Unveiling Airline Preferences for Pre-tactical Route Forecast through Machine Learning

An innovative system for ATFCM pre-tactical planning support

Manuel Mateos, Ignacio Martín,
Rubén Alcolea, Ricardo Herranz, Oliva G. Cantú-Ros
Nommon Solutions and Technologies
Madrid, Spain
nommon@nommon.es

Xavier Prats
Department of Physics - Aerospace Division
Technical University of Catalonia - BarcelonaTech
Castelldefels, Spain
xavier.prats@upc.edu

***Abstract*—In this work we describe a novel approach for the prediction of the flight plan to be sent by airspace users during the pre-tactical phase of Air Traffic Flow and Capacity Management (ATFCM). The proposed approach uses machine learning algorithms to extract airspace user preferences in terms of route characteristics, allowing the prediction of new routes not observed during the model training phase. We present the results obtained from applying this approach to short and medium range KLM flights for 52 weeks. Results show that the proposed solution is robust, scalable and capable of reducing the number of wrong predictions provided by the current Network Manager operational solution by 24.3% (4.5% increment on accuracy).**

Keywords—ATFCM, pre-tactical trajectory forecast; machine learning; airline preferences.

I. Introduction

The continued increase of air traffic experienced in the last decades, now temporarily stopped by the impact of COVID-19, was already stretching airspace capacity to its limits in many areas worldwide. The goal of Air Traffic Flow and Capacity Management (ATFCM) is to ensure that airport and airspace capacity meet traffic demand while optimising traffic flows to avoid exceeding the available capacity when it cannot be further increased, following a seamless process that spans from strategic planning to operations.

In Europe, ATFCM is handled by EUROCONTROL, in its role of Network Manager, and comprises three phases: strategic planning, which covers the planning phase between 18 months and 7 days before operations; pre-tactical flow management, applied during the six days prior to the day of operations; and tactical flow management, which takes place on the day of operations. In order to detect demand and capacity imbalances, the Network Manager estimates the expected demand at a given timeslot and compares it with the expected available capacity in different airspace volumes (sectors) and airports. The information to estimate the demand comes from the airspace users (AUs), while the information on possible ways to sectorize the airspace and the expected capacity provision is given by the air navigation service providers (ANSPs).

During the pre-tactical phase, few or no flight plans (FPLs) have been filed by the AUs and the only flight information available to the Network Manager are the so-called flight intentions (FIs), which include the flight call sign, the airline, the origin and destination airports, the estimated departure time, and the aircraft model to be used. Trajectory information becomes available only when the AUs send their FPLs. To forecast this information, the Network Manager relies on the PREDICT tool, which is used to predict the FPL before it is filed and provide the Network Manager Operations Centre (NMOC) with the information required to ensure a correct allocation of resources in coordination with the ANSPs [1]. PREDICT generates traffic forecasts according to the trajectories chosen by the same or similar flight codes in the recent past, without taking advantage of the information potentially encoded in historical FPLs and trajectory data.

This paper proposes an Artificial Intelligence system that emulates the airline decision making process by exploiting the characteristics of the route that intervene in such decision. The use of machine learning models has two main advantages: the identification of trends and patterns which are not directly observable and the continuous improvement of the models with the increase of the available data. Nevertheless, the development of a machine learning system is not a trivial task and it requires detailed knowledge of the problem to be modelled. We argue that only the adequate combination of ATFCM-related expertise and machine learning can yield successful predictions.

Our previous work in [2] and [3] demonstrated that machine learning can be successfully applied for FPL prediction during the ATFCM pre-tactical phase, outperforming the PREDICT tool. However, the models reported in these papers have a number of limitations related with the fact that the models were generated individually by Origin-Destination (OD) pair, so there was no real generalisation of the airline decision making process. Additionally, the proposed models were constrained to predict one of the previously observed trajectories, and therefore are not able to predict a newly created trajectory.

Following a different approach, the work in [4] applied machine learning techniques to calculate the probability of choosing different routes according to route charges, ground distance and the percentage of regulated flights in each one of the potential routes. In this case flights were segmented according to airline type and arrival time. This model was applied to the OD pair Istanbul-Paris. The paper shows that the model provides a fair performance. Nevertheless, it concludes the need of including more variables in the model.

Regarding the use of additional variables, the research carried out in [5] addresses the route prediction problem by considering the influence of 17 features. The authors observe that the most relevant variables are wind, thunderstorms and rain, followed by the miles in trail.

Similar conclusions were obtained in [6] and [7], which predict the trajectories during the tactical phase. The presented experiments were performed using Hidden Markov Models (HMM) to select the most probable 4D trajectory. According to the authors, the probability of observing a certain trajectory depends on the weather, in particular temperature, wind speed, wind direction, and humidity.

Based on the lessons learnt from previous studies, this paper proposes a novel machine learning model that predicts the routes flown by different airlines using a single model for each airline (independent of the OD pair) that, based on the recognition of the airline preferences, simplifies the prediction task and facilitates the training process. This methodology aims to provide a general and consistent solution able to predict routes that have not been previously observed.

The rest of this paper is structured as follows: Section 2 details the proposed methodology; Section 3 presents and discusses the main results of the experiments; finally, Section 4 summarizes the main conclusions of the study and discusses future steps.

II. Approach and Methodology

A. *Approach*

A preliminary data exploration revealed that airlines tend to select the same route among the set of feasible routes for a given OD pair. The selected route is often the shortest one, although route charges can sometimes compensate the cost of a longer route. Additionally, we have observed that, for certain OD pairs, route choices depend on specific conditions, such as the day of the week, the route wind or the airport configuration.

Overall, the common patterns determining airline route choice behaviour are apparently too complex to be modelled by simple rules. The approach followed in this work proposes to train a machine learning model which is fed with all potentially relevant variables.

As the described behaviours have been observed to be different by airline, even for the same OD pair, the proposed approach contemplates the generation of an independent model by airline. From a machine learning perspective, the model attempts to predict the chosen route by performing route-based binary predictions to determine the probability to file a particular route given its characteristics (i.e., the probability to flight each route is predicted independently). In terms of the observations feeding the model, the decision of the airline to fly a particular route given its characteristics is an observation, but the decision of not flying another available route is also a valid observation. The use of both observations does not only provide more valid observations, but also helps to identify which routes are less likely to be flown under certain circumstances (e.g., during a storm). Ultimately, the model will provide the probability of flying each one of the available routes, so that the most probable route for each flight is finally selected.

B. *Data*

The necessary condition for the proper training of machine learning models is the availability of sufficient data, especially when the feature space is large. The present research has used the following data sources:

1) *EUROCONTROL's Demand Data Repository 2 (DDR2)*

This is the main data source. The extracted data, extracted from [8], covers AIRAC cycles 1801 to 1813, which correspond to 52 weeks of traffic in 2018. The information extracted from the DDR2 includes:

- Flight Plans: they contain the route filed by the airline, which is the expected output of our prediction model. The flight plan available at the DDR2 is the last filed Flight Plan (also known as M1).
- Route charges: unit rates by ANSP updated monthly.
- Airport location: geodesic reference location of each airport.

2) *Copernicus Climate Data Storage service (CDS)*

CDS ([9]) provides geospatial weather information contained in different products. The ERA5 data product has been used. ERA5 data contains dozens of weather variables, particularly wind and severe weather variables, among others. Data is available by pressure level; nevertheless, as Flight Level (FL) is not taken into account in this work, a typical cruise level value of 200mb (~FL380) is used as default pressure altitude in the model.

3) *IOWA MESONET*

The IOWA MESONET provides access to the airports METAR files through [10]. METAR files contain a time record of the airport's meteorological station.

4) *Socioeconomic statistics*

Gross Domestic product has been obtained using the gridded dataset provided by [11], which combines national and regional data and is provided with 0.5 geodesic degree resolution.

Population density has been obtained from NASA´s Socioeconomic Data and Applications Center ([12]). The data is based on counts consistent with national censuses and population registers with respect to relative spatial distribution and is also provided with 0.5 geodesic degrees resolution.

Finally, kerosene daily prices are extracted from the Federal Reserve Economic Data ([13]).

C. Route clustering

Route clustering is implemented to group routes that have minor geographic variations but do not involve a relevant difference from ATFCM perspective.

The application of clustering facilitates data processing as routes are afterwards identified with a cluster tag and the process just needs to keep track of a unique route (also called central route) per cluster tag.

Clustering techniques were already used in [2], which performed a clustering using the area between two routes as distance metric and DBSCAN as clustering algorithm. Two changes have been implemented with respect to this work:

- The minimum number of routes to consider a cluster is now one. This means that all routs belong to a cluster, in contrast with the case of [2], where in order to define a cluster, a minim number of 5 routes was required. This leads to a number of cluster-less routes which were grouped in a cluster defined as noise.
- The clustering is performed independently by AIRAC, i.e., the central routes obtained for each AIRAC are now the routes considered available for that AIRAC.

It is important to note that, while the training process has access to historical routes, the list of routes to be flown is not known beforehand, in other words, the algorithm cannot have the available route options for prediction. This issue could be solved in the operational domain by using a route catalogue or a path finder algorithm, which could generate a set of valid routes, given an OD pair, providing the different route options for the prediction problem. However, these tools are not publicly available. To overcome this limitation, we assume that the available routes are the observed routes in the validation AIRAC cycle. Therefore, the clustering follows the same approach both for the training and the prediction datasets.

D. Cluster variable asignment

The model developed considers two kinds of features: general variables and cluster variables.

Cluster variables are dependent on the route under study. A simple route characteristic, e.g., route length, cannot provide information to the model by itself, as it is actually the length difference with other available routes what is relevant for the route choice problems. In other words, the model needs a reference.

It is also important to highlight that route variability in the FPLs is relatively low. Around 80% of the flights of an airline for a given OD pair follow the same route, i.e., airlines tend to consistently take the same route and select a different one only under specific conditions. It thus seems logical to take the most flown route as reference. For each AIRAC cycle, we have considered as a reference route the most flown in the previous cycle. For example, if the length of a route is 1,000 km and the length of the most flown route is 1,100 km, the reference value for the first route will be -100 km.

The cluster variables considered in the model are listed below.

1) Ground distance

The ground distance is probably the first variable motivating airline's choice. It is calculated by summing the projected ground length of the different segments composing the route. Following EUROCONTROL's recommendations, we have discarded the trajectory waypoints that were located closer than 40 NM to the origin and destination airports, in order to avoid introducing noise in the analysis of routes with minor differences in the terminal airspace.

2) Air distance

The wind length is calculated by adjusting the ground distance with the wind extension. The wind extension is calculated using the average wind projected along the flight path for each cluster central route and multiplying this average wind by the central route flight duration.

The air distance could be shorter (net tailwind) or longer (net headwind) than the route length.

3) Fuel cost

Fuel cost is known to be one of the main direct costs in the aviation industry. According to [14], the cost of the fuel alone can represent more than 30% of the airline operating costs. The calculation of fuel cost has two major components: fuel consumption and fuel price. One of the proposed features, air distance, is a used as a basis to calculate fuel consumption. Although climbs and descents are typically longer than 40NM, in this paper we assume that the air distance computed above is entirely flown in cruise conditions. Then, fuel consumption can be approximated by multiplying the air distance by the typical economic cruise fuel consumption. The typical economic cruise consumption for the Boeing 737-800, obtained from [15], has been taken as a reference value, as it is the most common aircraft in KLM's fleet; for other aircraft, fuel consumption has been assumed to be linear with the Maximum Take-off Weight (MTOW). Finally, fuel cost is estimated according to daily kerosene price.

4) Route charges

AUs pay different charges to cover different ATM services. These charges can be airport or route charges. As origin and destination airports are already fixed for the prediction, the only possible differences are in the route charges. The work done in [16] suggests that European airlines take into account the route charges when filling their flight plans. European route charges are calculated according to the entry and exit points on the different national airspaces that the flight navigates in.

Under the valid route charging scheme for the analysed period (2018), airlines paid charges according to the FPL, not the flown route. This situation changed in January 2020, when AUs started to get charged for the actual flight path.

5) Direct cost

The variable "direct cost" aggregates the charges and the fuel cost.

6) Convective phenomena

Convective phenomena features are calculated along the central routes. For each meteorological indicator, the average and the maximum values observed are calculated as features. The meteorological indicators used are:

- **K-index:** also known as George´s index, it is a measure of thunderstorm potential. It is a function of Temperature and Dew Point at several altitudes.
- **CAPE:** convective available potential energy. It is a measure of the instability in the atmosphere.
- **Humidity:** the presence of a relatively high fraction of water in the atmosphere is a necessary condition for some events such as storms to happen.

7) Local wind at origin/destination airport

Local wind is extracted from the origin and destination airports METAR files for the expected departure and arrival time. Based on the analysis of previous work, the airport configuration, which is highly dependent on the local wind, we have identified airport runway configuration as a feature having significant influence on route selection. This effect cannot be seen in every OD pair as it appears to be related with those cases in which arrival/departure points are rather separated in the terminal area, the ground distances are almost equally large for both options, and the convenience of using one of them depends on the airport configuration.

There are two components of the wind to be taken into account: the wind speed and the wind direction. Wind speed is just a scalar magnitude, so it can be directly used as a feature, while wind direction requires further processing. The hypothesis behind the wind direction is that, assuming the availability of similar routes that depart from/approach the airport in different directions, the airline will select the route that minimizes the need for manoeuvres. In other words, the airline will select the route whose first/last segments of the trajectory are more aligned with the take-off/landing direction. As operations are preferred to be done with head wind, especially landing, the wind may serve as a proxy of the airport runway configuration and the angle between the wind and the last/first segments should indicate the alignment with the airport configuration. An example of this calculation is presented in Figure 1. Ideally, the value of this angle would be 180 degrees if the last segment and the airport configuration (wind) were fully aligned.

8) Military zones

The impact of the military zones on aircraft trajectories has been addressed in [17] under a tactical scope. While it is clear that airspace restrictions will have a significant impact on

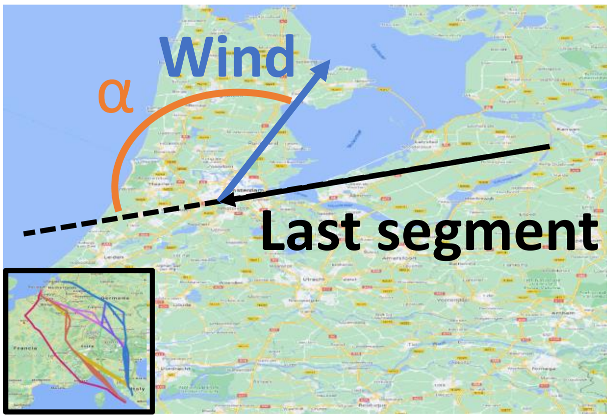

Figure 1. Local wind direction calculation for the pair LIRF-EHAM

pre-tactical planning, it is important to discuss the particularities of the European military airspace.

The European ATM system works under the Flexible Use of Airspace (FUA) concept, which means that airspace is no longer designated as purely "civil" or "military" and any necessary segregation is temporary, based on real-time usage within a specific time period. As a result of the application of FUA, the routes going through military airspace receive the name of Conditional Routes (CDR). Depending on the usability of these routes, they can be divided in three types:

- CDR 1 **-** Permanently plannable CDR during the times published. Available most of the time, not available under specific conditions (e.g., activation of a military training zone)
- CDR 2 **-** Non-permanently plannable CDR. Available under specific conditions, such as to facilitate traffic flow and increase ATC capacity
- CDR 3 **-** Not plannable CDR. Available on short notice, useable only on ATC instructions.

For the purpose of this work, CDR 3 routes have no impact as they can never be included in the FPLs. As for CDR 1 and 2, there is no practical difference, as both are announced to be opened or closed in advance to the flight planning phase, so their usage is supposed to be known the day before operations and therefore, both are treated equally in our model.

The airspace information included in the DDR2 repository contains the geographic description of the different military zones in Europe. Yet, it does not include the schedule of activation/deactivation of these zones or CDR time availability. The following approach was used to estimate the activation of the military zones:

1. Calculate occupancy (based on FPLs) for each military sector, day and hour.
2. Calculate the average occupancy for each sector and hour of the day.

3. If, for a particular sector, day and hour, occupancy drops are below a certain threshold, a military activation is flagged.

Regarding the time windows in which the occupancy is calculated, selecting a large time (e.g., 6 hours) could lead to misdetections of the military closure, while a short period (e.g., 5 minutes) would generate a large number of false positives. After discussion with several ATM experts, the time window was set to one hour.

Once the closure of military zones is estimated, each of the available routes is intersected with the active military zones at each given time and they are discarded as an observation if any of the crossed military zones was active.

It is important to highlight that the estimated closure of the military zones is just a workaround developed in the frame of this research due to data access restrictions. A hypothetical operational deployment of the proposed solution will not need to estimate the airspace closure as this information should be available for the Network Manager.

E. General variable asignment

Taking into account the results reported in [2] and [3], the following general variables have been selected:

1) Day of week (DoW)

It is broadly accepted that air traffic has a strong weekly component. The DoW has been used in two ways:

- Model feature: an integer number from 0-Monday to 6-Sunday
- Route filter: routes only flown during weekdays were not considered on weekends and the other way around.

2) Time of flight

We have used the Estimated Take-off Time (ETOT) of the flight. To capture the continuity between consecutive days, i.e., the fact that a flight departing at 23:55 will behave similarly to another departing at 0:10, a sin-cos transformation has been applied. The sin-cos transformation consists in the generation of two new features for each variable (see (1)), so they are always continuous.

$$h_c = \cos\frac{2\pi V}{T} \qquad h_s = \sin\frac{2\pi V}{T} \qquad (1)$$

where V is the variable to transform, the time of flight, and T the period (T=24 for the time of flight).

3) Day of Year (DoY)

The DoY is the ordinal position of any day of the year starting from the 1st of January (e.g., 21st of May 2018 is DoY 141). Following the same approach as for the time of flight, the DoY is also considered using a sin-cos transformation with T=365 (366 for a leap year).

4) Flight direction

The airline behaviour is not expected to be uniform for all the flights. One of the variables that might capture these variations is the flight direction. Flight direction is composed by two variables, the geodesic longitude difference between the origin and destination airports and the latitude difference. Following the usual conventions, North and East are considered positive. As an example, the flight direction for the OD pair Roma Fiumicino (LIRF) – Amsterdam Schiphol (EHAM) will be (-10.51, 7.47).

5) Airport socioeconomic variables

Nowadays, many airlines, especially legacy airlines, are profitable thanks to business travelling. Business travellers are often treated differently, so ultimately airline behaviour could be different for those flights that carry a significantly larger amount of business travellers.

It is not possible to estimate the amount of business tickets in each OD pair with the information at our disposal. Since business trips typically have origin and/or destination in densely populated and richer areas, we have used the local population and GDP in the origin and destination airports as proxies, taking the closest point of the grids defined in Section II.B.

6) OD pair competition

Following a similar approach as for the airport socioeconomic variables, it is possible that the competition in the OD pair might be affecting the airline behaviour. To take this into account, two proxy variables are considered: the OD pair frequency (computed as the number of flights) and the share of flights for each particular airline.

7) Maximum take-off weight (MTOW)

The MTOW of the aircraft is used as a numerical variable that characterizes the aircraft model. It is expected that the airline decision is influenced by the aircraft model but also that similar models (similar MTOWs) will lead to similar decisions. Other characteristics of the aircraft, such as age or engines, could affect the decision process as well. Nevertheless, this information is not easily accessible, so it has not been included in the model.

F. Model training

The machine learning algorithm used is a decision tree classifier. Although Decision Trees usually have a lower predictive power than more sophisticated algorithms, such as Random Forest or Neural Networks, they provide interpretability of results, which allows evaluating the effects of different conditions (weather, charges, etc.) on the final route choice decision. Cross-validation is used for hyperparameter tuning over the following hyperparameters:

- **max_depth:** the maximum depth of the tree.
- **criterion:** the function to measure the quality of a split, to be selected between the Gini index and entropy.
- **min_samples_leaf:** the minimum number of samples required to split an internal node.

III. Results and Discussion

The proposed methodology has been applied to the KLM flights. The reason to select KLM is that it has a significant number of flights with heterogeneous characteristics (length, zones, schedules, etc.). This allows exploring a wider range of casuistic.

All short and medium range flights of the airline with origin and destination inside the ECAC area have been considered. AIRAC 1813 has been chosen as a validation dataset and AIRACS 1802-1812 have been used to train the model.

A. Results of the KLM Probability Model

This experiment is intended to provide an overview of the model results. The model has been trained for all the available KLM observations and several temporal scopes have been considered depending on the AIRAC cycles used for the training dataset. Results are presented in Table 1, which shows that accuracy does not increase consistently with the number of AIRAC cycles used for the training. The explanation to this behaviour seems to be related with the airline´s winter/summer seasonal strategies. Our hypothesis is that airline behaviour is slightly different in each season, so the performance is better when training only with AIRAC data from the same season as the testing dataset. This hypothesis would explain why Model 7, which is trained including several weeks from September in AIRAC 1810, shows worse performance than those Models that do not include AIRAC 1810 (5, 6 and 8).

B. Model performance benchmark

In order to evaluate the performance of the proposed model, the accuracy has been measured and compared against that of PREDICT, the tool currently used by the European Network Manager. The PREDICT tool has been emulated following the information available from the Network Manager documentation [1] and the indications from EUROCONTROL experts. For each flight, the following workflow is applied:

1. select the FPL of the previous flights with the same call sign on the same day of the week. If this is not possible, the FPL of the flight operated by the same company at the closest time of the day is selected;
2. if no previous flight for the company is available, the same operation is repeated regardless of the company;
3. if no flight meets the previous requirements, the most recent FPL for the same OD pair is selected.

Additionally, the results are benchmarked against the machine learning model developed in [2] extended to the whole ECAC area, which we have called "Enhanced" model. The model presented in this paper, that will be called "Probability model". The main differences between both models are:

- The Enhanced model considers every OD pair as an independent decision problem (i.e., a different model is trained for each pair) in contrast with the current models which includes all airline decisions in the same model.

TABLE 1 – NETWORK MODEL RESULTS

Model ID	Training AIRACS	Validation AIRACS	Accuracy
1	1812	1813	0.814
2	1810-1812	1813	0.831
3	1807-1812	1813	0.834
4	1802-1812	1813	**0.852**
5	1802,1811,1812	1813	0.849
6	1802,1803, 1811,1812	1813	**0.854**
7	1802,1803, 1804,1810, 1811, 1812	1813	0.844
8	1802,1803, 1804,1811,1812	1813	**0.860**

- For each OD pair the Enhanced model consider flight information and weather variables, route characteristics such as length or charges are not included in these models since they present no variation within the same OD pair.

The analyzed results are not the binary classification results but the trajectory prediction (i.e., the most probable trajectory), which is the final outcome of the system. The goodness of these predictions will be assessed using only an accuracy metric per flight, which considers that a prediction is correct only if the route prediction matches the observed route.

Model 8 from Table 1 has been selected for the Probability model benchmark as it is the best performing model. To ensure the comparability, the validation analysis has been performed over the same flights taken from AIRAC 1813. Comparative results are shown in Table 2. The accuracy, as already explained, represents the number of correct predictions divided by the total number of predictions for each model. As for the comparisons against PREDICT, the increment reflects the relative increase of correct predictions against PREDICT's correct predictions, while the Error reduction is the relative decrease of wrong predictions against PREDICT's wrong predictions

Table 2 shows a significant improvement of the prediction in comparison with both the Enhanced model and PREDICT. The

TABLE 2 – BENCHMARKING RESULTS

Model	Accuracy	Increment (vs. PREDICT)	Error reduction (vs. PREDICT)
PREDICT	0.815	-	-
Enhanced	0.825	1.23%	5.41%
Probability model	0.860	5.52%	24.30%

reduction in the error is quite relevant as this model will avoid almost 1 out of 4 PREDICT's erroneous predictions.

C. Ad-hoc analysis, non-observed routes

One of the key improvements brought by the proposed modelling approach is that the model is capable of predicting new routes not previously observed in the training set. Since it is not necessary to include such routes in the training, the model is perfectly capable of calculating the probability of flying any new route just by deriving its features.

To exemplify this feature, we have chosen the OD pair connecting Kristiansand (Norway) and Amsterdam (ENCN-EHAM). This OD pair shows a new route in AIRAC 1813 that has not been flown previously in the training dataset. This new route is Route 3 (in purple) in Figure 2, which is used twice during AIRAC 1813.

The predictions of the model for this OD pair are shown in Table 3. Results detail the number of times each route (ID) was predicted by each model, specifying how many of these predictions were correct and how many were wrong (e.g., the Enhanced model predicted route 2 on 16 occasions; from those 16 times, 6 were correct predictions and for the rest, the model predicted route 2 but another route was actually selected).

As it can be seen in Table 3, not only the two assignments to the route with ID **3*** were correctly predicted by the model (while PREDICT does not forecast it correctly and the Enhanced model cannot even consider this prediction outcome), but also the general results outperform those from the two other models. The accuracy for all the predicted flights for the ENCN-EHAM OD pair shows an outstanding performance (75.9%) in comparison with the Enhanced model (63.0%) and PREDICT (51.9%).

IV. Conclusions

In this paper we have proposed a route prediction solution based on using machine learning to extract airline preferences. The forecasts correspond to pre-tactical ATFCM operations, where the route shall be estimated before the AUs send their flight plans. The most relevant conclusions of the study are summarized below:

- The proposed approach tested on the KLM airline, which account for the 3% of flights in the ECAC area, shows an 86% prediction accuracy, outperforming current operational approach (PREDICT) by 5.52% and reducing the prediction error by 24.3%.
- The model has been able to forecast non-observed routes in the training set.
- The model accuracy is significantly affected by the dataset used for training. Preliminary analyses suggest the importance of using data from the same season.

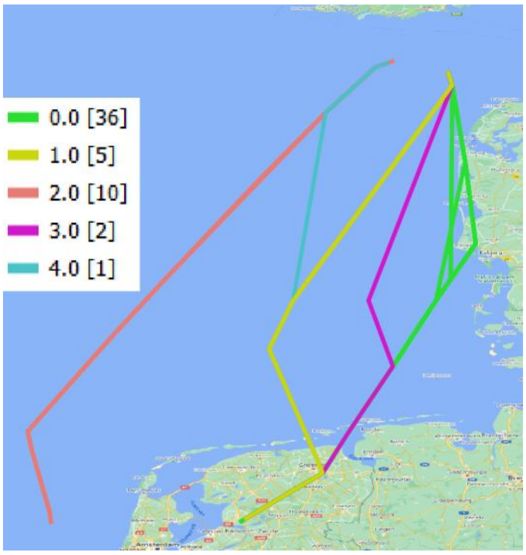

Figure 2. ENCN-EHAM OD pair routes for AIRAC 1813. In brackets the number of times the rote has been used in this AIRAC.

TABLE 3 – ENCN-EHAM PREDICTION

ID	Route observations	PREDICT		Enhanced		Prob. model	
		Correct assignments	Wrong assignments	Correct assignments	Wrong assignments	Correct assignments	Wrong assignments
0	36	23	12	27	10	27	2
1	5	1	4	1	0	3	1
2	10	4	8	6	10	9	10
3*	2	0	0	0	0	2	0
4	1	0	0	0	0	0	0
Noise	0	0	2	0	0	0	0
Total	54	28	26	34	20	41	14
Perc.	-	51.9%	48.1%	63.0%	37.0%	75.9%	24.1%

Additionally, the proposed approach may still be improved. Some of the planned future steps include:

- Extend the evaluation of the proposed methodology to the main airlines operating in the ECAC area.
- Explore other segmentations (e.g., group of airports, airline business model, aircraft type).
- Explore other machine learning algorithms: the decision tree algorithm has been implemented as it fitted conceptually with the airline decision process and it is computationally efficient; nevertheless, other algorithms such as neural networks might improve prediction performance.

ACKNOWLEDGMENTS

The authors would like to acknowledge EUROCONTROL's support to this work by facilitating data access and expert advice. Particularly, the authors would like to thank Dr. Ramon Dalmau, Ms. Stella Saldana, Mr. Stefan Steurs, Mr. Eric Allard and Mr. Francis Decroly for their advice on the definition of the models and the design of the evaluation experiments.

Manuel Mateos´ PhD is funded by the 1st SESAR ENGAGE KTN Call for PhDs and is developed in collaboration between Nommon and the Technical University of Catalonia. This PhD study has received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 783287. The opinions expressed herein reflect the authors' view only. Under no circumstances shall the SESAR Joint Undertaking be responsible for any use that may be made of the information contained herein.

REFERENCES

[1] EUROCONTROL, "ATFCM operations manual. edition number: 22.1," EUROCONTROL, Bretigny, France, Tech. Rep., 2018.

[2] M. Mateos, I. Martín, P. García, R. Herranz, O. García, and X. Prats, "Full-scale pre-tactical route prediction," in 9th International Conference for Research in Air Transportation (ICRAT), 2020.

[3] M. Mateos Villar, I. Martíın, P. García, R. Alcolea, R. Herranz, O. García, and X. Prats, "Predicting requested flight levels with machine learning," in 10th SESAR Innovation Days, virtual event, 2020.

[4] R. Marcos, O. G. Cantú Ros, and R. Herranz, "Combining visual analyticsand machine learning for route choice prediction," 7th SESAR InnocationDays, Beograd, 2017.

[5] Y. Liu, M. Hansen, D. J. Lovell, and M. O. Ball, "Predicting aircraft tra-jectory choice–a nominal route approach," inProc. of the InternationalConference for Research in Air Transportation, 2018.

[6] S. Ayhan and H. Samet, "Aircraft trajectory prediction made easywith predictive analytics," inProceedings of the 22nd ACM SIGKDDInternational Conference on Knowledge Discovery and Data Mining,2016, pp. 21–30.

[7] S. Ayhan and H. Samet, "Time series clustering of weather observations in predictingclimb phase of aircraft trajectories," inProceedings of the 9th ACMSIGSPATIAL International Workshop on Computational TransportationScience, 2016, pp. 25–30.

[8] EUROCONTROL, Demand Data Repository, retrieved from https://www.eurocontrol.int/ddr (Accessed 8/11/2021)

[9] Climate Data Storage https://cds.climate.copernicus.eu/ (Accessed 8/11/2021)

[10] Environmental Mesonet https://mesonet.agron.iastate.edu (Accessed 8/11/2021)

[11] Kummu, M., Taka, M., & Guillaume, J. H. (2018). Gridded global datasets for gross domestic product and Human Development Index over 1990–2015. Scientific data, 5(1), 1-15.

[12] https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-adjusted-to-2015-unwpp-country-totals-rev11 (Accessed 8/11/2021)

[13] https://fredhelp.stlouisfed.org/fred/about/about-fred/what-is-fred/ (Accessed 8/11/2021)

[14] https://www.statista.com/statistics/591285/aviation-industry-fuel-cost/ (Accessed 8/11/2021)

[15] https://web.archive.org/web/20140725005129/http://www.boeing.com/-assets/ pdf/commercial/startup/pdf/737ng_perf.pdf (Accessed 8/11/2021)

[16] L. Delgado, "European route choice determinants," in 11th USA/Europe Air Traffic Management Research and Development Seminar, 2015.

[17] H. Naessens, T. Philip, M. Piatek, K. Schippers, and R. Parys, "Predicting flight routes with a deep neural network in the operational air trafficflow and capacity management system,"EUROCONTROL MaastrichtUpper Area Control Centre, Maastricht Airport, The Netherlands, Tech.Rep, 2017.