Journal of Open Aviation Science (2023), Vol.**1**
doi:10.59490/joas.2023.7374

JOAS

EDITORIAL

Reviews and Responses for Aircraft Wake Vortices Affecting Airport Wind Measurements

Authors: Jan Krummen, Lena Noelke, and Raphael Monstein

Reviewers: Antonio Franco, David Lovell, and Raúl Sáez

Editor: Tatiana Polishchuk

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2023.7212

2. Review - round 1

2.1 Reviewer 1

In this paper, the authors address a novel field of research in Aviation, namely, the effects of wake vortices on operational wind measurements at airports. They apply a methodology (strongly based on human intervention) to identify, within a large dataset, instances where a wind measurement was impacted by wake vortices, i.e., what authors call a wake incident or a hit. Afterwards, they perform an exploratory analysis leading to the finding that both the prevailing wind conditions and the category of the involved aircraft are the key drivers for a hit. Then, with this finding in mind, they construct a model to predict the probability of wake vortices hitting the anemometer and present an example of several wake incidents in a row in order to show the magnitude of induced changes in the wind vector. Subsequently, the authors elaborate more on the operational implication of having an artificially pessimistic value as wind measurement. Finally, some reflections on both results obtained and limitations of their analysis are made in the following sections.

Despite the novelty of the research field and the fact that the paper is well structured and well written, one could identify a major weakness, namely, that the paper is mainly focused on the description of the phenomenon under study. Prediction or modelling of affected wind measures and the assessment of possible mitigation measures are either very immature or absent, respectively, so the value of the contribution is not fully shown.

As for the modeling, it is not sufficiently clear why the authors do not propose directly using the DLR P2P wake vortex model to predict wake incidents, but they go for developing an occurrence model. Even assuming there are good reasons for that (which they should highlight in a revised version of the paper), the authors do not follow customary guidelines for both model development and model testing against unseen data. A validation must have been included, for instance, by setting aside a fraction of observations and then testing the model against them (including classification metrics such as accuracy, precision, recall, etc.).

As for the assessment of possible mitigation measures, the authors could have easily explored the effect of increasing the time window of the moving average, and/or the selection of a location that minimizes the hits in a statistical sense (for instance, minimizing the expected number of hits during the considered 15-month period).

Finally, I suspect that the lack of openness in both data and codes could lead to the article's being out of the scope of the journal. This is something I have brought to the attention of the Editor, just in case it could prevent the paper from being published, irrespective of its technical quality.

Taking all these comments into account, my opinion is that, if the previous concerns were successfully addressed (and the Editor did not deem the lack of openness as blocking), the work would be worth it to be published.

2.2 Reviewer 2

This is a well-done study that should interest other airports that suspect they have similar situations. Just a couple of small items that would make things clearer:

1. End of Section 3.2. The authors assert that the A340s and Boeing 777s have a greater impact because of their lower altitude upon passing the anemometer. Elevation data are in the ADS-B data stream, so couldn't this assertion be checked more rigorously?

2. There isn't much discussion on the magnitude and direction of anomalous wind gusts (thus that we know were NOT generated by a passing aircraft). These could create false positives but also make the goal of classifying a hit more complicated. It would be nice to see this discussed more thoroughly.

2.3 Reviewer 3

In this paper, the authors study the effect that wake vortices have on measured wind in airports, and discuss the operational implications of such phenomena. The aim of the research is clear, and the paper is well-written and structured. However, I have some comments and suggestions that I would like the authors to address:

a) I am wondering why the authors discarded the Ilyushin IL76 from the subset containing take-offs of type codes that caused hits (Line 185).

This aircraft lies in the "Other Heavy" category, as defined by the authors. It represents 2.32% of all the aircraft in that category (1 hit for a total of 43 occurrences). On the other hand, if we focus on the B777 family, 4.9% of the aircraft led to hits on the wind sensor.

The difference in the relative amount of hits between both typecodes is not that big and, in my opinion, considering the Ilyushin IL76 to be an outlier might not be a good assumption (one could argue following the same reasoning that the B777 hits are also outliers). Furthermore, are all the aircraft in the "Other Heavy" category Ilyushin IL76? I guess not, right? Being in the same wake vortex category does not necessarily mean that the aircraft are identical, and maybe the Ilyushin IL76 is one of the aircraft types within the Other Heavy category that is more likely to affect the wind sensor measurements.

I recommend the authors reconsider their assumptions or provide another reason to justify not using the Ilyushin IL76 in their analysis. I guess the results would not change that much even if the Ilyushin IL76 is added to the subset but, still, I think it should be added.

b) In Section 4, the authors present a hypothetical situation to show how having an error in the wind measurements due to wake vortices could affect airport operations. I think the example is fine to better understand the implications. However, I think it would be more interesting to see how the effect of wake vortices on wind measurements actually affects the operations of Zurich Airport.

I do not expect the authors to provide a complete analysis focusing on this topic, but I would like to see some discussion and maybe some preliminary analysis in that direction. I think that would really improve the contribution of this paper.

c) Figures 2 and 6 have a very low resolution, I recommend the authors regenerate the figures with a higher resolution to improve readability. In addition, I think the font size of most of the figures is a bit too small, I think increasing it will improve readability. Mainly the font size of Figures 2, 3, 4, 5 and 6.

3. Response - round 1

3.1 Response to reviewer 1

Despite the novelty of the research field and the fact that the paper is well structured and well written, one could identify a major weakness, namely, that the paper is mainly focused on the description of the phenomenon under study. Prediction or modelling of affected wind measures and the assessment of possible mitigation measures are either very immature or absent, respectively, so the value of the contribution is not fully shown.

Response

We tend to disagree with the reviewer on this point. The main objective of this paper is to describe the phenomenon of aircraft wake vortices affecting wind measurements, to show its potential operational impact and to raise awareness of this issue. This issue seems to be unexplored in the literature and we aim to demonstrate the existence and give an indication of the magnitude of the effects. Subsequently, the exploratory analysis of the occurrences (in Sec. 3.1) and the occurrence model (in Sec. 3.2) are intended to aid in developing an intuition for the phenomenon, not as a predictive model. Since our study is very much focused on, and in fact driven by, the operational impact, having a validated predictive model is of little practical use. As mentioned in the conclusion of the paper (Sec. 6), a number of different mitigation measures were discussed with the stakeholders. While a predictive model might be appealing from an academic view, complex technical mitigations such as using a predictive model in operation have significant drawbacks compared to, say, using a different sensor for the take-off clearance.

Notwithstanding our comment above, we agree with the reviewer that this distinction should be made more clearly in the paper. Sec. 3 was edited to reflect this: The aim of the exploratory analysis and the modelling was to develop an intuition for the issue at hand, not to produce a predictive model. Additionally, the already existing text in the introduction (Sec. 1) was extended: "This paper investigates the suspected effects resulting from take-offs on runway 34 at Zurich Airport. The aim is to identify cases where wake vortices have interfered with the wind sensor and to provide a following-up analysis of the frequency of such incidents and the extent to which they affect the measurements. Furthermore, the research delves into the factors contributing to these occurrences and offers insights regarding their potential effects from an operational point of view. These insights are intended to inform the development of mitigation measures to alleviate this issue.

As for the modeling, it is not sufficiently clear why the authors do not propose directly using the DLR P2P wake vortex model to predict wake incidents, but they go for developing an occurrence model. Even assuming there are good reasons for that (which they should highlight in a revised version of the paper), the authors do not follow customary guidelines for both model development and model testing against unseen data. A validation must have been included, for instance, by setting aside a fraction of observations and then testing the model against them (including classification metrics such as accuracy, precision, recall, etc.).

Response

We partially agree with the reviewer on this point. We are about to start working on directly using DLR's P2P to predict the changes in wind measurements due to wake turbulences. However, we would argue that for the stated goal of gaining insight into the issue, our approach is sufficient.

With regards to the comment on the out-of-sample error: The reviewer has a point that this is not addressed in the paper. We intentionally skipped most of the technical details on the GAM and did not report the basic statistics of the model (such as p-values, adjusted R-squared, deviance explained etc.). We also left out the fact GAMs are a form of penalized regression. In the model that we fitted to the data (with the R package 'mgcv') the smoothing parameter that controls the 'wiggliness' of the fitted smooth (i.e., the bias-variance trade-off to avoid overfitting) is estimated from the data. This, to some extent, helps to avoid overfitting. The reviewer is correct in that cross-validation would increase the confidence in the predictive power of the model, but as mentioned earlier, the model is not intended for prediction.

As for the assessment of possible mitigation measures, the authors could have easily explored the effect of increasing the time window of the moving average, and/or the selection of a location that minimizes the hits in a statistical sense (for instance, minimizing the expected number of hits during the considered 15-month period).

Response

Investigating the effects of increasing the window size of the moving average could indeed be interesting and would be easy to implement. However, the two-minute window size used for the surface wind displayed to air traffic controllers is specified in Annex V of EASA Regulation 2017/373 and is unlikely to be easily changed as a mitigation strategy for an airport struggling with this issue. For this reason, it has not been investigated any further.

ATM-ANS (Regulation (EU) 2017/373) ANNEX V – Part MET:
(3) Averaging
The averaging period for surface wind observations shall be:
2 minutes for local routine report and local special report and for wind displays in ATS units;

Identifying an 'optimal' sensor position to reduce the impact of wake turbulence was also of great interest to the stakeholders we've engaged with. However, especially the meteorologists pointed out that repositioning a sensor to mitigate wake turbulence carries the risk of introducing other unwanted effects (wakes and jet exhaust from different runways, local wind differences due to terrain). Determining the ideal position for a sensor, therefore, requires a comprehensive analysis of many factors and was therefore outside the scope of this project. However, this topic is on the list of things we would like to address for the continuation of the work on wakes.

Finally, I suspect that the lack of openness in both data and codes could lead to the article's being out of the scope of the journal. This is something I have brought to the attention of the Editor, just in case it could prevent the paper from being published, irrespective of its technical quality.

3.2 Response to reviewer 2

1. End of Section 3.2. The authors assert that the A340s and Boeing 777s have a greater impact because of their lower altitude upon passing the anemometer. Elevation data are in the ADS-B data stream, so couldn't this assertion be checked more rigorously?

Response

This is a very valid point. An additional Figure (Figure 5), along with the corresponding text, has been added to Section 3.1. The Figure shows a violin plot of altitudes above the runway threshold at which the different typecodes pass by the wind sensor, also distinguishing between hits and non-hits.

A potential explanation for these differences might be the varying strength of wake turbulence generated by different aircraft types. Another aspect to consider is the altitude at which the aircraft passes the sensor. Such variations are due to different aircraft performance, take-off mass, head and tail wind variations and other meteorological conditions. Fig. 5 shows the distribution of altitudes above the runway threshold at the moment of passing the sensor, distinguishing between typecode categories and also between departures that caused hits and those that did not. The data appears to support this hypothesis, showing that aircraft types involved in occurrences tend to pass the sensor at lower altitudes. In particular, all recorded hits correspond to departures passing at altitudes less than 200 metres above the runway threshold. This preliminary analysis strongly indicates a link between the probability of a wake hit and factors such as the prevailing wind conditions and the aircraft type. Notably, the aircraft type's influence seems to be, among other factors, associated with the altitude at which the aircraft passes the sensor. Further explanations about the possible causes are provided at the end of the next section.

2. There isn't much discussion on the magnitude and direction of anomalous wind gusts (thus that we know were NOT generated by a passing aircraft). These could create false positives but also make the goal of classifying a hit more complicated. It would be nice to see this discussed more thoroughly.

Response

We agree with the reviewer, and we covered this indirectly in the last part of the section on visual classification (Sec. 2.2): "The primary justification for using visual classification as a first step is based on the assumption that any event that is not clearly visible in the wind measurements may be considered operationally irrelevant. In addition, the diverse manifestations of hit events within the measurements made the development of an automated identification method very challenging. Although visual classification is to some extent subjective and relies on the judgement of the assessor, this potential bias was minimised by conservative classification and a second step using a wake model which is described further." We would like to avoid going more into detail on this to keep the focus of the reader on the relevant parts of the work.

3.3 Response to reviewer 3

a) I am wondering why the authors discarded the Ilyushin IL76 from the subset containing take-offs of typecodes that caused hits (Line 185).

This aircraft lies in the "Other Heavy" category, as defined by the authors. It represents 2.32% of all the aircraft in that category (1 hit for a total of 43 occurrences). On the other hand, if we focus on the B777 family, 4.9% of the aircraft led to hits on the wind sensor.

The difference in the relative amount of hits between both typecodes is not that big and, in my opinion, considering the Ilyushin IL76 to be an outlier might not be a good assumption (one could argue following the same reasoning that the B777 hits are also outliers). Furthermore, are all the aircraft in the "Other Heavy" category Ilyushin IL76? I guess not, right? Being in the same wake vortex category does not necessarily mean that the aircraft are identical, and maybe the Ilyushin IL76 is one of the aircraft types within the Other Heavy category that is more likely to affect the wind sensor measurements.

I recommend the authors reconsider their assumptions or provide another reason to justify not using the Ilyushin IL76 in their analysis. I guess the results would not change that much even if the Ilyushin IL76 is added to the subset but, still, I think it should be added.

Response

We agree with the reviewer that the justification to exclude the one IL76 from the GAM was insufficient. We edited the paper to better explain the reasoning in Sec. 3.2.

The Ilyushin IL76 was excluded from the GAM since it is a rare guest at Zurich airport and is with its single departure operationally not as relevant as, say the Airbus A340 with 550 departures. One could include the whole group 'Other Heavy', but this makes little sense due to its heterogeneous composition.

b) In Section 4, the authors present a hypothetical situation to show how having an error in the wind measurements due to wake vortices could affect airport operations. I think the example is fine to better understand the implications. However, I think it would be more interesting to see how the effect of wake vortices on wind measurements actually affects the operations of Zurich Airport.

I do not expect the authors to provide a complete analysis focusing on this topic, but I would like to see some discussion and maybe some preliminary analysis in that direction. I think that would really improve the contribution of this paper.

Response

We absolutely agree with the reviewer that a more in-depth analysis of the operational consequences on the airport would be worthwhile. Unfortunately, such an analysis would be well beyond the scope of this paper. If such an analysis is to be based on more than just anecdotes, we would have to analyze the ground movements to detect runway changes where an aircraft switched from runway 34 to another one. The quality of the ground coverage in Zurich for the 15 months used in this study is mixed. This would make it hard to obtain reliable results with the data from OpenSky. We are, in collaboration with the airport, about to install an ADS-B receiver at the airport to improve the ground coverage in the future. We expect the sensor to be up and running in Q1 of 2024.

c) Figures 2 and 6 have a very low resolution, I recommend the authors regenerate the figures with a higher resolution to improve readability. In addition, I think the font size of most of the figures is a bit too small, I think increasing it will improve readability. Mainly the font size of Figures 2, 3, 4, 5 and 6.

Response

All the mentioned figures have been regenerated at a higher resolution and with increased font size to increase readability.

4. Review - round 2

4.1 Reviewer 1

The revised version provided by the authors clarifies the doubts raised by this reviewer. Hence, my recommendation is to accept the paper.

4.2 Reviewer 3

I am still not fully convinced with the author's response, especially regarding my second comment (comment "b").

The analysis of how wake vortex can affect the wind measurements is fine, no complaints about that. However, I think section 4 is not enough to show the operational implications of such phenomena. This is just a "tailored" example that could or could not represent reality at Zurich Airport.

Even if the wake vortex are affecting wind measurements, is this then translating into effects on operations? Or are changes to an alternate runway or offloading cargo (Line 282) rare occurrences?

Therefore, I leave the decision to the Editor regarding this issue: is the analysis of the wake vortex effect on wind measurements enough contribution for the paper to be published in JOAS? Or does the editor think that a more detailed operational implication study should be also included?

As I said, the analysis of the wake vortex on wind measurements is very good, the methodology is clear and the results presented are fine. If this is enough to be published in JOAS, I would be ok to accept the submission.

5. Editorial notes

The OpenSky symposium chairs and JOAS editors acknowledge that the tool utilized in this paper cannot be openly shared. An exception has been granted, as the authors had communicated from the outset their efforts to obtain permission for sharing the tools and data prior to the OpenSky conference.

Moreover, we recognize the authors' substantial efforts in negotiating with the data provider, which enabled them to make the originally private data publicly available for this paper.