EDITORIAL

Reviews and Responses for
Improving Efficiency Through the Publication of Expected Distances for Standard Terminal Arrival Routes

Authors: Jan Krummen, Timothé Krauth, Jan Allendorf, and Raphael Monstein

Reviewers: Tatiana Polishchuk and Junzi Sun

Editor: Xavier Olive

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2024.7894

2. Review - round 1

2.1 Reviewer 1

In this paper, the authors are investigating the benefits of publishing the expected flight distances for STARS, a new concept not previously addressed in scientific literature. The idea is deemed promising in providing noticeable fuel savings due to the fact that, in practice, aircraft are given shortcuts within certain TMAs and very rarely follow the full STAR procedures.

The paper is very well written, with all the analyses explained in sufficient detail. However, I have a number of questions and comments that I recommend addressing to improve the quality of the paper even further.

Safety considerations regarding the situations when aircraft may end up arriving at the TMA with insufficient fuel should be addressed and investigated further. How the contingency fuel is calculated? Will it always be sufficient to cover both bad weather events and congestion in TMAs?

The authors claim that they extracted the data corresponding "using geographical boundaries wide enough to cover the full extent of all STARs". How large are the resulting areas? Are they different for the airports in consideration? EUROCONTROL usually uses similar areas around the TMAs, with a radius of 40 NM, 50NM, or 100 NM. This simplifies the comparison between the airport's TMAs, which differ in size and shape (though this is probably not among the purposes of this research). The authors may want to adopt this idea in future research.

It would be interesting to understand how the proposed expected distances are actually calculated in different airports. In case this idea is adopted, what method would you recommend for the estimation of these values? The presented results show that in some airports, the expected distance is positioned within the overall distribution, in others, further away. Could you kindly comment on the possible consequences of over/under-estimation of these distances and suggest a method or express your opinion on what would be the correct way of defining them.

The analyses of the factors impacting the actual distance aircraft fly in the TMA revealed that visi-

bility may have a negative impact on the distance flown. Most probably, not the visibility itself, but the enforced airport procedures for low visibility, which reduce the traffic intensity, may cause this effect. However, there should be some threshold after which the procedures are enforced, and the impact becomes noticeable. Do you think your research can capture these thresholds?

No details are provided about the method used for the fuel calculation in Section 5.

Overall, the presented idea is very promising, and it would be interesting to discuss it with the operational experts at ANSPs. Would they be interested in implementing this in practice?

2.2   Reviewer 2

This is a well-written paper with a clear structure that quantifies the distance of standard terminal arrival procedures and the difference between actual distances and expected distances among three airports in Europe. The paper presents a very nice and innovative use case of OpenSky data for evaluating ATM operations. My comments are as follows:

1. What is the motivation for choosing these airports? Is it due to data availability, or are these airports known to have significant deviations between actual flight distances and expected distances?

2. In the introduction of this paper, line 29 states that STAR cannot be readily known to flight planning crews. Can these be estimated in advance using historical data?

3. Line 51 indicates an excess of 50 NM. How much fuel does that approximately equate to?

4. In Figure 1(b), what is the purpose of the loop (GVA-BILLO->GG502->INDIS->GVA) in the STAR procedure?

5. Line 53 indicates that some ANSPs publish expected STAR distances. Could you please provide a list of which ANSPs provide such information?

6. In section 4, a linear regression model is used to analyze the importance of the features. The R2 score is relatively low. Could you try a different approach, such as a random forest model, for feature importance analysis?

7. How could the uncertainty in actual STAR distances influence the calculation of the top-of-descent? Could this also affect the efficiency of continuous descent operations?

8. In section 5, fuel savings on example flights are provided. Would it also be possible to provide aggregated fuel savings for all the flights? Perhaps using performance models like OpenAP?

9. Judging from Figures 7-11, the difference between rush and non-rush hour seems marginal. Does this mean that efficiency is not a priority for air traffic controllers when managing arrival trajectories?

10. Following the previous comment, can we also conclude that the STAR procedures are not efficient and need to be redesigned for the airports studied?

3. Response - round 1

3.1   Response to reviewer 1

a) Safety considerations regarding the situations when aircraft may end up arriving at the TMA with insufficient fuel should be addressed and investigated further. How the contingency fuel is calculated? Will it always be sufficient to cover both bad weather events and congestion in TMAs?

> **Response**
>
> We fully agree with reviewer on the critical importance of safety considerations in fuel planning. Ensuring the safety of flight operations is, and must always remain, the first priority when proposing or implementing any procedural changes. With regard to contingency fuel, its calculation is strictly governed by ICAO guidance and regional regulations, such as ICAO Document 9976 and Regulation (EU) No 965/2012. Contingency fuel is specifically designed to account for various operational uncertainties, including deviations from planned routes, adverse weather conditions and air traffic congestion.
>
> It is important to emphasise that the concept of expected distances only affects the calculation of trip fuel, better aligning it with the actual route typically flown. This adjustment does not affect other components of the fuel planning, such as contingency fuel, alternate fuel or discretionary fuel, which remain important safeguards against unforeseen circumstances. Therefore, situations where unexpected weather conditions and air traffic congestion within the TMA lead to longer than expected flight routes (e.g. flying the full STAR instead of the expected distance for procedures with expected distance or entering holding patterns for those without) may occur under both scenarios: with and without the publication of expected distances. Such situations should be covered by appropriately calculated contingency fuel and other fuel reserves, such as extra or discretionary fuel, to ensure that operational safety is not compromised.
>
> STAR design varies significantly across airports, reflecting specific operational needs such as traffic volume, geography, and adjacent airspace constraints. Different approaches to path stretching exist, which impact flight planning in different ways:
>
> - Holding Patterns: Typically not accounted for as additional track miles or fuel requirements in flight planning tools.
> - Closed-Path Designs (e.g., Trombone/Point Merge STARs): These allow aircraft to follow predefined, extendable paths, and EASA fuel regulations explicitly permit flight planning based on a "reasonably expected route" rather than the full procedure.
> - Open-Path Designs: Where vectoring by air traffic control determines the final route, introducing further variability in flown distances.

b) The authors claim that they extracted the data corresponding "using geographical boundaries wide enough to cover the full extent of all STARs". How large are the resulting areas? Are they different for the airports in consideration? EUROCONTROL usually uses similar areas around the TMAs, with a radius of 40 NM, 50NM, or 100 NM. This simplifies the comparison between the airport's TMAs, which differ in size and shape (though this is probably not among the purposes of this research). The authors may want to adopt this idea in future research.

> **Response**
>
> We acknowledge that the description of the geographical boundaries used for data extraction is not very detailed in the paper and we appreciate the opportunity to clarify this point. The areas considered were indeed different for the three airports studied. Specifically, the approximate dimensions of the rectangular areas were as follows:
>
> - Geneva (LSGG): 130x130 NM
> - Rome Fiumicino (LIRF): 180x180 NM
> - Munich (EDDM): 130x130 NM
>
> These differences reflect the different extent of STARs at each airport and were chosen to ensure full coverage of all STARs for data collection purposes only. To clarify that different boundaries were used and to give an idea of their size, the following sentence has been adjusted in the paper:

"The data was retrieved using geographical boundaries tailored to each airport, with sizes adjusted to ensure complete coverage of the full extent of all STARs. These boundaries vary according to the extent of the STARs at each location and correspond to sizes ranging from approximately 130x130 NM to 180x180 NM."

However, we would like to emphasise that the size of the extraction areas does not directly affect the primary results of the study. All trajectories analysed were clipped between the first waypoint defining the STAR and the runway threshold. This ensures that the comparison is strictly focused on the STAR length flown relative to the published full procedure and expected distances, rather than including additional flight segments outside the STARs.

We appreciate the suggestion to consider adopting standard radii (e.g. 40 NM, 50 NM or 100 NM) as used by EUROCONTROL in future research. While it was not the purpose of this study to compare the size or shape of different TMA areas, adopting such a standardised approach could facilitate cross-airport comparisons in future analyses.

c) It would be interesting to understand how the proposed expected distances are actually calculated in different airports. In case this idea is adopted, what method would you recommend for the estimation of these values? The presented results show that in some airports, the expected distance is positioned within the overall distribution, in others, further away. Could you kindly comment on the possible consequences of over/under-estimation of these distances and suggest a method or express your opinion on what would be the correct way of defining them?

Response

We appreciate the reviewer s interest in the methodology behind calculating expected distances and the potential implications of over- or under-estimating these values.

From our findings, we can see that for those airports where expected distances are in place, the corresponding percentiles of flown distances vary slightly, but consistently fall between the 85th and 95th percentiles over the relevant periods for which they are valid. As mentioned in the discussion of the paper, his suggests that the expected distances are probably determined by analysing historical trajectories and selecting a certain percentile (around the 90th percentile) of the observed distances as the expected distance. However, this assumption is based on our observations and we have no confirmation of the exact methodology used. For future work, it would be valuable to request this information from the respective ANSPs to better understand their processes.

The presumed approach used is also directly in line with the methodology we would recommend for determining expected distances at airports where this practice could be introduced. Specifically, the use of a high percentile (e.g. 90th) of observed distances over a sufficiently long observation period. This will ensure that the expected distance reflects operational reality, while providing an adjustable percentage of flights that may fly longer distances than the expected one.

However, the choice of percentile must be made carefully, as both over- and underestimation of expected distances have different consequences:

Overestimation: Expected distances that are too high result in overly conservative trip fuel calculations. This minimises the potential fuel savings and operational efficiencies that the practice aims to achieve.

Under-Estimation: An underestimation of expected distances would result in a greater proportion of flights flying longer distances than expected one that is published in the AIP. This could lead to situations where the carried trip fuel is more often insufficient to cover the actual distance flown. In such scenarios, aircraft may have to rely more heavily on contingency or discretionary fuel. This could undermine the confidence and reliability of expected distances as a planning tool, discouraging operators from fully adopting the practice.

d) The analyses of the factors impacting the actual distance aircraft fly in the TMA revealed that visibility may have a negative impact on the distance flown. Most probably, not the visibility itself, but the enforced airport procedures for low visibility, which reduce the traffic intensity, may cause this effect. But then there should be some threshold after which the procedures are enforced, and the impact becomes noticeable. Do you think your research can capture these thresholds?

> **Response**
>
> Indeed, it is possible that low visibility itself does not directly reduce distances, but rather triggers specific operational measures, such as reduced traffic intensity or different approach patterns, which ultimately influence the distances flown.
>
> In our current analysis, visibility was treated as a continuous variable, and while the results indicate a statistically significant negative relationship between visibility and distances flown, our methodology did not explicitly focus on identifying thresholds at which, for example, low visibility procedures (LVP) are activated. Identifying such thresholds would require integrating data on the precise operational status of the airport (e.g. whether LVPs are active) or correlating visibility with significant changes in traffic patterns. This could however be an interesting path for further research.

e) No details are provided about the method used for the fuel calculation in Section 5.

> **Response**
>
> We appreciate the reviewer's comment on the lack of detail on the method used for fuel calculations in Section 5. As mentioned in the text, the fuel calculations that form part of the operational flight plans (OFP) were performed using CAE's Flight Plan Manager v6.5.1, a widely used and industry-standard flight planning software. This tool incorporates detailed aircraft performance models and takes into account operational variables such as aircraft weight, weather conditions, and routing.
>
> For clarity, the following sentence in the paper has been slightly adjusted:
>
> "Using CAE's Flight Plan Manager v6.5.1 software, three versions of an Operational Flight Plan (OFP), each detailing the required fuel amounts for different scenarios, were generated for this flight."
>
> It should be noted, however, that the exact details of how the calculations are performed in the background are not known to us as the software is a proprietary commercial tool and its source code is not available. This tool is widely used and trusted by airlines throughout the aviation industry for flight and fuel planning and forms the basis for many operational decisions. Given its widespread use and proven reliability in real-world applications, it is reasonable to assume that it provides accurate and reliable results, although the exact calculation processes are not accessible to us.

3.2   Response to reviewer 2

a) What is the motivation for choosing these airports? Is it due to data availability, or are these airports known to have significant deviations between actual flight distances and expected distances?

> **Response**
>
> The choice of Munich (EDDM), Rome Fiumicino (LIRF), and Geneva (LSGG) was based on the following considerations:
>
> Munich and Rome: These airports were chosen because we were aware that they had published expected STAR distances in their respective AIPs, making them ideal candidates for studying the impact of such measures. In addition, the ADS-B coverage provided by the OpenSky network around these airports is sufficient to ensure data collection and analysis.

Geneva: Geneva was selected because it provides an interesting counterpoint to Munich and Rome. Unlike these airports, Geneva does not publish expected distances, although there is evidence of significant discrepancies between full STAR distances and actual distances flown for certain STAR procedures. This allowed us to investigate the potential benefits of introducing expected distances at an airport with significant shortcut potential.

b) In the introduction of this paper, line 29 states that STAR cannot be readily known to flight planning crews. Can these be estimated in advance using historical data?

Response

While it is true that the exact STAR to be flown cannot be determined with certainty by flight planning crews, it is possible to make informed estimates based on historical data and current conditions.

By analysing historical flight data, crews or flight planning systems can predict the likelihood of certain STARs being assigned for a given city-pair under certain conditions (e.g. prevailing wind directions, time of day or runway usage patterns).

However, these predictions are not guaranteed and remain subject to real-time factors such as short-term weather changes, air traffic control instructions or unexpected congestion within the TMA.

To make this clearer, we have slightly revised the relevant sentence in the introduction to better reflect this aspect:

"During pre-flight planning, crews cannot predict with certainty which specific SID and STAR will be used, resulting in some uncertainty regarding the required fuel demand for a flight."

c) Line 51 indicates an excess of 50 NM. How much fuel does that approximately equate to?

Response

While exact fuel consumption varies depending on factors such as aircraft type, weight, and weather conditions, we can provide an estimate based on the specific case of an Airbus A220-300, which was used in the sample calculations in Section 5.

According to the operational flight plans presented in Section 5, flying the full STAR approach compared to a straight-in approach requires an additional 244 kg of trip fuel. This corresponds to approximately 244 kg of fuel for the additional 50 NM.

d) In Figure 1(b), what is the purpose of the loop (GVA-BILLO->GG502->INDIS->GVA) in the STAR procedure?

Response

We understand that this loop exists due to the proximity of French controlled airspace to the west of Geneva. Aircraft are typically handed over from French controllers just before entering the STAR while already close to Geneva airport. The loop provides controllers with additional flexibility and "breathing space" to manage incoming traffic, particularly in high traffic or complex situations. If needed, it acts as an extra buffer to ensure that air traffic control can safely sequence and space aircraft before they approach the runway.

e) Line 53 indicates that some ANSPs publish expected STAR distances. Could you please provide a list of which ANSPs provide such information?

> **Response**
>
> We appreciate the reviewer's question about which ANSPs publish expected STAR distances. Based on our knowledge:
>
> DFS (German ANSP): Publishes expected STAR distances for most major German airports, such as Munich (EDDM), Frankfurt-Main (EDDF), Berlin Brandenburg (EDDB) and Duesseldorf (EDDL) to name a few.
>
> ENAV (Italian ANSP): Provides expected STAR distances for a few selected airports, notably Rome Fiumicino (LIRF) and Milan Malpensa (LIMC).
>
> However, it is possible that other ANSPs of which we are currently unaware may also publish expected distances.
>
> We agree that this is information that could be useful to potential readers of the paper which is why we adjusted the following sentence in the paper:
>
> "To address this, some Air Navigation Service Providers (ANSPs), such as DFS and ENAV, publish expected flight distances for different STAR procedures at specific airports in the Aeronautical Information Publication (AIP), allowing operators to use these expected rather than full STAR distances for flight and fuel planning purposes."

f) In section 4, a linear regression model is used to analyze the importance of the features. The R2 score is relatively low. Could you try a different approach, such as a random forest model, for feature importance analysis?

> **Response**
>
> The R2 value provides insight into how well the selected regressors explain the variability in the data and is primarily used as a metric for the predictive power of the model. Its value depends heavily on the nature of the dependent variable, any transformations applied to it, and the choice of explanatory variables.
>
> In this study, while we noted the R value, it is not the primary focus. Our objective is to analyse relationships and dependencies between specific variables and the flown distance of STARs, rather than to build a highly predictive model. This perspective aligns with an econometric approach, where the emphasis is on deriving reliable coefficient estimates and understanding causal relationships, rather than maximizing predictive performance. Low R values are common in econometric studies, particularly in microeconomic models, where outcomes are influenced by numerous unobservable factors. Similarly, in our case, many unobserved variables likely influence flown distance, and our goal is to focus on the variables of interest that we can observe.
>
> While alternative approaches, such as a random forest model, could indeed provide insights into feature importance, we opted for linear regression as it is more suitable for an econometric analysis. Random forests are better suited for predictive modeling and feature selection within a machine learning framework, which was not the primary goal of this study.

g) How could the uncertainty in actual STAR distances influence the calculation of the top-of-descent? Could this also affect the efficiency of continuous descent operations?

> **Response**
>
> Uncertainty in actual STAR distances can lead to inaccuracies in the determination of the top of descent, as flight crews or flight management systems rely on planned distances to calculate the optimal descent point. If the actual distance flown differs significantly from the planned distance (e.g. due to shortcuts

or deviations), the aircraft may descend too early or too late, potentially resulting in inefficient altitude profiles. This could result in the need for less efficient corrective actions such as leveling. Such disturbances can also negatively affect continuous descent operations, as it becomes more difficult to maintain an uninterrupted descent profile if the planned and actual distances do not match.

To highlight the benefits of expected distance information on continuous descent operations, an additional sentence was added to the paper's discussion:

"Having the information about expected track miles helps pilots to manage in-flight descent planning, especially at unfamiliar or infrequently visited airports. This factor is especially significant for continuous descent operations, as accurate STAR distance estimates can greatly enhance the determination of the top of descent point."

h) In section 5, fuel savings on example flights are provided. Would it also be possible to provide aggregated fuel savings for all the flights? Perhaps using performance models like OpenAP?

> **Response**
>
> While OpenAP would indeed be a valuable tool for modelling the fuel savings of flying a shorter route (e.g. comparing the fuel consumption of a full STAR versus a shortcut or expected distance), it is less suited to the specific example calculations presented in this paper.
>
> The focus of the example calculations in Section 5 was to highlight the savings from not carrying excess fuel for the full STAR when flying a straight-in approach. In this case, the savings come from reducing the weight of fuel carried over the entire flight, rather than from differences in the distance flown alone. While OpenAP could theoretically be used to model this by simulating flights with different initial weights, operational flight plans would still need to be generated for each scenario to determine the exact fuel requirements and the resulting take-off weight for the trip based on the full STAR or expected distances. This process would need to be repeated for each flight to account for variability in aircraft type, route and operating conditions.
>
> While this approach is possible, it would require additional resources and data inputs beyond the scope of the current study.

i) Judging from Figures 7-11, the difference between rush and non-rush hour seems marginal. Does this mean that efficiency is not a priority for air traffic controllers when managing arrival trajectories?

> **Response**
>
> We wouldn't agree with that interpretation. Rather, the plots indicate that the difference between the full and observed distances is consistent across both peak and off-peak periods. This demonstrates that even during peak periods, the full STAR procedure is rarely followed, with significant shortcuts still being provided. If anything, this highlights that air traffic controllers prioritize efficiency even during rush hours by optimizing routes and providing shortcuts wherever possible.

j) Following the previous comment, can we also conclude that the STAR procedures are not efficient and need to be redesigned for the airports studied?

> **Response**
>
> While the data suggest that full STAR routes are rarely flown at the airports considered, it is important to recognise that these procedures were probably originally designed in this way for specific operational

reasons.

Full STAR routes provide controllers with additional flexibility to manage spacing, sequencing and unforeseen circumstances such as sudden traffic surges or adverse weather conditions. In certain scenarios, the ability to use the full STAR is essential to maintain safety and operational efficiency.

If long-term observations consistently show that aircraft rarely, if ever, fly the full STAR distance, it may be worth considering whether procedures could be adapted. However, changing STAR procedures is a complex and time-consuming process that requires careful coordination between regulators, air navigation service providers and operators. In contrast, updating the AIP to include expected distances offers a much more practical and efficient solution to optimise fuel planning while maintaining operational flexibility.