Journal of Open Aviation Science (2024), Vol.**2**
doi:10.59490/joas.2024.8170

JOAS

EDITORIAL

Reviews and Responses for Evaluating Potential Fuel-Savings of External Alternative Ground Propulsion Systems

Authors: Manuel Waltert, Benoit Figuet, and Michael Felux

Reviewers: Xavier Olive, Martin Strohmeier, and Tatiana Polishchuk

Editor: Junzi Sun

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2024.7888

2. Review - round 1

2.1 Reviewer 1

The paper is well-written and engaging, analysing potential fuel savings associated with external alternative ground propulsion systems (AGPS). The topic is relevant to the OpenSky Symposium and well suited for the proceedings.

Please find below some comments intended to further refine and improve the paper.

Line 61 (ICAO Emission Databank and Fuel Flow/Consumption Calculations): While the paper uses the ICAO Aircraft Emission Databank as a source, there is no explicit mathematical explanation connecting this data to the calculated fuel flow or fuel consumption. Consider adding equations or a more detailed explanation.

Line 92: Correct the reference to Section 2.4, which should be Section 2.1.

I recommend mentioning that some aircraft activate their transponders midway through taxiing, which could pose a limitation to achieving complete data coverage. This insight would add depth to your discussion of data reliability.

Section 2.2.1 (Take-Off Runway Detection): Given the paper's emphasis on excellent coverage, you might consider leveraging trajectory intersections with runways for take-off runway detection (e.g., using the ‘aligned_on_runway‘ method).

Section 2.2.2 (Pushback Patterns): Including map plots of S-shape, U-shape, and L-shape pushback patterns would be interesting.

Figure 2 (Fuel Consumption During Queueing): Clarify how fuel consumption is modelled when aircraft are stationary or queueing. Additionally, the transition between AGPS tugs and engine taxiing is not entirely clear. For example, do tugs transport aircraft to the runway threshold, or do they/could they transition earlier?

Figure 3 (Whisker Plot Outliers): The high number of outliers for runways 28 and 32, some showing taxi times of up to 60 minutes, needs further explanation. Are these outliers realistic? If so, contextualize them (e.g., atypical operational scenarios, extreme traffic congestion).

Figures 4 and 5 (Fuel Consumption Representation): The scale inconsistency (values in multiples of 10^6 but the graph scaled in 10^7) hampers the clarity of the results. I recommend switching the units to tons for easier interpretation. Additionally, providing average taxi fuel consumption per aircraft (by dividing total values by the number of trajectories?) would offer better context.

Line 291 (Tug Usage Recommendation): The suggestion to use tugs exclusively for runway 16 raises operational questions about scenarios where departures are from runway 28. What would the tugs do in such cases?

Line 316 (Optimizing Limited AGPS Units): Bin-packing might not yield the optimal solution for utilizing limited AGPS units. Consider discussing alternative optimization techniques, such as a constrained resource allocation model or a heuristic approach.

Replace all instances of “the OSN” with “the OpenSky Network” for clarity in the text; you may maintain “OSN” in tables and plots.

The paper relies heavily on footnotes. Consider integrating these comments into the main text where feasible, or removing them when not really necessary in the text. URLs in particular, could be included in the bibliography using the @misc BibTeX format.

Recommendation: Revisions Required

2.2 Reviewer 2

This paper discusses fuel savings with alternative ground propulsion systems, using the example of Zurich airport. Its main contribution is to use open ADS-B data in order to accurate calculate taxi times instead of using simplistic models. It is well-written and referenced. Overall, it is a nicely written paper, which I am happy to recommend for accept. My only gripe is that it focuses in its exposition first on the literature baseline but later only discusses the results in context of potential fuel savings of AGPS. I would like to see the discussion (and the evaluation) circle back to the baselines from the literature and compare how the assessment changed now that it is more accurate. This could be done by comparing the outcomes directly with numbers from these baseline methods.

My further notes and comments:

- Some things may seem obvious (more savings potential for aircraft with longer taxiing times; more towing units are better) but are clearly underpinned by the data here.
- The potential operational challenges could have been discussed a bit more. Future work might include these into the simulation.
- Did you submit `get_toff_runway()` to the traffic library?
- My main question is, did you validate your new functions (also pushback()) and how? Did you properly compare them to the existing ones in traffic?
- You nicely justify and reference most of your data and assumptions but not all, e.g. "start-up process of a single engine takes 60 seconds"
- "In the same observation period, Zurich Airport Ltd. (FZAG) reported 59,222 departures [19]" You consider only certain aircraft types, is it the same basic types in the FZAG dataset?
- What explains that the delta is not consistent across runways? Especially 10 and 34?
- Absolute numbers of fuel reduction vs relative? Still absolutely more for longer taxi times?

Nitpicks:

- Footnote 1 seems at least partly superfluous?
- Please check the date format, e.g. 01. May, 2024 in Fig. 5/Table 3 is wrong and inconsistent with other date usage in the paper.
- "However, as shown in Figure 5, significant fuel-saving potential can already be achieved with a much smaller number of external AGPS units. Nonetheless, a significant part of the fuel-saving potential can still be realized with a much smaller number of external AGPS units in use, making it a more cost-effective approach." These sentences seem to be the same?

Code:

- https://github.com/m-waltert/osn24_agps is not accessible yet!

Recommendation: Accept Submission

2.3 Reviewer 3

The study evaluates the potential benefits of utilizing external Alternative Ground Propulsion Systems (AGPS) at Zurich Airport, analysing the observed ADS-B-based service trajectory data for 5 months. Additional contribution is the development of the alternative classification method (additional to the existing one in traffic library) for identifying the parts of the trajectory corresponding to aircraft pushback.

The paper is well written and provides all the necessary background information and the details about the work. My comments and suggestions for improvement are rather minor:

1. Turboprop commercial aircraft were deliberately not considered in the studies, but the authors did not explain why.
2. line 224: the authors introduce a threshold value t_{AGPS}. Explain how it is determined.
3. line 307: in future work, the selection of which aicraft are to be chosen taxiing with AGPS, should include considerations regarding feasibility of scheduling of such vehicles.
4. line 328: I suggest adding a couple of references to the recent work related to the case when ground trajectories have gaps:
   - Martin Schlosser, Hannes Braßel, and Hartmut Fricke. Analysis of Aircraft Ground Trajectories: Map-Matching with Open Source Data for Modeling Safety-Driven Applications. ICRAT 2024
   - Xavier Olive, Manuel Waltert, Ryota Mori, Philippe Mouyon. Filtering Aircraft Surface Trajectories Using Information on the Taxiway Structure of Airports, ICRAT 2024.
5. Again, for future work. Conducting a cost-benefit analysis for the scenarios with different number of AGPS vehicles can naturally complement the studies. Of course, different stakeholders will benefit from the provided improvements, as fairly noted in the conclusions, but the overall environmental goal should always be a good motivator ;)

3. Response - round 1

3.1 Response to reviewer 1

Line 61 (ICAO Emission Databank and Fuel Flow/Consumption Calculations): While the paper uses the ICAO Aircraft Emission Databank as a source, there is no explicit mathematical explanation connecting this data to the calculated fuel flow or fuel consumption. Consider adding equations or a more detailed explanation.

Response

We have completely rewritten Section 2.3. As part of this revision, we added the formulae required for estimating the taxi fuel consumption, among other things.

Line 92: Correct the reference to Section 2.4, which should be Section 2.1.

Response

The reference has been corrected. It now states "Section 2.1".

I recommend mentioning that some aircraft activate their transponders midway through taxiing, which could pose a limitation to achieving complete data coverage. This insight would add depth to your discussion of data reliability.

Response

The 5th paragraph of our discussion explicitly states this fact: *Another source of error is that pilots sometimes do not switch on their transponders until they have started taxiing, resulting in incomplete trajectories.*

Section 2.2.1 (Take-Off Runway Detection): Given the paper's emphasis on excellent coverage, you might consider leveraging trajectory intersections with runways for take-off runway detection (e.g., using the `aligned_on_runway` method).

Response

In collaboration with Xavier Olive, we have implemented the method described in our study for classifying the take-off runway in function `takeoff()` of the *traffic* library. Applied to the example of Zurich Airport, this function can be applied to a `flight` as follows:

```
flight.takeoff('LSZH', method='track_based')
```

Section 2.2.2 (Pushback Patterns): Including map plots of S-shape, U-shape, and L-shape pushback patterns would be interesting.

Response

Figure 1 contains three specifically selected pushbacks: (i) a S-shaped pushback on the eastern part of the midfield dock, (ii) a U-shaped pushback on the western part of the midfield dock, and (iii) an L-shaped pushback from Terminal B.

Figure 2 (Fuel Consumption During Queueing): Clarify how fuel consumption is modelled when aircraft are stationary or queueing. Additionally, the transition between AGPS tugs and engine taxiing is not entirely clear. For example, do tugs transport aircraft to the runway threshold, or do they/could they transition earlier?

Response

Aircraft being stationary or queueing are not considered specifically. Our model for estimating fuel consumption during the taxi process assumes that the aircraft engines are always operating in idle conditions. Moreover, the mass of the aircraft, acceleration of the aircraft on the ground, etc. is not taken into account when calculating taxi fuel consumption.

Even though our assumptions are in line with the literature, e.g., see [1], [2], [3], [4], [5], it represents a limitation, which we address in the discussion.

Figure 3 (Whisker Plot Outliers): The high number of outliers for runways 28 and 32, some showing taxi times of up to 60 minutes, needs further explanation. Are these outliers realistic? If so, contextualize them (e.g., atypical operational scenarios, extreme traffic congestion).

Response

The long taxi times mentioned are due to (i) "legitimate" outliers (e.g., aircraft waiting for de-icing pads for reasons unknown to us, or cases of airport congestion), (ii) measurement errors (e.g., noisy data), or (iii) flights that remain on the stand for a very long time before taxiing. To reduce the number of outliers, we have taken the following measures:

- To determine the ‘start taxi’ time of flights, we only consider the part of the trajectory at which the aircraft is a certain distance (0.05 NM) away from the initial position of the trajectories. This allows us to filter noisy data from aircraft waiting on stands.
- Before we perform the fuel reduction potential analysis, we filter our data set according to the following additional conditions:
  - Flights that have jumps in the computed groundspeed (`compute_gs`) of more than +200 and less than −200 between two consecutive measurements are not considered for the analysis.
  - Flights that have jumps in the computed cumulative ground distance of more than 0.1 between two consecutive measurements are not considered for the analysis.

  Using these filter conditions, we can remove a large part of the outliers (i.e., noisy surface trajectories) from the dataset used for the analysis. However, this more stringent filtering of the data set contributes to the fact that the delta between the number of flight movements reported by FZAG and the number of flight movements in our data set, as shown in Table 1, becomes larger compared to the delta reported in our initial submission.

  We would also like to note at this point that we have replaced the box plot in Figure 3 with a violin plot. In our opinion, a violin plot allows for a better visualisation of the distribution of the data.

Figures 4 and 5 (Fuel Consumption Representation): The scale inconsistency (values in multiples of 10^6 but the graph scaled in 10^7) hampers the clarity of the results. I recommend switching the units to tons for easier interpretation. Additionally, providing average taxi fuel consumption per aircraft (by dividing total values by the number of trajectories?) would offer better context.

Response

We adjusted Figures 4 and 5 so that the Y-axis now shows the fuel consumption in tons (i.e., in the unit of 10^3 kg). We also adjusted Tables 2 and 3 accordingly and converted all references in the text.

Additionally, we also extended the results by providing the estimated average fuel consumption per taxi movement (190.69 kg per movement) as well as the average fuel consumption per movement of Airbus A320 family aircraft types (152.42 kg per movement).

Line 291 (Tug Usage Recommendation): The suggestion to use tugs exclusively for runway 16 raises operational questions about scenarios where departures are from runway 28. What would the tugs do in such cases?

Response

Given its length, runway 16 at Zurich Airport is typically used by aircraft that cannot take off from runway 28 due to field length restrictions. This primarily includes long-haul flights which require a long runway due to their high take-off weight. This is our motivation behind focussing exclusively on runway 16 departures for AGPS usage. To explicitly state this fact in the paper, we added the following paragraph to Section 2.4:

We analysed the taxi fuel-saving potential at Zurich Airport for two different groups of aircraft eligible to be towed by an AGPS. Specifically, we examined the impact of AGPS when applied to all departing flights versus only those departing from runway 16. Due to its length, runway 16 is primarily used by long-haul aircraft. Since long-haul aircraft typically have higher specific fuel consumption due to their larger engines and flights departing from runway 16 have considerably longer taxi durations compared to departures from other runways at Zurich Airport, see Table 1, distinguishing this group is particularly valuable for assessing the taxi fuel-saving potential of AGPS.

As we wrote in our discussion, the use of the operational reserve of pushback vehicles (i.e., vehicles that are on stand-by and could step in if needed) would be particularly suitable for scenarios in which aircraft are only towed from the stand to the runway very selectively. This answers your question regarding ‘what would the tugs do in such cases’: they would be available as an operational reserve and could be used for pushback/AGPS duties if necessary.

Line 316 (Optimizing Limited AGPS Units): Bin-packing might not yield the optimal solution for utilizing limited AGPS units. Consider discussing alternative optimization techniques, such as a constrained resource allocation model or a heuristic approach.

Response

We completely agree with you. However, we have deliberately chosen bin-packing for our OSN-contribution, as we would like to discuss alternative optimisation techniques in a future publication of ours. Indeed, we will address the question of optimal AGPS usage policies in our (accepted) contribution for the 28th Air Transport Research Society (ATRS) World Conference in Hong Kong from July 1st to July 4th, 2025.

Replace all instances of “the OSN” with “the OpenSky Network” for clarity in the text; you may maintain “OSN” in tables and plots.

Response

All instances of "the OSN" in the text have been replaced with "the OpenSky Network".

The paper relies heavily on footnotes. Consider integrating these comments into the main text where feasible, or removing them when not really necessary in the text. URLs in particular, could be included in the bibliography using the @misc BibTeX format.

Response

- We removed Footnotes 1, 2, and 9.
- Footnotes referring to www.wheeltug.com and https://taxibot-international.com/ were included into the bibliography using the @misc BibTeX format.
- Footnotes 5, 6, and 8 were integrated into the main body of the text.

3.2 Response to reviewer 2

My only gripe is that it focuses in its exposition first on the literature baseline but later only discusses the results in context of potential fuel savings of AGPS. I would like to see the discussion (and the evaluation) circle back to the baselines from the literature and compare how the assessment changed now that it is more accurate. This could be done by comparing the outcomes directly with numbers from these baseline methods.

> **Response**
>
> We have expanded the discussion by additionally referring to four sources mentioned in the literature review: In the first paragraph of the discussion we compare our results with [6]. In the second paragraph with [1] and [3]. In the third last paragraph, we make a reference to [5].

Some things may seem obvious (more savings potential for aircraft with longer taxiing times; more towing units are better) but are clearly underpinned by the data here.

> **Response**
>
> We completely agree with you on this. For this reason, we have (attempted to) address these points in the discussion. The second paragraph of the discussion deals, among other things, with the influence of longer taxi times, while the third paragraph deals with impact of the number of AGPS units.

The potential operational challenges could have been discussed a bit more. Future work might include these into the simulation.

> **Response**
>
> We have tried to address or point out the operational challenges resulting from the use of AGPS in Section 5. To limit the scope of our OpenSky Symposium paper, we have deliberately decided to address operational challenges as well as usage policies in a separate study. Indeed, we will address the question of the optimal usage policy of AGPS in our (accepted) contribution for the 28th Air Transport Research Society (ATRS) World Conference in Hong Kong from July 1st to July 4th, 2025.

Did you submit `get_toff_runway()` to the traffic library?

> **Response**
>
> The `get_toff_runway()` function mentioned in our original submission has been implemented in the `takeoff()` function of the Traffic Library in collaboration with Xavier Olive. Applied to the example of Zurich Airport, the takeoff classification function mentioned in our paper can be applied to a `flight` as follows:
>
> ```
> flight.takeoff('LSZH', method='track_based')
> ```

My main question is, did you validate your new functions (also pushback()) and how? Did you properly compare them to the existing ones in traffic?

> **Response**
>
> We validated all the new functions we developed. To this end, we proceeded as follows:
>
> - We applied the function to a large number of flights (at least one entire month of observations).
> - Then we randomly selected flights to which we applied our functions as well as the existing ones in

traffic. Subsequently, we both plotted and visually compared the outputs of "our" functions as well as the ones of the existing function. Of all the randomly selected flights, we couldn't find flights with wrongly classified runways / wrongly identified pushback-manoeuvrers (as long as the underlying trajectory data was free of outliers, noise, and gaps).

To explicitly describe this validation procedure, we added the following section to our paper (see Section 2.2.4):

To validate our classification methods for take-off and pushback, we selected a random sample of 1000 *surface trajectories from our dataset. On these trajectories, we applied both our proposed classification algorithms as well as those readily available in the traffic library for pushback and take-off. We then plotted and visually compared the results of our classification algorithms with the legacy ones of the traffic library. In this process, no take-off runway misclassification were observed. However, the accuracy of pushback classification proved to be highly sensitive to trajectory data quality. In particular, noisy trajectories frequently led to misclassification. This is one of the reasons why we decided to remove trajectories subject to significant noise from the dataset used in this study, as already explained in Section 2.1*

You nicely justify and reference most of your data and assumptions but not all, e.g. "start-up process of a single engine takes 60 seconds"

Response

The durations required for the MES and WUP are based on information from Subject Matter Experts (pilots). In the course of our adjustments to Section 2.3 (Estimation of Fuel Consumption) carried out as requested by Reviewer 1, we also explicitly pointed out that the duration of the MES and WUP period are based on inputs from subject matter experts.

"In the same observation period, Zurich Airport Ltd. (FZAG) reported 59,222 departures [19]" You consider only certain aircraft types, is it the same basic types in the FZAG dataset?

Response

The comparison with the 59,222 departures reported by the FZAG refers to all observed aircraft movements contained in our data set (including turboprops, small aircraft, etc.). To emphasize this fact, we added a further note to Table 1. Besides, we modified the entire body of the text of our paper so that it becomes clear that the estimated taxi fuel consumption only refers to a set of selected turbojet aircraft types (as specified in Footnote 2).

What explains that the delta is not consistent across runways? Especially 10 and 34?

Response

We can only speculate about the reasons for this delta. We assume that a large part of the delta are trained noisy trajectories, which are removed from our data set as described in Section 2.1.

Absolute numbers of fuel reduction vs relative? Still absolutely more for longer taxi times?

Response

We changed the corresponding sentence in Section 4 as follows: *In absolute numbers, towing all departing turbojet aircraft to all runways in the observation period from May to end of September 2024 could have saved* 5178.6×10^3 *kg of fuel, while limiting towing to runway 16 departures would have resulted in* *savings of* 1574.0×10^3 *kg.*

Footnote 1 seems at least partly superfluous?

Response

We removed Footnote 1.

Please check the date format, e.g., 01. May, 2024 in Fig. 5/Table 3 is wrong and inconsistent with other date usage in the paper.

Response

The date format is now consistent throughout the entire paper.

"However, as shown in Figure 5, significant fuel-saving potential can already be achieved with a much smaller number of external AGPS units. Nonetheless, a significant part of the fuel-saving potential can still be realized with a much smaller number of external AGPS units in use, making it a more cost-effective approach." These sentences seem to be the same?

Response

We agree. We have simplified the mentioned sentences as follows: *However, as shown in Figure 5, a significant part of the fuel-saving potential can still be realized with a much smaller number of external AGPS units in use, making it a more cost-effective approach.*

https://github.com/m-waltert/osn24_agps is not accessible yet!

Response

Thank you very much for pointing this out. In the meantime, we have switched the repository to public access.

3.3 Response to reviewer 3

Turboprop commercial aircraft were deliberately not considered in the studies, but the authors did not explain why.

Response

We have provided the following explanation for not including turboprops in Section 2.3: *Turboprop commercial aircraft, business jets, and helicopters have been deliberately excluded due to the lack of open-access fuel consumption data for these aircraft types.*

line 224: the authors introduce a threshold value t_{AGPS}. Explain how it is determined.

Response

We added the following sentence to the manuscript in order to specify the range of threshold values considered: *For this purpose, we analysed threshold values in the range of* $t_{AGPS} = \{0, 1, 2, \dots, 20\}$ *minutes. These threshold values were selected on the basis of the observed taxi durations at Zurich Airport as summarized in Table 1 and Figure 3.*

line 307: in future work, the selection of which aircraft are to be chosen taxiing with AGPS, should include considerations regarding feasibility of scheduling of such vehicles.

Response

We agree with you! We will address the question of optimal AGPS usage policies in our (accepted) contribution for the 28th Air Transport Research Society (ATRS) World Conference in Hong Kong from July 1st to July 4th, 2025.

line 328: I suggest adding a couple of references to the recent work related to the case when ground trajectories have gaps:

- Martin Schlosser, Hannes Braßel, and Hartmut Fricke. Analysis of Aircraft Ground Trajectories: Map-Matching with Open Source Data for Modeling Safety-Driven Applications. ICRAT 2024
- Xavier Olive, Manuel Waltert, Ryota Mori, Philippe Mouyon. Filtering Aircraft Surface Trajectories Using Information on the Taxiway Structure of Airports, ICRAT 2024.

Response

We added the references suggested. For this purpose, we added the following sentence to Section 4: *To mitigate the influence of noisy trajectory data and trajectories with gaps in future studies, one could apply certain filter algorithms such as a Kalman filter aligning trajectories with the geometries of taxiways and runways [7] or match-making techniques making use of open-source geospatial airport data [8].*

Again, for future work. Conducting a cost-benefit analysis for the scenarios with different number of AGPS vehicles can naturally complement the studies. Of course, different stakeholders will benefit from the provided improvements, as fairly noted in the conclusions, but the overall environmental goal should always be a good motivator ;)

Response

We will consider the topic of cost-benefit analysis for a future study.

References

[1] Robert Camilleri and Aman Batra. "Assessing the environmental impact of aircraft taxiing technologies". In: *32nd Congress of the International Council of the Aeronautical Sciences*. Vol. 9. 2021.

[2] Milos Lukic, Paolo Giangrande, Ahmed Hebala, Stefano Nuzzo, and Michael Galea. "Review, challenges, and future developments of electric taxiing systems". In: *IEEE Transactions on Transportation Electrification* 5.4 (2019), pp. 1441–1457.

[3] Emanuel Fleuti and Silvio Maraini. *Taxi-emissions at Zurich Airport*. Accessed: February 11, 2025. 2017. URL: https://www.flughafen-zuerich.ch/-/jssmedia/airport/portal/dokumente/das-unternehmen/politics-and-responsibility/environmental-protection/technische-berichte/2017_taxi_study_zurichairport.pdf?vs=1.

[4] Ming Zhang, Qianwen Huang, Sihan Liu, and Huiying Li. "Assessment method of fuel consumption and emissions of aircraft during taxiing on airport surface under given meteorological conditions". In: *Sustainability* 11.21 (2019), p. 6110.

[5] M Ithnan, Ted Selderbeek, Wouter Beelaerts van Blokland, and Gabriel Lodewijks. *Aircraft taxiing strategy optimization*. Tech. rep. Technology University of Delft, the Netherlands, 2013.

[6] Iniobong Edem, Owunna Ikechukwu, Ikpe Aniekan, and Ememobong Ikpe. "The future of conventional aircraft ground propulsion systems in relation to fuel consumption and CO2 emission". In: *International Journal of Thermal and Environmental Engineering* 13.2 (2016), pp. 91–100.

[7] Xavier Olive, Manuel Waltert, Ryota Mori, and Philippe Mouyon. "Filtering aircraft surface trajectories using information on the taxiway structure of airports". In: *Proceedings of the 111th International Conference on Research in Air Transportation (ICRAT 2024)*. Singapore, July 2024.

[8] Martin Schlosser, Hannes Braßel, and Hartmut Fricke. "Analysis of aircraft ground trajectories: Map-matching with open source data for modeling safety-driven applications". In: *Proceedings of the 111th International Conference on Research in Air Transportation (ICRAT 2024)*. Singapore, July 2024.