights that either departed from or arrived at one of the ten busiest airports” – how were the data related to descent and climb identified?

Page 3: Figure 1 and surrounding text. Figure 1 caption: Specify more details about the trajectories, such as that they belong to Canadian airspace, and mention the period of time.

Line 74: “flights on more frequently traveled routes” – how are these defined?

Lines 90–92: Paragraph should not consist of a single sentence only. Since its content is strongly related to the previous paragraph, I suggest merging them together.

Line 94: “calculation of KPIs a given TMAs” – a preposition is missing. Should it be “of” or “for”?

Page 4: Figures and TMA definition. Figure 2 caption and all figures: Replace ICAO codes with airport names wherever possible to improve readability. Figures should be readable without referring to the text.

Line 105: The first sentence belongs better to the previous paragraph by meaning, while the second one can start the next paragraph.

Page 5: Data filtering. Too many flight trajectories were excluded from the data analysis, which risks making the analysis incomplete or incorrect. OSN data is known to be noisy and incomplete, but multiple preprocessing techniques exist (moving median smoothing, interpolation for vertical anomalies, Gaussian approximation for lateral deviations, etc.).

Line 136: Add the reference to GANP guidelines already here when first mentioned.

The KPIs used have recognized ICAO numbers (KPI17, etc.). You may want to mention them in the text.

Equations (1) and (2) and their corresponding variable descriptions are absolutely equivalent. There is no reason for repetition.

Page 7: Results section. Table 1: Use the unique KPI ICAO numbers instead of or in addition to specifying their names in parentheses.

Figure 5 seems to be missing, with only one number used instead.

Line 194–198: Link airport names and ICAO codes together in parentheses to improve readability of the text and Figure 6.

Figure 6: Increase font sizes. The title below the table is redundant; all information should be in the figure caption.

Page 8: Textual issues and comparisons. Line 207: “farther out” → “further out”.

Line 208: “the framework allows visualization”. The two KPIs do not constitute a framework. Replace with “performance analysis” or similar.

Line 210: “From left to right each line represents a flight”. The authors likely refer to curves, not lines. Also, from left to right on the x-axis we have distances, not flights, which is confusing.

Line 216: Comparing the same calendar days in April and January is problematic. Comparing same days of the week with similar traffic patterns, or the busiest days of the months, would be more appropriate for fair comparison.

Line 222: Heightened thunderstorms are suggested as the reason for increased KPI values. Are thunderstorms common in April in Canada? Check weather statistics from that year to support the hypothesis.

Line 226: Stable altitudes for level-offs may be attributed to operational inefficiencies, as suggested in Pasutto *et al.* (2021).

Line 231: “the framework generates boxplots” – not the framework, but the authors generate the plots.

Line 239: “which demonstrate” → “demonstrates” (subject-verb agreement).

The authors refer to ICAO's vertical flight efficiency standards. Which exact standards are mentioned? Please provide references or further explanations.

Page 10: Figure quality and grammar. Figure 11: Increase font sizes and figure sizes. In Figure 11b there is excessive empty space; reducing the x-axis scale and centering the figure would improve presentation.

Add KPI ICAO codes to figures to make them more readable.

Line 249: The sentence starting with "Particularly, weather factors..." is grammatically incomplete.

Line 255: High traffic volumes are mentioned. Please provide the numbers.

Discussion section. It is known that weather and traffic intensities are the main reasons for inefficiencies in TMAs. You may also want to analyze what procedures are used at the corresponding airports. For example, trombone and point merge arrivals feature stable vertical constraints at specific levels.

Page 11: Clarity and grammar. Line 280: "Fragmented descent profile" – what does this mean? Please clarify.

Line 281: "From right to left" – see my comment to Line 210 with similar content.

Line 283: "Warmer colors represents" → "Warmer colors represent" (subject-verb agreement).

2.2 Reviewer 2 – Raúl Sáez

Comment 1: Clarification of TMA definition (Line 105). In line 105, the authors wrote "This allowed us to determine whether an aircraft was operating within a TMA of 200 NM at the destination airport, for example". What do the authors mean by that statement? Are the authors assuming the TMAs in the Canadian airports are circles of a 200 NM radius? I recommend the authors clarify this statement, as it seems a very strong generalization for all airports.

Comment 2: Percentage of excluded flights and justification. In line 126, the authors explain that they excluded several flights from the dataset. Could the authors clarify what percentage of flights were excluded? The reasons given to exclude those flights do not seem sufficient to justify their exclusion. For instance, trajectories having "repeated waypoints over advancing timestamps" could be preprocessed to remove these waypoints, ending up with trajectories suitable for the efficiency study. If the percentage of flights exhibiting these issues is negligible, removal might be acceptable, but otherwise some preprocessing should be applied. I saw in Line 186 that 76% of departure flights are classified as complete. Does this mean that 24% of flights were excluded for the stated reasons? If so, the number of removed flights seems too high, and some preprocessing might be needed to include more flights.

Comment 3: Effect of US border proximity on TMA definition. As described in Line 149, the TMA radius considered is 200 NM. What happens for airports close to the US border and flights departing/arriving from that direction? I understand Canadian TMAs do not extend into US airspace, right? Does this affect the results? Is this accounted for by the authors?

Comment 4: Effect of published procedures on level-offs. Could the published departure/arrival procedures influence the level-offs observed during climb/descent? If procedures are more restrictive, aircraft might need to level off. So it may not only be ATC instructions or weather causing level-offs, but also strategic constraints in the published charts. I recommend checking published charts and discussing this.

Comment 5: Canadian airspace particularities and airport characteristics. The assessment of climb and descent performance via ICAO indicators is clear, but the authors should explain Canadian airspace particularities and the characteristics of the major airports studied. For instance, published procedures, coordination with US airspace, etc. A thorough analysis would strengthen the contribution.

Comment 6: Typographical corrections. Typos:

- Line 36: “aircrafts” → aircraft.
- Line 186–187: “After data filtering, approximately 76% of departure flights at the ten airports in Canada are classified as complete, while 85.3% of arriving flights.” → ...complete, compared with 85.3% of arriving flights.

2.3 Reviewer 3 – Ryota Mori

Comment 1: Data mining techniques in abstract. The abstract states “By applying data mining techniques...”. However, no specific data mining techniques are mentioned or described in the manuscript. The statement “the analysis reveals that vertical efficiency is strongly influenced by traffic volume, geography, and seasonality” appears speculative at this stage, as the analysis does not provide sufficient evidence to support these claims.

Comment 2: Seasonality analysis. To assess seasonality, the authors should analyze data over a full year, not just 6 months. Although traffic volume data can be easily obtained, they are not presented in the manuscript. Is seasonality defined as the temporal distribution of traffic across the year, or by another metric? At present, “seasonality” and “traffic volume” appear to be used interchangeably.

Comment 3: Geographic analysis. Is any analysis of geographic factors included in this paper?

Comment 4: Table 1 terminology clarification. Table 1: The meaning of each metric is ambiguous. Do “filtered flights” refer to the flights analyzed in the research? Are “complete flights” distinct from “filtered flights”? What do “KPI flights (level-off climb)” mean? Flights that include pre-defined level-off segments among departure flights?

Comment 5: Quantitative support for procedure/sector claims. Line 205–206: The authors claim that “suggesting no concentration in a specific procedure or sector of the TMA”, but is this conclusion supported by quantitative analysis? Please provide the ratio of level-off flights per procedure, rather than relying solely on visual representation of the figure.

Comment 6: Weather data and supporting evidence. Line 221: The authors discuss weather-related factors and wind patterns, but no supporting data are presented. If they are considered, the relevant data must be included and analyzed.

Comment 7: SID structure and mean altitude of level-offs. Fig. 10: The authors discuss mean altitude of level-off, which may be affected by SID structure at each airport. This factor should be acknowledged and discussed in the context of the findings.

Comment 8: Vertical efficiency vs. inefficiency labeling. Fig. 12: The vertical axis is labeled “vertical efficiency during descent”, which implies that higher values are better. However, lower values actually indicate better efficiency. Using “vertical inefficiency” would be more appropriate. This correction should also be applied to other figures.

Overall work scope and methodology positioning. The conclusions drawn by the authors are not sufficiently supported by the current analysis. To strengthen the validity and impact of the findings, a more in-depth analysis with expanded data and rigorous statistical validation is strongly recommended.

3. Response - round 1

3.1 Response to Reviewer 1 – Tatiana Polishchuk

We sincerely thank the reviewer for the detailed and constructive feedback. The suggestions have significantly improved the clarity and scientific rigor of the manuscript. Below, we provide responses organized by page and section.

Abstract – ICAO codes.

Response

We have revised the abstract to remove ICAO codes, presenting the airports by their common names instead. This improves accessibility for a broader audience while maintaining clarity.

Introduction – recent open-data references.

Response

We have expanded the literature review section with citations of recent works on vertical flight performance assessment using open-source data, including recent publications addressing vertical efficiency analysis from operational datasets.

Introduction – optimal vertical profiles claim.

Response

We have clarified this point. The paper does not model optimal vertical profiles; rather, it evaluates the actual vertical profiles flown by aircraft using standardized ICAO KPIs. We describe observed flight profiles and their efficiency characteristics relative to continuous climb and descent operations, not theoretical optimal profiles. We have reframed the Introduction to avoid confusion and now clearly state

that the contribution focuses on assessing operational outcomes (inefficiencies in climb and descent, the occurrence of level-offs) rather than optimizing trajectories.

Methodology – OSN abbreviation and references.

Response

We have revised the text to introduce OpenSky Network by its full name on first mention, provide the corresponding reference, and then use the abbreviation “OSN” subsequently.

Identification of climb and descent phases.

Response

We have clarified that the descent and climb phases are identified using the arrival and departure tags from the OpenSky Network database. Flights are separated into two categories based on these labels: departure trajectories (climb phase) and arrival trajectories (descent phase). This identification is performed automatically by the OSN database and ensures consistent labeling of operational phases.

Figure 1 caption – airspace and period.

Response

We have expanded the Figure 1 caption to specify that the trajectories are from Canadian airspace and include the temporal coverage of the dataset.

“Frequently traveled routes” definition.

Response

We have clarified that more frequently traveled routes are identified based on the origin and destination airport labels in the OpenSky Network database. The dataset is organized by the frequency of flight counts between airport pairs, allowing us to focus on routes with the highest traffic volumes.

Single-sentence paragraph (Lines 90–92).

Response

We have merged this single-sentence paragraph with the preceding paragraph to improve text flow and readability.

Missing preposition (Line 94).

Response

We have corrected this to read “calculation of KPIs for given TMAs.”

Replacing ICAO codes with airport names in figures.

Response

We appreciate this suggestion. We have made efforts to improve figure clarity, including adding airport names alongside ICAO codes in key figures. However, we have retained ICAO codes as the primary identifier because: (1) they are standardized globally and ensure consistency for international comparisons; (2) the small figure size in some cases makes using full airport names impractical for legend and label sizing; and (3) ICAO codes are the standard notation in aviation research and familiar to the target audience. We have balanced this by including airport name references in figure captions.

Restructuring Line 105.**Response**

We have restructured this content, moving the first sentence to the preceding paragraph and using the second sentence to introduce the next section for improved logical flow.

Data filtering – exclusion percentages and justification.**Response**

We have expanded the discussion of data filtering with explicit percentages and stronger justification. The excluded flights primarily correspond to trajectories that fundamentally cannot be used for terminal-area vertical efficiency assessment (e.g., overflights containing only cruise segments, trajectories not crossing the 200 NM terminal boundary) rather than correctable data quality issues.

In total, 24.0% of departure trajectories and 14.7% of arrival trajectories were excluded. Despite these exclusions, the final dataset comprises more than 170,000 analyzed flights, significantly larger than typical studies reported in the literature (which often rely on one week of operations). Furthermore, not all complete flights are used in KPI calculations: for example, flights whose entire departure trajectory remains below 3,000 ft AGL are excluded, as this is the minimum altitude threshold specified in GANP guidelines.

Rather than applying preprocessing techniques to trajectories that lack the necessary terminal segments for valid efficiency assessment, our filtering approach ensures methodological consistency and transparent KPI computation.

GANP reference at first mention (Line 136).**Response**

We have added the GANP reference at the first mention in this section.

ICAO KPI numbers in text.**Response**

We have updated the manuscript to reference the specific ICAO KPI numbers (KPI17, KPI19) in addition to their descriptive names throughout the text and figures.

Redundant equations.

Response

We have removed the redundant equation and retained only one equation with its complete description.

Table 1 – KPI ICAO numbers.

Response

We have updated Table 1 to include KPI ICAO numbers (KPI17, KPI19) alongside the descriptive names for consistency with the methodology section.

Figure numbering.

Response

We have corrected the figure numbering to ensure all figures are properly referenced and numbered sequentially.

Linking airport names and ICAO codes (Lines 194–198).

Response

We have updated the text to consistently present airport names with their corresponding ICAO codes on first mention (e.g., “Toronto Pearson (CYYZ)”), improving readability and cross-referencing with figures.

Figure 6 font sizes and redundant title.

Response

We have increased the font sizes in Figure 6 for improved legibility and moved all relevant information to the figure caption, removing the redundant title below the table.

Grammar/style corrections (Lines 207, 208, 210, 231, 239, 283).

Response

We have corrected “farther out” to “further out”; replaced “framework” with “performance analysis” or “the analysis”; clarified that each curve represents the vertical profile of a single flight, with distance (x-axis) and altitude (y-axis); fixed subject–verb agreement (“demonstrate” → “demonstrates”; “represents” → “represent”); and tightened wording where the term “framework” was used loosely.

Comparison of same calendar days (Line 216).

Response

We have clarified that this comparison is primarily intended for visualization purposes to demonstrate the tool’s capability to generate trajectory plots, rather than to draw substantive conclusions about traffic pattern differences. The visualization serves to illustrate how the analysis plots trajectories and identifies differences in operational patterns. For rigorous traffic pattern analysis, comparing same days of the week or busiest days across months would indeed be more appropriate, which may be addressed

in future work.

Thunderstorms hypothesis (Line 222).

Response

We have revised this section to clarify that while meteorological factors may influence vertical efficiency indicators, the paper does not establish a direct causal relationship between specific weather events and KPI values. Many factors can affect KPI values beyond weather (ATC procedures, traffic intensity, aircraft type mix), and our analysis focuses on describing observed efficiency outcomes rather than isolating individual causal factors. A comprehensive weather impact analysis is identified as a direction for future research.

Operational inefficiencies and Pasutto et al. (2021).

Response

We have added a reference to Pasutto, Zeghal, and Hoffman (2021) on flight inefficiency in descent and have incorporated their perspective into our discussion of level-off patterns.

ICAO standards references.

Response

We have added explicit references to ICAO's standards, particularly the GANP (Global Air Navigation Plan) guidelines and the specific KPI documentation that defines the vertical efficiency metrics (KPI17, KPI19) used in this study.

Figure 11 – font/figure sizes and layout.

Response

We have made improvements to Figure 11 to enhance readability, including increasing font sizes and adjusting the layout. Some limitations remain due to the visualization libraries used; we have worked to optimize the balance between information density and visual clarity within those constraints.

ICAO KPI codes in figures.

Response

We have updated figures to include KPI ICAO codes (e.g., KPI17, KPI19) in addition to descriptive names where space permits.

Incomplete sentence (Line 249).

Response

We have corrected this grammatical issue to form a complete sentence.

Traffic volumes – specific numbers (Line 255).

Response

We have added specific numbers and statistics regarding traffic volumes in the discussion to support the claims made about traffic intensity at the studied airports.

Discussion – procedures (trombone, point merge).

Response

We have incorporated discussion of the role of weather and traffic intensity as significant factors influencing terminal area efficiency. Detailed procedural analysis (e.g., trombone, point merge) is beyond the scope of the current work, which focuses on performance assessment rather than causal attribution to specific procedural constraints. As discussed in our response to Reviewer Sáez’s Comment 4, observed efficiency metrics reflect the combined outcome of all operational factors, including procedural design. A detailed procedure-level analysis is identified as a direction for future research.

“Fragmented descent profile” (Line 280).

Response

We have clarified this terminology. By “fragmented descent profile,” we refer to descent trajectories that are not continuous, featuring multiple level-offs and altitude changes rather than a smooth descent from cruise altitude to landing. We now use clearer language such as “non-continuous descent profile” or “interrupted descent profile” to avoid ambiguity.

Direction of reading (Line 281).

Response

We have clarified the directional reference to avoid confusion with axis interpretation and data representation.

3.2 Response to Reviewer 2 – Raúl Sáez

We sincerely thank the reviewer for the thorough and constructive feedback. We have carefully addressed all comments and believe the revisions have significantly strengthened the manuscript.

Comment 1: Clarification of TMA definition.

Response

We have revised the manuscript to explicitly explain that the 200 NM terminal radius is not an assumption about actual TMA boundaries, but rather a standardized spatial reference defined by ICAO GANP guidelines for vertical efficiency KPI computation.

The revised text now reads (Lines 117–122): *“This allowed us to determine whether an aircraft was operating within a TMA of 200 NM at the destination airport, for example. According to the GANP guidelines, KPIs related to climb and descent profiles should be computed using standardized terminal areas defined by 200 NM radius circles, ensuring that all terminals have the same spatial extent so that the time spent inside the terminal airspace is comparable and not biased by differences in terminal size.”*

This standardized approach ensures comparability across airports with different actual TMA configurations and is consistent with ICAO's recommended methodology for vertical efficiency assessment.

Comment 2: Percentage of excluded flights and justification.

Response

We have explicitly stated the exclusion percentages and provided stronger justification for the filtering criteria. The revised manuscript now includes (Lines 218–229): *“The excluded flights mainly correspond to incomplete or inconsistent trajectories, including cases where the trajectory does not cross the 200 NM terminal boundary (e.g., trajectories containing only cruise segments) or presents temporal inconsistencies that would compromise KPI calculation. In total, 24.0% of departure trajectories and 14.7% of arrival trajectories were excluded. Despite these exclusions, the final dataset still comprises more than 170,000 analyzed flights, which is significantly larger than what is typically reported in the literature, where studies often rely on much shorter time windows (e.g., one week of operations). Therefore, the exclusions do not compromise the robustness or representativeness of the results, while ensuring methodological consistency and transparency in KPI computation. Note that not all complete flights are used in KPI calculation. For example, flights whose entire departure trajectory remains below 3,000 ft AGL are not considered for analysis, as this is the minimum altitude threshold based on GANP guidelines.”*

The exclusions are primarily due to flights that inherently cannot be used for terminal-area vertical efficiency assessment (e.g., overflights, missing terminal segments), rather than data quality issues that could be addressed through preprocessing.

Comment 3: Effect of US border proximity on TMA definition.

Response

We have added the following explanation to the manuscript (Lines 123–131): *“The use of a fixed 200 NM terminal radius does not introduce bias for airports located near the US border, as the KPI is defined relative to the arrival or departure airport, not national airspace boundaries. The efficiency assessment is airport-centric: for instance, if a flight lands at Vancouver, the KPI measures aircraft behavior within the 200 NM terminal centered on that Canadian airport, even if part of this circular area overlaps US airspace (e.g., near Seattle). FIR or national airspace sectorization does not affect KPI definition or interpretation, as the metric focuses exclusively on time or distance spent in climb or approach phases within standardized 200 NM circular terminals. This ensures consistency and comparability across all airports, independently of proximity to international borders.”*

The GANP methodology intentionally uses this airport-centric approach to enable consistent global benchmarking regardless of national airspace boundaries.

Comment 4: Effect of published procedures on level-offs.

Response

We have added the following discussion (Lines 366–373): *“The scope of this study is intentionally limited to assessing performance trends and identifying relatively more or less efficient airports and regions, rather than attributing causality to specific operational factors. While published SID and STAR procedures may influence level-offs, as more restrictive procedures can limit the feasibility of CCOs/CDOs, this effect is considered part of the structural operating environment of each airport. As such, the KPI captures the effective outcome of climb and descent operations as executed in practice, independently of whether con-*

straints originate from procedural design, ATC actions, or other factors. Procedure-level analysis is outside the scope of this work and is left for future research.”

This approach aligns with the GANP framework, which evaluates operational outcomes rather than isolating individual causal factors.

Comment 5: Canadian airspace particularities and airport characteristics.

Response

The Canadian airspace system presents several distinctive features relevant to vertical efficiency assessment. Canada operates one of the world’s largest airspace domains, managed by NAV CANADA as a privatized air navigation service provider. The ten airports examined in this study represent diverse operational contexts: major international hubs (Toronto Pearson, Vancouver, Montréal), regional centres (Calgary, Edmonton, Ottawa), and airports with significant proximity to US airspace boundaries (Vancouver, Toronto).

Key operational factors now discussed in the manuscript include: transborder coordination (airports near the US border operate under bilateral coordination agreements); weather variability (from maritime conditions at Vancouver to continental extremes at Calgary and Edmonton); traffic complexity (mixed wide-body international, domestic narrow-body, and regional aircraft at major hubs); and airspace structure (Canadian terminal procedures generally follow ICAO standards with altitude and speed constraints that may influence CCO/CDO feasibility).

The standardized GANP methodology enables meaningful comparisons across Canadian airports and against international benchmarks. Differences observed between airports reflect the complete operational environment, including infrastructure, procedures, traffic mix, and meteorological conditions.

Comment 6: Typographical corrections.

Response

We thank the reviewer for these corrections. Both typographical errors have been corrected in the revised manuscript.

3.3 Response to Reviewer 3 – Ryota Mori

We sincerely appreciate the reviewer’s thoughtful comments regarding the scope and validation of our analysis. The feedback prompted us to clarify important aspects of our work’s objectives and methodological approach.

Comment 1: Data mining techniques in the abstract.

Response

We have revised the abstract accordingly. The reference to “data mining techniques” has been removed, as the current work focuses on applying standardized ICAO KPIs to analyze climb and descent profiles using established trajectory analysis methods rather than exploratory data mining approaches.

The manuscript implements a deterministic analysis framework based on ICAO’s GANP vertical efficiency KPIs, which is fundamentally different from data mining. Data-driven machine learning approaches (such as LSTM-based techniques for anomaly detection or new KPI development) represent valuable future research directions but are intentionally beyond the scope of the current work.

The abstract has been revised to present the findings more precisely. Rather than claiming that efficiency “is strongly influenced by” these factors, we now state that the analysis “identifies variations in vertical efficiency across airports, associated with differences in traffic patterns, geographic characteristics, and seasonal periods.”

Comment 2: Seasonality analysis.

Response

We have revised the manuscript to distinguish clearly between seasonality and traffic volume, which are related but conceptually distinct factors.

Seasonality definition: In this work, seasonality refers to systematic variations in vertical efficiency patterns across different seasons or temporal periods. Rather than analyzing the full calendar year, our 6-month dataset captures operational variations across different weather patterns and seasonal conditions (e.g., winter versus spring). We acknowledge that a full-year analysis would provide more comprehensive insights into seasonal patterns, and we explicitly note in the future work section that extended temporal analysis is a natural extension of this study.

Traffic volume data: We have included traffic volume statistics in the revised manuscript. These data clarify the relationship between traffic intensity and observed efficiency patterns. Specifically, traffic volume refers to the number of flight operations at each airport and on each route, while seasonality refers to how efficiency metrics vary across time periods.

We stop short of claiming direct causal relationships, as multiple factors (weather, procedures, aircraft type mix) simultaneously influence vertical profiles.

Comment 3: Geographic analysis.

Response

We have clarified and expanded the geographic analysis component. By “geographic factors,” we refer to the spatial distribution of airports across Canada and how their geographic location influences operational characteristics and efficiency outcomes.

Specifically, the geographic analysis includes: coastal vs. continental positioning (Vancouver benefits from maritime weather influences, which differ from continental airports such as Calgary and Edmonton); clustering and proximity effects (the Eastern corridor airports – Toronto Pearson, Montréal, Ottawa – are geographically proximate, leading to potential coordination effects and shared airspace characteristics); and regional characteristics affecting weather patterns, traffic routing, and connection to major US hubs.

These geographic contextual factors are now explicitly discussed in the manuscript to help explain observed differences in efficiency metrics across the ten studied airports.

Comment 4: Table 1 terminology clarification.

Response

We have expanded Table 1’s caption and the surrounding text to precisely define each category. The terminology reflects successive stages of data processing:

- **Filtered flights:** Flights that remain within the scope of the analysis after removing trajectories that cannot be meaningfully analyzed for vertical efficiency in the terminal area (overflights, trajectories not entering the 200 NM terminal area, fundamental data quality issues).
- **Complete flights:** A subset of filtered flights with complete trajectory information from cruise altitude to landing (or gate to cruise, for departures), with no significant gaps or temporal inconsistencies.
- **KPI flights:** The final subset of flights for which the KPI calculation function successfully computed the specific metric. This requires not only completeness but also meeting altitude thresholds and other KPI-specific criteria. “KPI flights (level-off climb)” specifically refers to departure flights where the KPI calculation identified and quantified level-off segments during the climb phase.

Comment 5: Quantitative support for procedure/sector claims.

Response

The original visualization was intended to illustrate that level-offs are not concentrated in spatially localized regions of the terminal area, but rather are distributed throughout the terminal airspace. We acknowledge that this visualization does not constitute rigorous statistical evidence of the absence of procedural concentration.

A detailed analysis of level-offs by specific SID/STAR procedures would require procedure-level data that is not the focus of the current work. We have revised the text to clarify that the figure demonstrates the spatial distribution of level-offs across the terminal area, and that detailed procedural attribution is beyond the scope of this work. Such analysis represents an important extension for future research.

Comment 6: Weather data and supporting evidence.

Response

We have significantly revised our discussion of meteorological factors to be more precise about our analytical approach and the scope of weather-related claims. The original text created an impression of direct causal analysis between weather and vertical efficiency, which is not the case.

We have clarified that the manuscript describes observed variations in vertical efficiency metrics across the time periods studied, without claiming that weather is the cause of these variations. A comprehensive weather impact analysis would require high-resolution weather data synchronized with flight trajectory data, statistical modeling to control for confounding factors, and methods specifically designed to handle complex interactions between weather and operational performance. This extended analysis is identified as a priority for future research.

Comment 7: SID structure and mean altitude of level-offs.

Response

We have incorporated this acknowledgement into the revised manuscript: *“The mean altitude of level-offs observed across airports reflects the operational environment at each airport, which includes SID design, airspace structure, ATC procedures, traffic flows, and weather conditions. The observed differences in level-off altitudes between airports likely reflect these combined procedural and operational factors. A procedure-level analysis examining the explicit altitude restrictions in each airport’s SIDs would provide additional insights and is recommended for future research.”*

Comment 8: Vertical efficiency vs. inefficiency labelling.**Response**

We have reviewed our KPI definitions and revised figure labelling for consistency and precision. For time-based metrics (KPI17, KPI19), higher values (more time in terminal area) indicate lower efficiency or higher inefficiency; figures now use labelling such as “Vertical Inefficiency Indicator (Time in Terminal)” to avoid ambiguity. Where ambiguity could arise, we have added explicit clarifications in figure captions (e.g., “higher values indicate greater inefficiency” or “lower values indicate better performance”).

Overall work scope and methodology positioning.**Response**

The current manuscript is intentionally designed as a descriptive analysis that applies established ICAO KPI frameworks to a comprehensive Canadian dataset. The contribution lies in three areas:

- **Methodological demonstration:** Showing how standardized ICAO KPIs can be systematically applied to operational trajectory data from open-source systems (OpenSky Network) at a national scale across ten major airports.
- **Baseline assessment:** Providing the first comprehensive characterization of vertical efficiency patterns at major Canadian airports using standardized metrics, establishing a baseline for future comparative studies.
- **Identification of variations:** Documenting significant differences in efficiency metrics across airports and temporal periods, which motivate deeper causal analysis.

The reviewer’s suggestion for “more in-depth analysis with expanded data and rigorous statistical validation” is entirely appropriate for follow-on research. The current work is explicitly positioned as Step 1 in a multi-stage research program, providing a foundation and roadmap for statistical hypothesis testing, machine-learning approaches, causal inference, and longitudinal analysis. This positioning has been clarified in the manuscript’s Introduction and Future Work sections.