Journal of Open Aviation Science (2024), Vol.**2**
doi:10.59490/joas.2024.7368

JOAS

EDITORIAL

Reviews and Responses for Hidden Markov Models and Flight Phase Identification

Authors: Rémi Perrichon, Xavier Gendre, and Thierry Klein

Reviewers: Emy Arts, Nicoletta Fala, and Junzi Sun

Editor: Mayara Condé Rocha Murça

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2024.7269

2. Review - round 1

2.1 Reviewer 1

This work discusses the approach of using Hidden Markov Models to segment a flight's ADS-B data into its flight phases. This approach shows benefits compared to similar work as it allows to us identify flight phases with lesser a priori knowledge and reduces the need for pre-processing. This work is written using clear language, and a clear structure and provides a solid mathematical basis.

The authors state that the ability to identify the probability of a point belonging to a specific phase is an advantage of this approach as well however, it is not clearly defined why this is of benefit (e.g. compared to probability-like values that can be obtained from the degree of membership).

Furthermore, this work allows us to identify a larger number of phases compared to the state-of-the-art fuzzy logic approach. However, the nomenclature of the phases is inconsistent with other research.

The descent phase is not present and is referred to as the approach, however, according to the ICAO ADREP nomenclature the approach is only the final part of the descent. The accuracy of the landing phase is discussed on several occasions, however, the landing phase as such is not identified, instead, the rollout phase seems to be referred to as landing. According to the ICAO ADREP nomenclature, the rollout is a subphase of the landing phase after touch-down.

Although this approach has the benefit of reducing pre-processing efforts, it is not suitable for large-scale applications in ADS-B data as of date. This is due to the fact that ground coverage is still relatively scarce, which means that some flights might have some or all of the lower altitude phases, but others do not. When dealing with the decoding methods offered this could lead to great computational efforts when running on a big scale.

Further minor comment: the colour scheme for the 2D image and the 3D image in Figure 1 are different, the 3D image has no legend.

2.2 Reviewer 2

This paper uses Hidden Markov Models to segment and identify phases of flight in-flight data recorder files from an airliner. The authors used open-access data from NASA DASHlink Tail 687, which is publicly available, and provided their algorithms to reproduce the work. Overall, the research is of high quality and the paper is both well-written and well-explained. My questions and comments to improve the communication of the research are as follows.

1. We should be further discussing how well this method would work with other operations, beyond airline flights. The authors have included the example of a helicopter operation, using ADS-B data. However, how would the method work on a small general aviation aircraft, where the pilot is not relying on autopilot and therefore introducing more variability in the parameters studied, especially the rate of climb? For example, in the time series in Figure 1, the flight appears to have five or six cruise segments, four or five descents, and three climbs, depending on how the algorithm treats the small level-off towards the end of the flight. Vertical speed is very noisy and instantaneous, and smaller aircraft climb/descent at slower rates, i.e. rates which are more similar to “level flight.” Both differences may make the task a lot more challenging. Is the high accuracy presented in this paper mainly due to the method, or because of the smoothness and nature of airline operations? As a follow-up question, then, how does the HMM method compare to other methods in the literature where researchers have used NASA DASHlink data?

2. The authors evaluate the accuracy of the end result using the ACMS-derived phases of flight as ground truth. However, if the ACMS already uses algorithms or exceedances to get the phase of flight that we accept as truth for this kind of operation, then what is the purpose of using any kind of other method which we can train to only do at most as well as the ACMS method?

3. ADS-B data was more central to the paper in the abstract and introduction but then the authors navigated away from it. I think the authors should change the introduction and keywords to remove the ADS-B focus since the work does not use ADS-B data until the helicopter example.

4. The authors’ discussion and use of multiple performance metrics is correct and I appreciate that they focused on the transitions as well as the data points themselves.

5. While some transitions are less likely (i.e. approach to climb), they are not completely unlikely. For example, an incoming flight could result in a go-around initiated during the approach, which would result in a climb following an approach segment. How can we account for such situations when evaluating accuracy?

6. The section on choosing initial values for the state-dependent density (page 9 lines 282-285) could be better explained to help the reader understand where the uniform distribution parameters come from.

7. While the paper is generally very well-written, it could still use some copy-editing. The use of the phrase “to be precise” is repetitive. There are some mistakes throughout the paper that will be corrected through careful proofreading (e.g. page 6 line 212 “the naïve approach separately consider,” page 4 line 129 “much points”).

2.3 Reviewer 3

The paper revisits the phase identification problem of in-flight data processing. It proposes a new HMM (Hidden Markov Model)-based approach to leverage dependencies in states across a sequence of flight data points. The HMM approach demonstrates higher accuracy compared to previous studies, as evidenced by the testing flight data used in this study. The paper is well-written and the concepts and examples are clearly articulated.

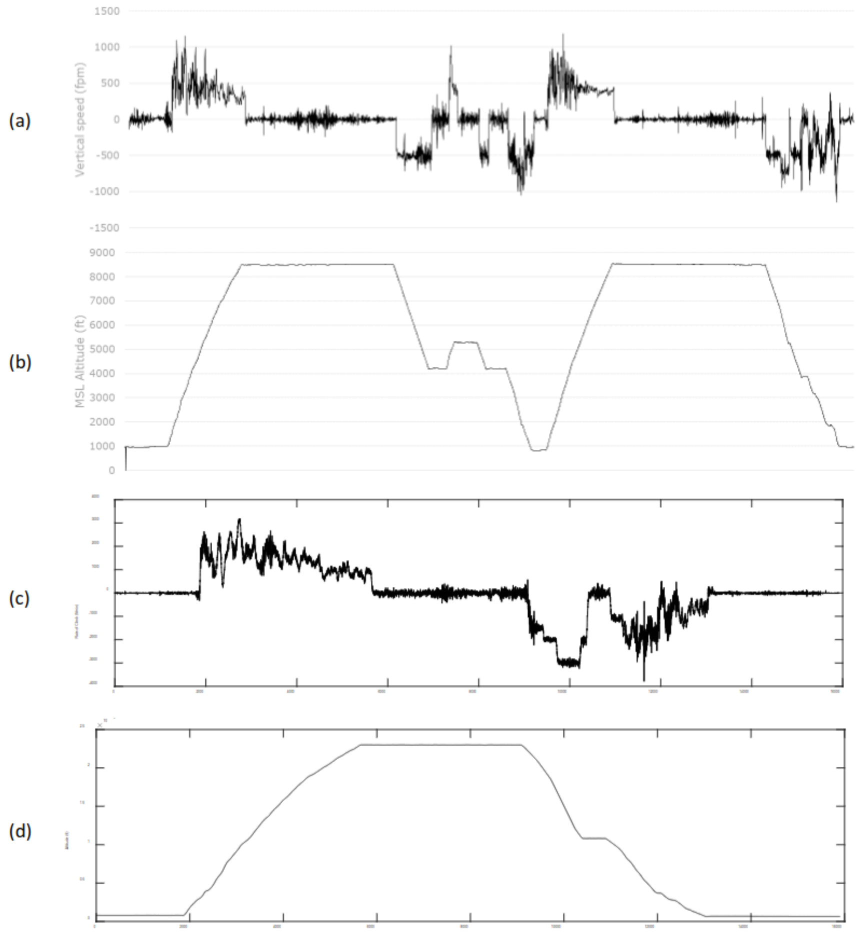

Figure 1. Data comparison for a general aviation (a, b) cross-country flight and a flight from the NASA DASHlink Tail 687 set (c, d). (a, c) show the vertical speed or rate of climb time series directly from the data files, and (b, d) show the altitude profiles for the two flights.

The following are my minor comments for further improving the paper:

1. The level-off phase during climb or descent is not explicitly addressed. Could the HMM approach be applied to identify level-off phases? Maybe this could be elaborated on in the discussion section

2. In the Data section, could more details be provided on the update rate of these flight data? Resampling may introduce inaccuracies that complicate phase identification.

3. Line 206: the equation lacks numbering, and the notation **1** should be clarified for readers.

4. Line 261: "assumption is relaxed is the following" appears to be a typographical error.

5. For Figure 2, I suggest replacing state numbers with state names for clarity.

6. Line 304: It states "unlikely transitions: going (directly) from climb to approach." However, in short flights, it is feasible to transition directly from climb to descent (approach) without a cruise phase.

7. Figure 3 suggests that the Naive segmentation approach also performs adequately in this case. A more detailed comparison between these two approaches would be beneficial.

8. Figure 4: Are level flight segments considered (or excluded) in the F1/accuracy calculation for the fuzzy logic approach?

9. Figure 6: I suggest replacing state numbers with actual state names would enhance clarity.

10. Line 401: The sentence "Indeed, ground speed primarily serves to segment the ground phase, a phase that is not of interest at this stage of the analysis," needs clarification or elaboration.

11. Section 6.2: It would be important to highlight that the probability of a phase is conditioned on previously estimated states.

12. Section 6.3 presents an innovative application of HMM in identifying arbitrary flight phases in helicopter operations, which is commendable.

13. In this helicopter use case, are speed and heading also considered in state determination?

14. Line 448: State 2 appears to represent the helicopter's circling, as indicated by the constant change in track.

15. Lines 492-495: Please replace the authors with the actual names of the authors.

16. Line 500: For the reproducibility statement, consider creating a persistent repository with a DOI, for example, using platforms like Zenodo or Figshare.

3. Response - round 1

We would like to first thank the reviewers for their attentive reading of our contribution as well as the editor who allowed us to make some corrections.

3.1 Response to reviewer 1

1. Degree of membership v. probability

Response

The degree of membership can only provide an ordinal measure of certainty: at each point, one can rank the states (flight phases) from the most 'plausible' to the least 'plausible'. Yet, from one point to another, a given state can be the most plausible without having the same degree of plausibility for each point. Having probabilities ensures that the degree of plausibility reaches a certain value at each point of the flight. One can then establish rules such as 'I label a given point in the flight if the most plausible state exceeds a threshold of 95%; otherwise, the data is too noisy to draw conclusions'. Conceptually and practically, it seems advantageous to measure the uncertainty of segmentation « in the same way » for all points of the flight.

2. Nomenclature of the phases is inconsistent with other research.

Response

Thank you for the remark. It has been taken into account in the revised version (lines 148-150)

3. ICAO ADREP nomenclature.

Response

Thank you very much for the link to the nomenclature. We believe that it now has been taken into account in the revised version (lines 148-150)

4. ADS-B and coverage

Response

It is true that the good performance of HMMs relies on the assumption that the initial sampling allows for the detection of each flight phase. For a given sample of ADS-B flights, if there are good reasons to believe that certain flights are not observed in their entirety, it is then preferable not to specify the number of states in advance, following the approach used for helicopter flights. Unlike fuzzy logic, it is true that our method is not parallelizable.

It has been taken into account in the revised version (lines 472-475)

5. Colors, legend

Response

Thank you, we used the right colors in the revised version. Yet, we did not add a second legend as it is now clear that the colors are the same.

3.2 Response to reviewer 2

1. Generalization

Response

It would be very interesting to apply HMMs to small general aviation aircraft flights. It is likely that the task is, indeed, more challenging. Yet, provided a good model specification, we believe that HMMs may provide an accurate segmentation. The main problem is the availability of such a dataset as well as a «ground truth» to quantitatively assess performances.

The high accuracy presented in this paper is both due to the method and to the smoothness and nature of airline operations.

To our knowledge, Liu et al. (2020) are the only ones who used NASA DASHlink data for flight phase identification (see references). The main problem is that they used more input variables, such as the engine fan speed, which could distort the comparisons. Additionally, the performance metrics differ.

2. ACMS

Response

The only way to study the performance of a segmentation model is to have a set of labeled flights. These labels can be assigned by domain experts or by an existing algorithm for which there is a consensus on its ability to be very close to the 'true' labels. For flight phases, it is entirely accurate to say that the ACMS itself is based on a set of algorithmic rules and thresholds. When ACMS-derived phases are available, there is indeed no interest in using another segmentation method if we trust the ACMS-derived phases. Here, we use the ACMS as a ground truth from which it is possible to obtain quantifiable performance measures. The reasoning is then as follows: 'if we did not have this ground truth, our segmentation would be satisfactory at X%.'

3. ADS-B v. FDR

Response

We have modified the keywords to make it clear that FDR data is at the core of the article.

While we have flight phases for FDR data (ACMS), this is not the case for ADS-B data. The paradox is as follows: it is challenging to find ADS-B datasets for which flight phases are known (and this is necessary to compare segmentation methods). The article then focuses on FDR data rather than simulated data (as often found in the literature), hoping that segmentation performance will be roughly similar. Since FDR data has a higher sampling rate than ADS-B data, the linear interpolation we perform (see Data section) adds more realism.

4. Performance metrics

Response

Thank you

5. Unlikely transitions

Response

We believe it is up to the end-user to specify what should be considered an 'improbable transition'. Additionally, it is possible to set a minimum number of allowed improbable transitions.

6. Initial values

Response

Initial values may be chosen by hand. More conveniently, it is often advantageous to set initial values within intervals that are likely to contain the true values. This leads to faster convergence. It would be challenging to justify drawing these initial values from any distribution other than a uniform one because we do not have information on this matter. The parameters of the uniform distribution are chosen based on what is to be expected from a typical commercial flight. The higher the transition parameters, the more likely it is to remain in the same flight phase.

7. Copy-editing

Response

Mentioned mistakes have been taken into account in the revised version.

3.3 Response to reviewer 3

1. Identification of level-off flight phases

Response

In the fuzzy logic approach, level flight subphases are identified which is not the case in our article and in some other contributions. From a conceptual point of view, HMMs may totally detect the level-off phases during climb or descent. To achieve so, one should specify a constrained multivariate HMM with 5 states that are labeled « ground », « climb », « descent », « cruise », and « level flight ». The model's good performance crucially depends on the input variables. As a first guess, one could use variables that are part of the fuzzy logic: - The altitude, potentially encoded as a binary variable (to differentiate the cruising phase from the level flight phase) - The rate of climb (to differentiate the climb from the descent) - The speed, encoded as a binary variable (to identify the ground phase).

HMMs are to be seen as a probabilistic extension of fuzzy logic that incorporates the temporal dependence between points in the trajectory.

Note that it has been taken into account in the revised version (lines 505-508)

2. Original sampling rate and the impact of resampling

Response

In the original data, the sampling rates vary depending on the variables (1 Hz for longitude, and latitude, 4 Hz for the pressure altitude and ground speed). These details are added in the Data section, as requested. To address this challenge, linear interpolation is suggested for its simplicity. Each flight is resampled to 1,000 points. This procedure is indeed not neutral and can, in theory, impact the model's performance. Linear interpolation acts as a pre-smoothing of the data, simplifying its erratic nature. The effects of this interpolation are numerous and challenging to quantify without a more thorough study. If the interpolation is too coarse, it is possible that information may be lost, and the results of segmentation procedures may be less accurate. Maintaining a high temporal granularity can lead to a non-negligible computational cost for HMMs. Outliers are also a concern. In practical terms, 1,000 points appear to be a good compromise between these two pitfalls (given our sample). The existence of an optimal resampling procedure is not addressed in this contribution, and we believe that it could be the subject of a separate study.

Note that it has been taken into account in the revised version (lines 151-152, lines 158-162)

3. Numbering, clarification

Response

It has been taken into account in the revised version.

4. Typo It has been taken into account in the revised version.

5 & 9. State numbers v. state names

Response

In Figures 2 and 6, the state numbers are indicated rather than the flight phase names. It is suggested to replace state numbers with actual state names. We made this choice because the reader needs to be aware that HMMs are unsupervised models. Unlike classification where we would have labels to learn, HMMs do not « learn » in the strict sense. The transition from the states of the model (theoretical, without a priori meaning) to flight phases (practical, with operational significance) is immediate due to imposed constraints, but they are not the same thing. This is particularly clear for helicopter flights where we have no a priori knowledge of the states. Thus, we originally kept the state numbers for the diagrams in Figures 2 and 6 (theoretical) and indicated the flight phase names for the plots.

6. Short flights may go from (directly) from climb to approach

Response

It has been taken into account in the revised version.

7. Results for the naive approach

Response

Yes, a more detailed comparison between approaches is beneficial. This is why we go from the visual result (fig 3 on a specific flight) to performance metrics (fig 4, whole sample).

8. Are level flight segments considered (or excluded) in the F1/accuracy calculation for the fuzzy logic approach?

Response

Yes, we take into account level flight segments (see Appendix 1). Yet, we believe that there are several ways to do it.

10. Clarification

Response

It has been taken into account in the revised version.

11. Probability of a phase

Response

Actually, the probability of a phase is not conditioned on previously estimated states (at least for local decoding). From Equation 11, one may see that the probability of belonging to state i depends on the observations $x_1, \ldots, x_T$, but not on previously estimated states. No errors are « accumulated » in the decoding process.

12. Innovative application

Response

Thank you

13. In this helicopter use case, are speed and heading also considered in state determination?

Response

Speed is taken into account (see line 452 of the revised version) but track angle is not per se. It would necessitate to introduction of circular distributions (such as the Von Mises distribution) which seems somewhat tangential to the core of the article. Instead, we use the longitude and latitude first differences that should contain similar information.

14. Additional information

Response

It has been taken into account in the revised version.

15. Names

Response

It has been taken into account in the revised version.

16. Persistent repository

Response

Thank you for the suggestion. The link to the Zenodo repository is:
https://doi.org/10.5281/zenodo.10512234