Abstract

This paper presents new idea how trajectory calculations could be improved in order to match real flights better. Exact trajectory calculation is important for future of air traffic control, because it is one of the enablers for safe traffic increase. Methods used to calculate trajectories are based on aircraft types and their performances mainly. However, we believe that there are many other influencing factors which should be taken into account. We collect available data about flights and store them into a multi-dimensional database. Knowledge accumulated in this database is the basis for aircraft performances prediction using machine learning methods. In that way the prediction is not based on aircraft type alone, but also on other attributes like aerodrome of departure, destination and operator. There attributes indirectly imply to procedures, operator’s best practices, local airspace characteristics, etc. and enable us to make better predictions of aircraft performances. Predictions in this case are not static but tailored to every particular flight.

Introduction

In order to effectively cope with the increase of air traffic, air traffic control will have to find new ways to manage denser airspace safely. Plans for future air traffic developed by the Single European Sky ATM Research (SESAR) and North American NextGen, rely heavily on precise 4D trajectory calculations. Precise 4D trajectory calculations are crucial for future development of air traffic [1, 2].

Present clearance-based operations rely on air traffic controllers to identify and resolve potential conflicts. This will not be sufficient in the future. We will have to find new ways to manage airspace in a safe way. There are several ways to accomplish this goal. One is introducing new tools for air traffic controllers, which will enable them to identify potential conflicts earlier. This leads to a gradual move towards trajectory based operations and throughput optimizations, which enables positioning aircraft closer to each other. All these methods require as exact as possible flight path calculation and prediction.

To predict aircraft position in the air and to calculate these 4D trajectories accurately, we need to know exact aircraft performances. Currently, the main source of aircraft performances is the Base of Aircraft Data (BADA) model developed and maintained by the European Organisation for the Safety of Air Navigation (EuroControl) Experimental Centre (EEC) [3]. Aircraft are grouped according to types such as Airbus 380, Boeing 747, etc. However, the same aircraft types are not equipped identically and there are many other attributes, which influence their performance.

We propose to use machine learning for predicting more realistic flight performances. Our model of aircraft performance is based on real flight recordings. They act as an accumulated knowledge on how aircraft perform in realistic circumstances stored in a multi-dimensional database. Based on known attributes about a given flight, such as, operator, aerodrome of departure, destination, etc., the prediction algorithm tries to predict aircraft performances based on data from similar flights from the database.

The paper is structured in four sections. The first one outlines practices how others cope with trajectory prediction and problems related to it. The second section states our idea how to approach the problem of aircraft performances with recording real traffic and employing machine learning techniques. The results of our experiments are presented in the third section and concluding remarks with plans for the future are the content of the fourth section.

The Current Situation

The main source for aircraft performances currently in Europe is the BADA model [4]. To the
best of our knowledge there are no usable alternatives.

Currently version 3 is being used. In 2005 the development of BADA version 4 started. The new version is based on better and more detailed information about aircraft types. However, there seem to be some problems. It is very difficult to obtain information from manufacturers, so the new BADA model covers only 60% of the current aircraft types operating in European Civil Aviation Conference (ECAC) area. Another problem is the publication of the model. Negotiations for publication outside EuroControl are in progress [4].

BADA is based on mass-varying, kinetic approach that models an aircraft as a point and requires modeling of underlying forces that cause aircraft motion [4]. With the help of complex formulas aircraft performances are calculated from aircraft characteristics. Using these calculations to get aircraft’s trajectory is a difficult task. That is why the majority of applications use pre-calculated tables. These tables provide some default values of air speed, climb and descent rates at different altitudes and are a good start for calculating trajectories.

BADA user manual [3] recommends improving conformance of the simulated behavior with real operations by modifying BADA default speeds. In that way, local operation characteristics could be taken into account. We believe however, that it is not a realistic scenario for users to do that.

Applications in air traffic control are required to calculate more and more accurate trajectories, but they don’t have the means to do it. To improve trajectory prediction accuracy, even more complex models have been developed. They use BADA and data from real flight trajectory recordings.

The model by Schuster, Ochieng and Porretta [5] uses a flight management system to make flight path more realistic. It combines aircraft performances and flight intent to predict and adjust trajectory accordingly. BADA also provides a generic aircraft behavior model called AirLine Procedure Model (ARPM), which focuses on how the aircraft is operated. Gillet, Nuic and Mouillet [6] are using radar recordings to get realistic data to fine tune ARPM. Different energy share factors and speed profiles have been calculated according to airline operator, operating airport, aircraft type, flight phase and flight range with the help of statistical processing. In that way they generate more realistic flight trajectories for simulation purposes.

De Leege, van Paassen and Mulder [7] are using machine learning methods to predict trajectories along one particular landing procedure of a 45 nautical miles length. The trajectory prediction is predicting time over points from first approach navigation point along significant points to the runway threshold (a total of 7 points). Model inputs are: aircraft type (heavy, medium), aircraft ground speed, altitude over initial point and winds. The model predicts with an approximately 5s error on the last 15 nautical miles and 20s error on the 45 miles trajectory. Using this model to calculate the approach schedule, the capacity of the airport was increased by four aircraft per hour.

Kun and Wei [8] use a similar approach in the context of ignoring aerodynamics and using radar data. The method consists of two phases. First, they predict the total flying time based on historical data of identical flights. The second phase of prediction is adjusting the trajectory based on real-time radar data after the flight takes off.

Cheng, Cui and Cheng [9] use a hybrid of neural networks and statistical analysis. The proposed prediction model was tested on air traffic flow collected by the Air Traffic Control Command Monitoring System (ATCCMS), which aims to give early conflict alert and advice of short-term air traffic flow management to human controllers in the Beijing center. By means of the analysis, the air traffic flow was classified into seven categories corresponding to daily difference in a week, which were trained and forecasted separately.

For faster adoption we propose to use a service which will issue the same tables for aircraft performances as the BADA model does. The difference will be that the service will try to predict aircraft performances based on given details about a particular flight. On the other hand, BADA tables are static and are based on aircraft type only. In that way applications could continue to use the tables in the same way, but with better overall predictions.
An Alternative Approach with Machine Learning

We propose an alternative way of predicting aircraft performances which is general and supports any trajectory calculation that takes aircraft performances as an input. Our machine learning model searches for similar flights in a database and predicts aircraft performances based on similar flights performed in the past.

The model is not relying explicitly on physical characteristics of aircraft. It extracts information from flight data and from attributes such as aerodrome of departure/destination, operator, day of the week etc. These attributes indirectly indicate some physical characteristics.

An example: A flight from Ljubljana to Brussels is usually full on Mondays and less loaded on Tuesdays. The same aircraft can be fully loaded again the next day when transporting tourists to Gran Canaria. The machine learning model can recognize these differences based on accumulated knowledge from the past and predict different performances suited exactly for the particular flight.

To be able to predict aircraft performances in this way, it is crucial to accumulate reliable and trustworthy historical data. For that, we are using three main data sources.

Data Sources

The first source for measuring the actual aircraft performances are flight tracks recorded by radars. We identify cruise, descent/climb phases and extract performances from these flight phases. Measured performances need to be as exact as possible. Air traffic control uses at least double radar coverage for safety reasons. Radars provide recordings of aircraft positions called plots. Every radar provides plots for all aircraft within range. These unrelated plots are combined into tracks with software called tracker. A tracker usually uses a Kalman filter and extensive settings with many details about radars to effectively minimize radar measurement errors and to produce smooth trajectories. These smoothed tracks with radar errors minimized are the sources for our performances extraction and calculation. In Europe, ARTAS (ATM suRveillance Tracker And Server) is used as a primary tracker in EuroControl member states.

However, performances extracted from flights alone are of no value, until we add additional information to them. Another important function of all air traffic controls is correlation. It correlates flight tracks identified by tracker with flight plans filed in by pilots prior to take off. We use flight plan data as the second source of data and use correlation to enrich each extracted performance. In that way we get much additional information attached to performances, which are going to be used for predictions.

The third source of data is weather data. It is very important for calculating correct performances from tracks. Tracks provided with the help of radars can only record ground speeds of aircraft. To get the actual air speeds, we need to subtract the wind speed from the ground speed. An important information about atmosphere is also the air temperature. It is added to each performance. We get weather data from two sources. The first source is weather forecast for aviation provided by national meteorological agency. These forecasts are made with numerical weather prediction (NWP) models and are issued every twelve hours.

There is also a relatively new way of acquiring weather data with the help of Mode-S radars. Aircraft are able to measure wind speed by comparing ground speed and air speed. Mode-S radars can retrieve this information from aircraft if they are equipped with an appropriate transponder. Extensive studies have been performed [10, 11], and they show that these data are of satisfactory quality. When fresh weather data from aircraft are available, we use them for our air speed calculations. Otherwise we still have numerical weather predictions. Both are good enough to be used.

Pre-Processing

Data from all the sources are combined in the pre-processing phase and stored in a relational database. This database is used for creating multi-dimensional database as a source for predictions. The whole path from data sources to predictions is outlined in Figure 1.
First, the identification of flight phases from radar flight tracks takes place. We decompose the flight to climb, descent or cruise phases. Then the aircraft performances are extracted from these phases. From the cruise phase only airspeed is calculated. Climb and descent phases provide also respective climb and descent rates. Although we are simplifying the calculation of performances considerably, we expect to get good results nevertheless. The first example of simplification is acceleration and deceleration. When calculating airspeed, we always assume that the aircraft is not changing speed. The same applies for climb/descent rates. We would have problems detecting these maneuvers in flight reliably. On the other hand, even if the aircraft is accelerating, calculating average speed for a given flight section is valuable information. When we have our flight sections identified we calculate the average aircraft performance for the whole phase.

Next, we enrich the data with flight plan data and some additional attributes. Attributes from flight plans provide details about a given flight such as aircraft type, operator, departure, destination, date and time of flight, aircraft equipment, etc. Other attributes which are also added are temperature, duration of the section, etc. Duration of the section, for example, could tell us how reliable the measurement is. For a 20 second flight phase the average speed measurement cannot be as accurate as for a ten minute phase. There are approximately 30 attributes being gathered. Some of them are available for all flights and some are not. Some attributes are also very closely related to each other. For instance wake turbulence provided in a flight plan is directly linked to an aircraft type. In that case we don’t expect that such an attribute will contribute any valuable information in the prediction process. With all the attributes attached, a record called fact is stored into the database.

From the relational database of facts a multi-dimensional database is created. We use the online analytical processing (OLAP) technology for multi-dimensional database. Queries into this database are fast and effective with multidimensional expressions (MDX) language. That is important because we can expect instant predictions from the machine learning enabling responsiveness required for real trajectory calculation needs.

An important feature of pre-processing is full automation. The process from data sources to the multi-dimensional database is fully automatic. Every day, flights from the previous day are processed automatically and added to the database. When predicting flights for the present day, the newest data are already available in the database for predictions. This gives the prediction an interesting feature that the newest trends are available and are affecting results. The machine learning algorithms that we will use must therefore not be trained on a fixed training dataset, because the knowledge database is constantly growing.

A lot of attention has been given to the correctness and accuracy of pre-processing and performances extraction. It is very important to extract realistic performances and have accurate data for prediction. Without reliable inputs we cannot expect realistic predictions.

**Prediction**

When data are pre-processed and stored in the multi-dimensional database, we can use it for various purposes. The information stored can give a skillful user useful knowledge about the flight characteristics in the observed airspace.

It has been shown in the past that physical laws can be learnt from experimental data with artificial intelligence [12, 13]. We are trying something similar.
in our case. With the help of accumulated knowledge about past flights in the database we try to predict the performances of future flights. Pilots file in flight plans before taking off and these flight plans are the source of information we have. With the help of attributes from the flight plan we look for similar flights in the database and assume that the new flights will fly in the same way. In that way we take attributes which don’t look as physical parameters at a first glance and predict physical characteristics of the flight.

The nature of the problem directs us in the field of unsupervised machine learning in the sense that our database does not hold labeled samples. All samples are equal and the algorithm does not have a checking mechanism, which can tell whether the prediction was close or not. We were choosing among many available machine learning algorithms [14] and the first choice was the nearest neighbors algorithm.

The nearest neighbor's algorithm is very simple. We find a representative set of similar flights; we take the average performances of this set and that is our prediction. However, the simplicity of the chosen algorithm does not make the task of prediction easy.

First, we have a problem of specifying similarity. Flights, which have all the attributes identical to the predicted flight, are most certainly similar. However, there aren’t many identical flights. Even if we ignore different dates of flights, weather conditions, etc., they may even be identical in the sense of attributes available to air traffic control, but the take-off weight and other attributes which are not available in our database are different. These unknown attributes make each sample different. We need therefore bigger sample sets to decrease the influence of unknown attributes by averaging.

Our task is therefore to find a sample set, which is big enough to give us trust in average values and small enough, so that we do not include the samples, which are quite different from the new example.

Beside unknown attributes about flights there is also a class of other influencing factors, which affect performances. These are clearances and other instructions given by air traffic controllers or induced by airspace specifics. We expect that our prediction will cope with such specialties better if they are performed regularly. For instance, performances during approach procedures will be also dependent on actual procedure and not on aircraft type alone.

If there are no data about flights performances from the past in the database, our method is designed to return the BADA values instead of empty ones. When a new aircraft type appears in our airspace, the prediction will return BADA performances.

**Results**

The database with accumulated knowledge for the airspace used in this experiment is being updated with newest flights every day. It holds air traffic from February 2011 until the present day. We typically get five to fifteen thousand new facts every day. It depends on the amount and complexity of traffic. Until August 2014 over 1.100.000 flights were pre-processed. If a flight just overflies the airspace on the same altitude, we get only one fact per flight. For departures and arrivals there are more maneuvers (climbs, descents) and we therefore get more facts – up to 30. On average there are 3 facts per flight which translates to over 9.300.000 facts in the database from February 2011 until August 2014.

Our test set is composed of flights from January to June 2014. That is around 110.000 flights which are being predicted and compared. These flights have been analyzed and split into phases again. We have dissected 110.000 flights into 700.000 sections. When the performances are predicted, we calculate the expected time and compare it to the measured one for the corresponding section.

Only flight sections are being compared for now. This means that for each section only the time is calculated according to the aircraft performances given and compared to real flight times. In the future we will combine the sections from one flight together and predict times also for the whole flight. The reason why whole flights are not calculated is the transition between phases. We can have a climb phase that is followed by a cruise phase. We cannot just summarize times from both phases because the predicted climbing can be faster or slower than the one measured. That means that in the calculation, an aircraft can the reach cruise level earlier or later than in the recorded sample. Therefore the next phases are
affected, because predicted cruise phase can start earlier or later than measured.

Table 1. Average Accuracies of Times Calculated from Predicted Air Speeds

<table>
<thead>
<tr>
<th>Prediction Method</th>
<th>Cruise [%]</th>
<th>Climb [%]</th>
<th>Descent [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>BADA</td>
<td>11.06</td>
<td>12.03</td>
<td>25.83</td>
</tr>
<tr>
<td>Aircraft type average</td>
<td>7.38</td>
<td>9.84</td>
<td>21.63</td>
</tr>
<tr>
<td>Machine learning</td>
<td>7.25</td>
<td>9.59</td>
<td>21.55</td>
</tr>
</tbody>
</table>

We are predicting flight times for the test set with three different methods. The first method uses aircraft performances from BADA tables. The second method is similar and uses performances tables based only on aircraft type average from our database. The third method is using machine learning and average values of the flights from the database with matching aircraft type, operator, aerodrome of departure and aerodrome of destination. If there are not enough samples with matching attributes in the database the the algorithm relaxes the condition by removing one attribute and tries again. It repeats with eliminating attributes until predictions are reliable or until it comes to aircraft type only condition. In the latter case the values match the ones from the second method.

This relaxing does not apply for the whole performances table. If there are enough samples for some value, it is kept. Values with relaxed condition are being sought for only where there are not enough samples. The final table can hold values from different iterations. For instance: Air speed on FL350 is based on all matching attributes. There are not enough samples for FL340, therefore operator is removed from conditions. For FL330 even without operator the sample set is not big enough, therefore aerodrome of destination is removed. In that way, the algorithm fills the table until it is full. In the worst case, when even aircraft type average is not available, the BADA values are being returned. In that way the algorithm should behave as BADA in the worst case scenario if the average value is better than BADA value.

The accuracies of all three methods are compared to real flight times. Table 1 shows the average relative differences of calculated times using different performances compared to actual flight times. Table 2 shows the same for climb and descent rates. Figure 2, Figure 4 and Figure 6 show the differences in a greater detail. Prediction errors in the figures are shown as root mean square error (RMSE). RMSE is a good indicator of prediction accuracy. However, all prediction methods have problems with rare large miss-prediction errors if data about a flight is wrong. Sometimes it happens that the track observed is actually not the flight from the flight plan. These errors come from limitations of the test environment where flight plans are not manually checked and corrected as they are in operational environment. Flights can be changed right before the take-off due to technical or some other reasons or the test environment wrongly couples tracks with flight plans. All these prediction errors induced by wrong input data have a great influence on RMSE. For a clearer picture about prediction accuracy we have included additional charts with average relative accuracy shown in percentage of a phase’s total flight time. These relative prediction errors are shown in Figure 3, Figure 5 and Figure 7. We can see in these charts that the prediction errors get relatively smaller with longer flight times.

Table 2. Average Accuracies of Times Calculated from Predicted Climb/Descent Rates

<table>
<thead>
<tr>
<th>Prediction Method</th>
<th>Climb [%]</th>
<th>Descent [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>BADA</td>
<td>27.31</td>
<td>40.75</td>
</tr>
<tr>
<td>Aircraft type average</td>
<td>22.28</td>
<td>38.76</td>
</tr>
<tr>
<td>Machine learning</td>
<td>21.16</td>
<td>38.47</td>
</tr>
</tbody>
</table>
Figure 2. RMSE Comparison of Cruise Predictions

Figure 3. Average Relative Prediction Error of Cruises

Figure 4. RMSE Comparison of Climb Predictions

Figure 5. Average Relative Prediction Error of Climbs

Figure 6. RMSE Comparison of Descent Predictions

Figure 7. Average Relative Prediction Error of Descents
We expected that flight times calculated with BADA performances would give the biggest discrepancies in predictions. The second worst would be the average for the aircraft type from our database, and the best method would be the machine learning nearest neighbor algorithm, which uses some kind of similarity criteria. The results prove our expectations to be correct. However, we expected a little better performance from the machine learning.

Predictions for air speeds in cruise phases show the smallest differences in methods. We can see in Figure 2 and Figure 3 that when calculating the times for cruising, we get the worst results with air speeds acquired from BADA tables. The other two methods are practically identical. Their lines are aligned and almost indistinguishable. A closer look at the results reveals a barely noticeable advantage of machine learning which does not make a difference in real day to day predictions.

With the climb rates in Figure 4 and Figure 5 the situation is different. BADA shows the worst results while a difference can be observed also on the other two methods. Machine learning gives notably better results in this case.

For the descent rates in Figure 6 and Figure 7 the machine learning is again providing the best results. BADA values are in this case showing the worst results compared to others.

The results show that predictions for air speeds are closest to each other. This is understandable because only small changes of speeds are possible in cruise phases, for example. For climb and descent rates the situation is different. Air traffic controllers can give clearances for climb or descent rates which influence results more than assigned speeds. With introducing more performance based navigation procedures which enables aircraft to descent in the most economical way, the situation will probably change.

Taken into account all results together, the machine learning method has proved to be the best choice. The average on aircraft type alone is the second best and BADA the last. We believe that we have proven, that machine learning has a promising potential for better trajectory prediction.

The results show also that there is still ample room for improvement in machine learning. The first improvement is to define a better similarity measure, which will help to identify similar flights. Next, the machine learning algorithm will be improved to find the optimal number of nearest neighbors. And also other machine learning approaches have already been evaluated for possible implementation. The final plan is to use a combination of different methods to get the best possible results. This is called ensemble learning. At the end we expect a much more notable benefit in prediction with machine learning.

**Side Effects and Other Uses**

With the multi-dimensional database in place there are many other possibilities on how to use the accumulated data. Manual queries into the database provide numerous possibilities to investigate patterns about practices in the observed airspace. We are presenting one simple example here. It shows the average performances of the most common aircraft in our airspace Airbus 320 (A320) extracted from this database and compared to BADA tables.

Figure 8 shows average measured airspeeds compared to BADA speeds. Figure 9 shows comparison of climb rates and Figure 10 shows the descent rates. It can be observed in the charts that some measured values are very close to BADA while others are far apart. The differences on lower altitudes can at least partly be attributed to airspace specifics, while it is difficult to say that for higher altitudes.

![Figure 8. Comparison of Air Speeds for A320](image)

The machine learning method has proved to be the best choice. The average on aircraft type alone is the second best and BADA the last. We believe that we have proven, that machine learning has a promising potential for better trajectory prediction.

8C4-8
One possible use of these observations is what the BADA developers also suggest. Since BADA parameters are global, they cannot take into account local particularities. Fine tuning of BADA parameters could be performed for the observed airspace.

Another idea is also to estimate or predict take-off weights. With that and possibly other valuable information from the database, trajectory calculations could be more precise. It could be some new method based on database information or already known methods using BADA parameters adjusted with prediction data.

There is a possibility to connect to the database with Microsoft Excel and use pivot tables to perform ad-hoc queries. A user can get any information with just a few clicks. With all the stored data and the ease of performing different, sometimes unimaginable views of it, the possibilities of various usages really are numberless.

**Conclusion**

We have shown in this paper how measurements and data recording can help with aircraft trajectory calculation and prediction. Our first attempts in machine learning gave the best results. However they are not satisfactory yet. We expect to find better methods of prediction from the existing database, which will improve results further. The results are comparable to other methods, but we expected somewhat better results and we are confident that we can get better results in the future.

For the future we plan to develop methods for predicting aircraft performances, which will enable prediction of trajectories closer to reality. The methods for evaluating the test results will also be improved to combine phases together into whole flights. As long as air traffic related software will use relatively simple algorithms for trajectory calculations using aircraft performances tables, the prediction from these tables for each flight seems like a very promising alternative. More complicated trajectory calculations data about real flights could also contribute vital information for fine tuning and better calculations. However, we believe that using our service for aircraft performances predictions is simpler than fine tuning some parameters. And since the database is updated regularly the new trends and changes can be quickly reflected in predictions.

We are accumulating the knowledge in a multi-dimensional database and this technology allows users with no extensive knowledge in computer technologies to exploit its possibilities in many ways. We have shown that on the example of Airbus A320 performances from the database. The possibilities of different views in the data are limitless for many users in air traffic control, where data about flight characteristics from other sources are not available.

**References**


33rd Digital Avionics Systems Conference
October 5-9, 2014