

Clustering Climb Profiles for Vertical Trajectory Analysis

Matthias Poppe, Jörg Buxbaum
DFS Deutsche Flugsicherung GmbH
Langen Germany
matthias.poppe@dfs.de

Abstract – Aircraft climb trajectories show a wide performance range and are difficult to predict. This study provides a systematic framework which clusters the climb profiles, collected during 2019, for all aircraft types from two major German airports. Clusters with similar climb performance could be distinguished with K-Means for up to 16 clusters. The cluster representatives could be used for prediction of the trajectory profile and the root mean square error of the deviation at FL250 could be reduced to the error of the cluster members if the correct cluster was predicted.

Keywords - Air Traffic Control, trajectory prediction, clustering, climb profiles

I. INTRODUCTION

The aircraft climb performance and the prediction of the aircraft trajectory are key enablers for future ATM trajectory based operations. The climb profile depends on many factors of different nature, being it the aircraft intent, the aircraft current state, the atmospheric conditions, and last but not least human controller and pilot interventions. This results in an operational and technical uncertainty [1] where many of these factors are not known or can only roughly be estimated.

As the aircraft mass is a critical factor and one of the major uncertainties to the climb performance, several attempts have been made to estimate the aircraft take off mass (e.g. [2][3][4]). With a better and more accurate mass estimate, parametric approaches like BADA [5] or [6] will provide more accurate climb trajectories. Alternatively, non-parametric approaches [7][8] try to estimate the aircraft trajectory not only in the climb phase.

Despite all the progress made during the last years, the resulting uncertainty may still be too high for certain safety critical applications like Conflict Detection or Controller Decision Support tools.

The focus of this study is to analyze the climb performance based on a clustering approach as it is proposed for trajectory

prediction e.g. in [9] or applied for air flow identification [10]. All aircraft types within a performance range typical for civil transport aircraft from Take Off to Flight Level FL 250 will be considered. This limit has been chosen because it is well below the typical crossover altitude¹ for most aircraft types. The resulting look ahead time is about 6-8 minutes. The question is to which extent this analysis can be used for predicting the future climb profile based on the first observed climb segment.

Section II explains the basic approach with K-Means. After discussing the data pre-processing in section III, sections IV and V and provide the obtained results and how to use them for trajectory prediction.

II. APPROACH

A. Unsupervised Learning

The field of unsupervised learning is used to discover structures in big data sets and try to extract useful information without knowing the corresponding output variable. On the other hand, it is hard to assess the results obtained from unsupervised learning methods since there is no universally accepted mechanism for performing cross validation. Therefore, the time to reach FL250 has been chosen as a benchmark for this study.

Clustering is an exploratory data analysis to understand the magnitude of the problem. It is based on a distance or similarity function. Multiple clustering approaches are known, e.g. based on density functions [11][12]. However, the authors have chosen the well-known k-Means partitioning algorithm [13] based on the Euclidean distance function. As we are interested in the climb behavior, the high-dimensional trajectory data were reduced to flight level over time and sampled every 10 FL. A vector of constant length has been derived which is required for using the Euclidean distance. One advantage of k-Means is that all input data will be assigned to a cluster, and on the same time a cluster representative will be provided which can be used for further analysis.

¹ The crossover altitude is the altitude at which a specified CAS and Mach value represent the same TAS value. Above this altitude the Mach number is used to reference speeds.

B. K-Means Clustering

In K-means clustering [14][15], we seek to partition the observations into a pre-specified number of clusters k . Each cluster C_k contain $|C_k|$ number of observations and each of the n observations is assigned to exactly one cluster:

$$C_1 \cup C_2 \cup \dots \cup C_K = \{1, \dots, n\} \quad (1)$$

The clusters are non-overlapping, two different clusters are disjunct:

$$C_k \cap C_{k'} = \emptyset \text{ for all } k \neq k' \quad (2)$$

The basic idea is to minimize the within-cluster variation, i.e. the amount to which extent the observation within a cluster differ from each other, in our case the squared Euclidean Distance:

$$\text{minimize}_{C_1 \dots C_K} \left\{ \sum_{k=1}^K \frac{1}{|C_k|} \sum_{i,i' \in C_k} \sum_{j=1}^p (x_{ij} - x_{i'j})^2 \right\} \quad (3)$$

The K-means algorithm solves the above equation and finds a local optimum (not necessarily the global optimum). Therefore, the algorithm will run several times with different initial configurations. One disadvantage of K-Means may be the distortion of clusters due to the presence of outliers since the squared distance is not very robust to perturbations.

III. DATA PREPARATION

Selected Mode S Enhanced Surveillance Data (EHS) and the Mode C altitude information were used from 2019 to extract the required information for data preparation.

The Mode-S EHS data [16] contain the following Binary Data Selector (BDS) register:

1. BDS 4.0 - Selected Vertical Intention
2. BDS 5.0 - Track and Turn Report
3. BDS 6.0 - Heading and Speed Report

An overview of the used raw data for this study can be seen in Table 1.

Index	Abbreviation	Explanation	Unit
0	t	Time	sec
1	FL	Flight level (Mode C)	FL
2	AID	Aircraft ID (24 bit address)	-
4	FCU	FMS Selected Altitude	feet
5	GS	Ground Speed	knots

² At the transition altitude, the aircraft switches from the use of local barometer derived altitudes to flight levels. In Germany, this altitude is reached at about 5000ft (depending on local conditions).

8	IAS	Indicated Airspeed	knots
10	RoC	Inertial Vertical Speed	ft/min

Table 1: Extracted data items

A climbing flight was defined when it shows a positive flight level change over time and when no apparent restrictions in terms of Level off segments are available. This means:

$$\begin{aligned} & \max_t \{FL(t) > 255\} \quad \text{AND} \\ & \operatorname{argmin}_t \{FL(t)\} < \operatorname{argmax}_t \{FL(t)\} \quad \text{AND} \\ & \forall FL(t) < 255: \quad FCU(t) - 100 \cdot FL(t) \leq \omega \end{aligned} \quad (4)$$

It was chosen to set $\omega := 1000$ [feet]. In total, more than 45.000 climbing flights from Frankfurt (EDDF) and Munich (EDDM) airport fulfilled the criteria and were selected for this study. The original Mode S and ADS-B data were extracted and downloaded with the 'traffic' toolbox [18] from the OpenSky [19][20] website during the period from February 2019 to August 2019 (approximately 104 days EDDF departures, 93 days EDDM departures).

Caveat: Mode C altitudes below the transition altitude² are also given in respect to 1013.25 hPa (and not to QNH) in our data so no deviations to the Flight Level information were obtained from these circumstances³. Yet, differing altitudes at Take Off due to varying airport elevations are not respected. Due to the selected quantization (see next section), the impact should be negligible.

A. Vector Extraction

For the identification of similar climb profiles, we need to cluster these profiles, i.e. we need to consider the flight level change over time. Figure 1 shows a synthetic profile and how the feature vector for each flight was extracted.

In order to get a vector of constant length for a application of the Euclidean Distance, we divided the climb phase from FL30 up to FL250 in 23 segments and recorded for each segment the time t_k needed to climb to the respective flight level. This has the advantage that the vector length is constant and independent of the individual flight time to FL250 which may vary considerably (please refer to Figure 4). FL30 has been selected as a lower bound in order to reduce noise effects from the initial take off phase and the different airport altitudes.

All data were scaled to a range between 0 and 1 (in the following figures the absolute values are provided for a better interpretation).

³ The QNH is the barometric pressure adjusted to mean sea level. It is dependent on current weather conditions and the difference to 1013.25 hPa can be large dependent on the current weather situation [17].

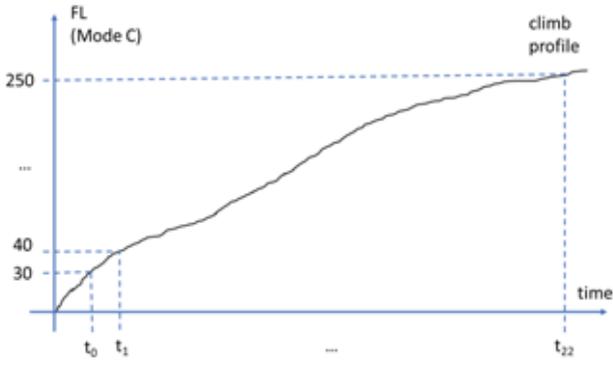


Figure 1: Extraction of the climb feature vector

The calculation of the times t_k is derived from:

$$t_k = \operatorname{argmin}_t \{FL(t) > 10 \cdot (k + 3)\}, k = 0, 1, \dots, 22 \quad (5)$$

The resulting feature vector describes the flight climb behavior to FL250. The starting point of FL30 and the division of 10 FL means a quantization which eliminates noise in the profile and eases the later clustering and processing of the data.

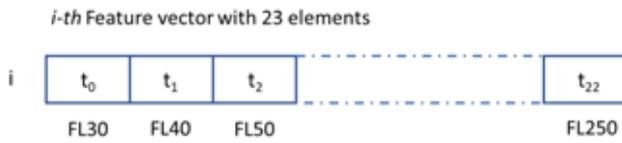


Figure 2: Structure of the feature vector

For a flight i , the vector structure is shown in Figure 2. Each element of the vector thus contains the time needed to reach the respective Flight Level.

In Figure 3, the resulting piecewise constant approximation of the climb time as a function of the Flight Level is outlined. The basis function t_m can be described with

$$t_m(FL) = I(L_m \leq FL_k < U_m) \quad k, m = 0, \dots, 22 \quad (6)$$

The upper and lower bounds L_m and U_m has been set to

$$L_m = U_m = 5 FL \quad (7)$$

As a reference for the spread of the time to climb to FL250 and in order to get a sense for the magnitude of the problem, the distribution of t_{22} is shown in Figure 4. All aircraft types are included, and the distribution has a mean of 707 seconds and a standard deviation of 108 seconds. The distribution has been cut at 450 seconds and at 1100 seconds to eliminate outliers in the data (because of the sensitiveness of K-Means regarding outliers). This results in an average Rate of Climb (RoC) after

the cut of about 3700 feet/minute and 1400 feet/minute respectively.

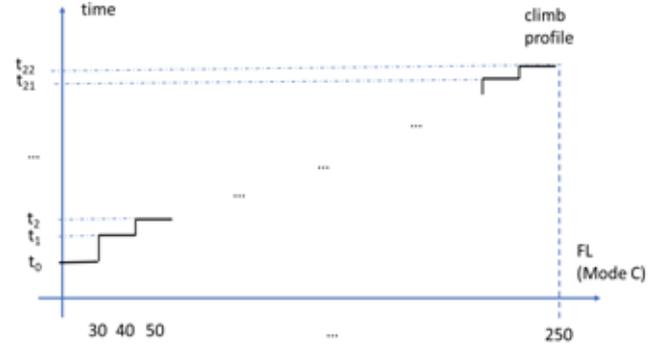


Figure 3: Piecewise constant FL segments

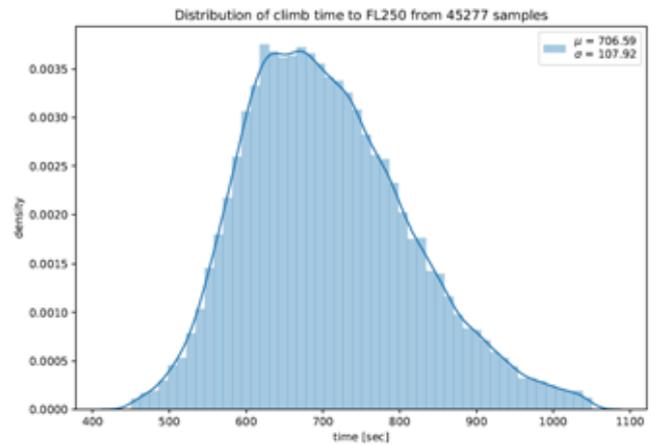


Figure 4: Climb time to FL250 (EDDF, EDDM)

IV. RESULTS

Because we are interested to which extent we can use the clustering approach for trajectory prediction, the clusters of the feature vector are presented with respect to the time from Take Off to FL250. In total, 35000 flights have been used for clustering while the remaining 10000 flights were used for prediction (they were not part of the clustering).

The number of clusters k is a free-to-choose parameter in K-Means. We need to find the number of clusters that show a small cluster variance but at the same time we should be able to predict the correct cluster for a new query. The challenge is to assign the correct cluster based on partial information, e.g. only from a feature vector with elements t_0 to t_{11} . Figure 5 and Figure 6 show the sample distribution for six and for 10 clusters. With more clusters, the overlap between cluster members increases while the general structure remains similar. Most members are in the clusters with a cluster mean around 600 to 700 seconds.

From these clusters, it can be seen that in particular the range between 500 and 600 seconds and between 900 and 1000 seconds will benefit from a higher number of clusters because the cluster range becomes narrower. For example, the intra-cluster standard deviation for six clusters with mean 549 seconds is 34 seconds while this value decreases for 10 clusters to 28 seconds.

In order to get a good value for the number of clusters, *Figure 7* shows the total sum-of-squares within the cluster for different numbers of clusters. This diagram shows a sharp decrease from one to six clusters while the curve flattens for a higher number of clusters.

Based on the shape of this curve, the examples provided refer either to six or to 10 clusters.

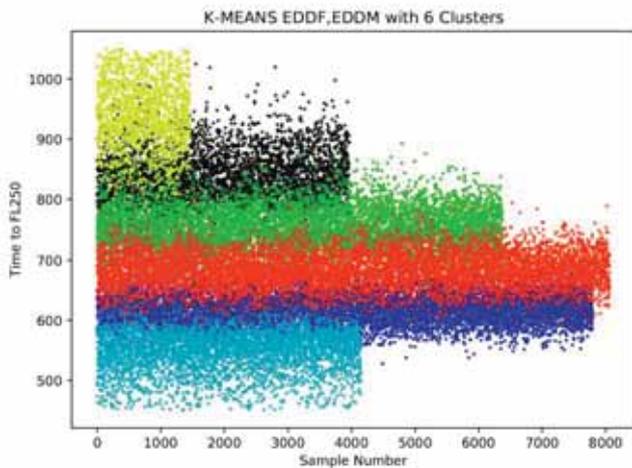


Figure 5: Distribution of samples to six clusters

In *Figure 8*, the cluster representatives C_{rk} are plotted for $k = 10$ clusters. Here, one of the representatives is crossing two other cluster centroids at sample numbers 10 and 17. This means that the climb profiles may exhibit different behavior. A low initial climb performance changes to a later high climb performance (curve flattens in *Figure 8*) when compared to the other characteristic profiles.

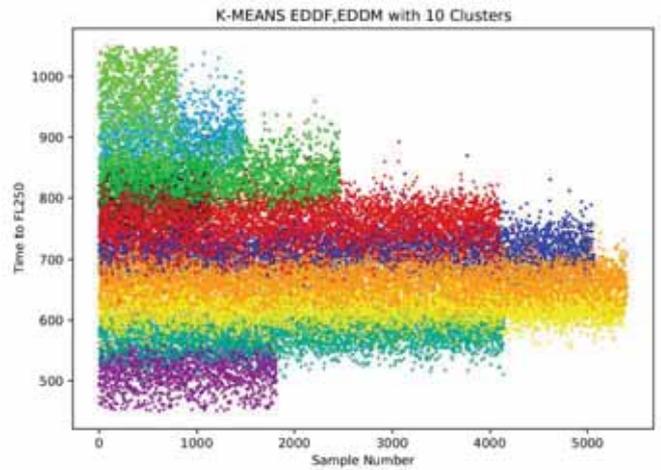


Figure 6: Distribution of samples to 10 clusters

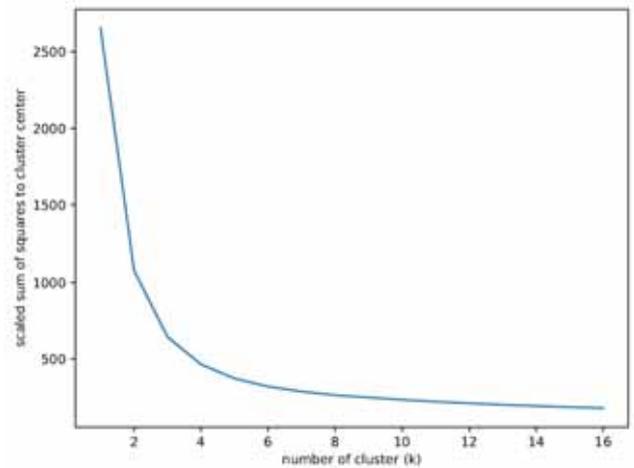


Figure 7: Total within-cluster sum of squares

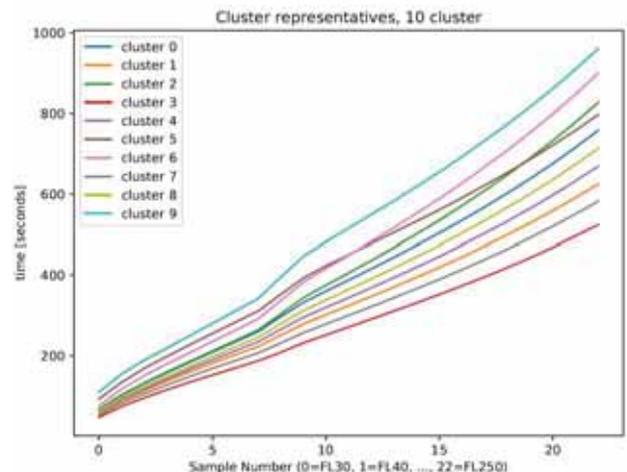


Figure 8: Cluster representatives for 10 clusters

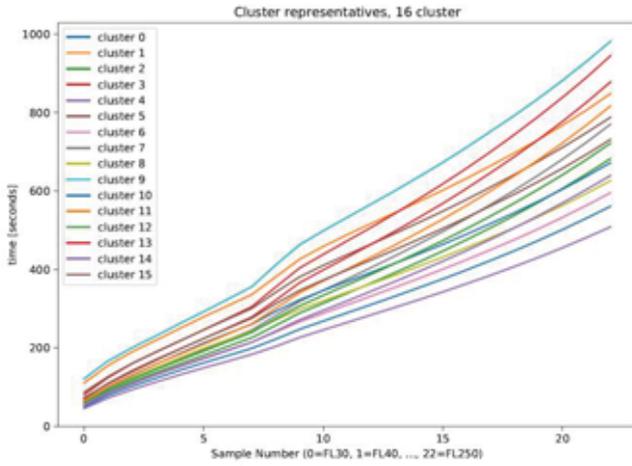


Figure 9: Cluster representatives for 16 clusters

This is not the case for less than 10 profiles, and the number of crossing representatives C_{rk} increases with the number of clusters. *Figure 9* shows the situation for 16 clusters. Only the outer cluster representatives are not crossed by others in this case. It is not clear what causes these overlaps and whether this is due to human controller or pilot intervention (operational uncertainty) or whether this is caused by the Flight Management System or the Cost Index, dependent on the aircraft mass (technical uncertainty) or may be atmospheric impact.

Another observation is that the cluster representatives are closer together at lower FL and spread more and more towards higher altitudes.

To better understand the results of the clustering, the mean and standard deviation were computed for all cluster members with respect to the time to FL 250, i.e. the vector element t_{22} . Table 2 shows an example for six clusters (the color corresponds to the colors in *Figure 5*).

Cluster	Mean	Std. Dev.	Number
0 – light blue	548.9	33.6	4267
1 - dark blue	619.0	26.0	8120
2 - red	686.1	29.0	8411
3 - green	756.2	33.7	6551
4 - black	837.9	44.7	3996
5 - yellow	940.7	54.4	1455

Table 2: Mean and standard deviation of all class members at FL250 for six clusters (in seconds)

While the standard deviation is around 30 seconds for the four most populated clusters, the two clusters with the highest mean (low climb performance) have a standard deviation around 45 and 55 seconds respectively.

V. ANALYSIS OF PREDICTION CAPABILITY

After having analyzed the different climb behavior with K-Means clustering and having found cluster representatives with a standard deviation in the order of 30 seconds, it would be beneficial to use this a-priori knowledge for predicting the time to FL250 for new yet unknown trajectories.

With the full knowledge of a new feature vector, the correct cluster can be found by choosing the one with the minimum squared Euclidean distance (ED) to C_{rk} . The challenge is to assign the correct cluster k with only partial knowledge of the feature vector, e.g. when only t_0 to t_{10} are known (i.e. the prediction takes place at FL130). *Figure 10* shows the percentage of queries that are assigned to the correct cluster centroid dependent on the knowledge of the length of the feature vector, i.e. when the prediction has been taken place. In addition, the number of clusters k has been varied from 2 to 10. A lower number k results in a higher number of correctly assigned queries to the correct cluster. For example, with a prediction at FL150 and 6 clusters, a bit more than 60% of all queries could be assigned to the correct cluster (based on 10000 queries).

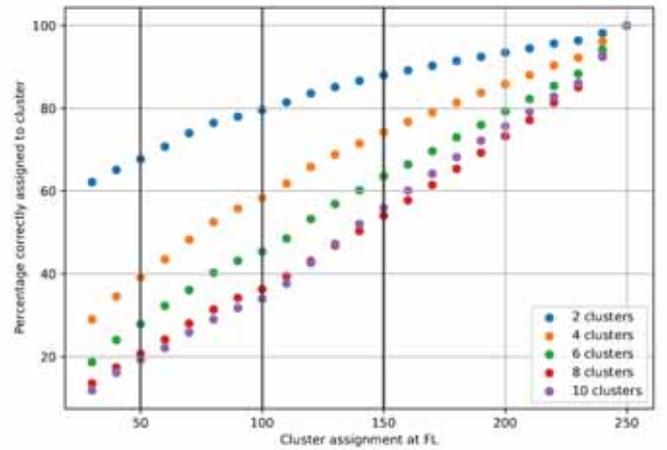


Figure 10: Correctly assigned clusters as a function of the flight level and the total number of clusters

In general, the percentage can be slightly increased to 70%, if the distance of the last available vector element (e.g. t_{10}) to the corresponding 10th cluster element is compared, instead of using the squared ED from t_0 to t_{10} . Reason for this seems to be the more noisy climb behavior in the beginning.

The resulting error distribution is shown in *Figure 11* for six clusters. If the correct cluster k is assigned to a new query, the standard deviation of the prediction error (true time to FL250 – predicted time based on the cluster representative C_{rk} at FL250) is about 37 seconds.

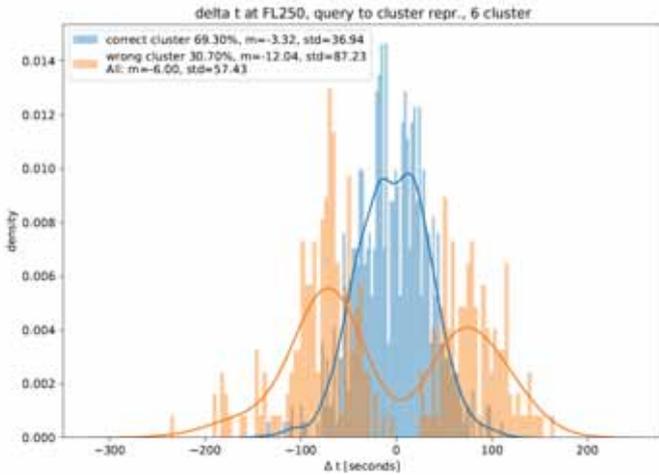


Figure 11: Error distribution at FL250 for correctly and wrongly assigned clusters

On the other hand, if the wrong cluster has been assigned, the standard deviation increases to 87 seconds (refer to orange double dipped distribution). The overall resulting standard deviation is about 57 seconds.

The quadratic confusion matrix M with elements m_{ij} (row i with true cluster, column j with predicted cluster) shows for the 10000 queries the distribution of predicted and true clusters. The class accuracy (sum of $m_{ii}, i = 1, \dots, k$) is about 70%. About half of the matrix elements m_{ij} are 0 or close to 0. Looking for example at the predicted cluster 4 in the fifth column, in most cases the prediction m_{44} is correct while in about 1/3 of the cases (144+365) the true cluster is one of the neighbor clusters, either cluster three or five. Note that the cluster number corresponds to those in Table 2.

Cluster number	Pred. 0	1	2	3	4	5
True 0	529	84	0	0	0	0
1	236	1257	291	0	0	0
2	28	419	1508	400	0	0
3	2	38	457	1663	365	0
4	0	0	21	318	1540	133
5	0	0	0	5	144	561

Table 3: Confusion Matrix for six clusters

Of course, it would be beneficial, if the class accuracy could be further improved. So far, we have only been using the climb behavior, i.e. the time needed to achieve a certain flight level. Open question is whether the not yet used information about the query flight (refer to Table 1, plus additional flight plan data, together with parametric models like BADA and e.g. mass estimation from [2] or [3], or using a framework like [22]) would

allow to discriminate between clusters 3, 4 and 5 if cluster 4 has been predicted. For this specific example, this is a classification problem with three classes (sometimes two classes or four classes, refer to matrix M) instead of six classes.

VI. CONCLUSIONS AND OUTLOOK

With a derived vector of constant length that describes the climb profile, K-Means is capable of identifying clusters with similar climb behavior. As all aircraft types have been considered in the sample of 45000 climbing flights, the cluster representatives present typical profiles. If the number of clusters is increased, the climb profiles cross each other and a high initial climb performance may switch to lower climb performance, in particular after the acceleration phase to constant Calibrated Airspeed. This indicates that a prediction of the climb profile based on initial climb segments (and even with good knowledge of an aircraft climb parameters) is very difficult.

The cluster representatives allow to compare new partially known trajectories, only based on the climb behavior, and to predict the correct cluster with a certain accuracy. This accuracy depends on the number of clusters - more clusters imply less intra-cluster standard deviation of the prediction error but it is more difficult to forecast the correct cluster.

Further analysis is required to look at similarities within each cluster, based on the static and dynamic aircraft parameters like aircraft type, Take Off mass, departure route, Indicated Airspeed or others. The authors believe that a combination with other available approaches like parametric models and/or mass estimations could further improve the correct cluster prediction.

In a second step, an intra-cluster search with algorithms like K-Nearest Neighbors (as proposed in DART [7]) could probably further improve the prediction accuracy.

ABBREVIATIONS

BDS	Binary Data Selector
CAS	Calibrated Air Speed
hPA	hecto Pascal
CI	Cost Index
ED	Euclidian Distance
EHS	Enhanced Surveillance
FCU	Flight Control Unit
FL	Flight Level
GS	Ground Speed
IAS	Indicated Air Speed
ML	Machine Learning
RoC	Rate of Climb
SESAR	Single European Sky ATM Reserach

REFERENCES

- [1] Common Trajectory Prediction Capability for Decision Support Tools, Sip Swierstra, EUROCONTROL HQ, Brussels, Steven M. Green, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, CA, 5th USA-Europe ATM Seminar, Budapest, June 2003
- [2] Sun, J., Ellerbroek, J. & Hoekstra, J., Bayesian inference of aircraft initial mass, in 'Proceedings of the 12th USA/Europe Air Traffic Management Research and Development Seminar', FAA/ EUROCONTROL, 2017
- [3] Machine Learning and Mass Estimation Methods for Ground-Based Aircraft Climb Prediction. Richard Alligier, David Gianza, Nicolas Durand. IEEE Transactions on Intelligent Transportation Systems, IEEE, 2015, pp.1-12.
- [4] Statistical Modelling of Aircraft Takeoff Weight. Y.S. Chati, H. Balakrishnan. 12th USA/Europe ATM Research and Development Seminar ATM 2017.
- [5] User Manual for the Base of Aircraft Data (BADA), Revision 3.8, EEC Technical/Scientific Report No. 2010-003. April 2010
- [6] Adaptive Trajectory Prediction Algorithm for Climbing Flights. Charles A. Schulz, David Thippavong, Heinz Erzberger. American Institute of Aeronautics and Astronautics.
- [7] Evaluation and Validation of Algorithms for Single Trajectory Prediction. DART D2.4 Data Science in ATM, ER-2-2025. University of Piraeus Research Center. August 2018
- [8] Flight Level Prediction with a Deep Feedforward Network. Matthias Poppe, Deborah Fieberg, Roland Scharff. SESAR Innovation Days Salzburg, December 2018
- [9] Review & Perspective for Distance Based Trajectory Clustering. Philippe Besse, Brendan Guillouet et.al. arXiv: 1508.04904v1. August 2015
- [10] Deep Trajectory Clustering with Autoencoders. Xavier Olive, Luis Basora, Benoît Viry and Richard Alligier. International Conference of Research in Air Transportation ICRAT 2020
- [11] OPTICS: Ordering Points to Identify the Clustering Structure. M. Ankerst, M. Breunig et.al. Proc. ACM SIGMOD 99 pp.49-60, Philadelphia PA, 1999
- [12] A density based algorithm for discovering clusters in large spatial databases with noise. M. Ester, H.P. Kriegel et.al., KDD, vol. 96, no.34, pp.226-231, 1996
- [13] Least squares quantization in pcm. S. Lloyd. IEEE transactions on information theory, vol. 28, no. 2, pp 129-137. 1982
- [14] The Elements of Statistical Learning. Trevor Hastie et.al. Second edition, Springer New York, 2017
- [15] Introduction to Statistical Learning. Trevor Hastie et al. Springer New York, September 2013
- [16] ICAO Doc 9871 AN/464 Technical Provisions for Mode S Services and Extended Squitter, First Edition, 2008
- [17] EUROCONTROL Guidance Material For Transition Altitude Change, First Edition, Brussels, September 2004
- [18] Traffic, a toolbox for processing and analysing air traffic data. Xavier Olive. Journal of Open Software, Vol. 4, 2019
- [19] Bringing up OpenSky: A large-scale ADS-B sensor network for research. Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, Matthias Wilhelm. ACM/IEEE International Conference on Information Processing in Sensor Networks, April 2014
- [20] The OpenSky Network, <http://www.opensky-network.org>
- [21] Deep Learning, Ian Goodfellow, Yoshua Bengio and Aaron Courville, MIT Press Cambridge, Massachusetts, London, England, 2016
- [22] Open Aircraft Performance Model and Python Toolkit. Junzi Sun, 2019. Accessible via <https://github.com/junzis/openap>

