The revenue and performance of a flight schedule is highly dependent on its robustness. Robustness can be thought of as the flexibility available within a schedule to enable a user to recover from disruptions to that schedule. A more detailed description of robustness can be found in [8]. A lack of robustness in a schedule might result in one delay leading to a series of knock on delays which can seriously affect the smooth operation of an airline. In order to mitigate against this, airline companies want to have accurate estimates of a schedule’s robustness before that schedule is brought into operation. Regular evaluation during the scheduling process enables schedule operators to improve the schedule quality by adding slack where necessary and by incorporating strategic patterns – for instance a favorable slack distribution - into the schedule. A more detailed description of robust scheduling within the airline industry can be found in [1, 2].

Our full paper will present custom-built and generic models which have been developed to evaluate schedule robustness. This work builds upon some of the conclusions arising from research carried out by Lee et al. [7], Schumacher [10], Rosenberger et al. [9] and Farrington et al. [5] on stochastic modeling which provides schedule operators with a detailed what-if analysis of the schedule. However, although stochastic models provide good estimates of schedule robustness, operators often find them unsatisfactory because of their heavy computational load. This prevents them from performing schedule evaluations on a regular basis. Our method represents a much quicker approach to help operators to build better models. Moreover, our initial experiments suggest that the approach is also more accurate but this hypothesis will be rigorously evaluated and discussed in the full paper.

We explore a prediction of robustness which is based upon certain robustness features which are drawn from schedule operators at KLM and from recent papers in the literature [1, 2]. This feature-based prediction enables us to forecast schedule robustness. The robustness features in the flight schedule which we are exploring include:

- slack distribution
- likeliness of delay propagation
- number of redundant paths [1]
- number of swap options
- presence of strategic patterns
- number of meeting points in the schedule

The goal is that these features (and others) will be identified and translated into mathematical parameters called explanatory variables. The selection of a relevant subset is carried out using wrapper feature selection [6] on historic real world data from KLM. In this iterative process an induction algorithm [6] is used to train a classifier using a selected subset of features as input. Model evaluation is carried out using n-fold cross-validation [6] on the training set. The iterative process finishes once the search
algorithm ends or certain stop-criteria are met. Finally, the individual classifiers are combined into a hybrid model which leads to a better classification accuracy due to the uncorrelated errors of the individual classifiers [4]. The resulting hybrid method is used to forecast performance of a particular schedule in the future based upon trends which are discernible from the data taken from recently implemented schedules and which include passenger loads, weather conditions etc.

In our analysis of KLM data, we defined parameters based on hub banks or waves in their schedule [2]. KLM typically serves a large number of transfer passengers. To deal with this, the airline creates four hub banks or waves per day into their schedules in which a significant number of aircraft arrive and take off in a short period of time. While the commercial benefit is represented in an increased number of possible connections for their passengers, banks could also lead to an increased robustness, resulting in a better performance of the schedule. A favorable distribution in time of the aircraft standing idle at KLM’s hub (at Schiphol airport) during the banks could result in an increased number of swap options leading to a greater flexibility within the schedule to enable us to recover from disruptions.

The model was built using 12 operational plans representing flight schedules for KLM’s European fleet. The historically measured five minutes departure punctuality – that is, the number of aircraft which have actually left five minutes after their scheduled departure time - is used as the evaluation criterion for schedule performance. Bagging (which is a method to construct a hybrid model [3, 4]) is used to hybridise our model of the individual classifiers. Two operational plans, S03-1 and S03-2, covering the summer of 2003 were used as a test set to measure the accuracy of the resulting model.

Figure 1 shows the historic and the predicted five minute departure punctuality, averaged for each weekday in the operational plan. Figure 2 represents historically measured values and the model’s output averaged for each operational plan. It can be seen that the historic and predicted curve show a strong correlation, although there is a serious discrepancy with consistently lower values for the predicted curve during the first part of summer 2003. However, further analysis leads us to the conclusion that exceptional external events – such as the war in Iraq and SARS - resulted in less passenger traffic at Schiphol leading to less passenger related delays and a higher historic punctuality. It is, of course impossible to predict such events accurately. Comparing the results of our model with KLM’s simulation model shows similar results. Further refinement of the model, taking into account external influences, might overcome this deficiency in precision.

![5 minute Departure Punctuality averaged per day per Operational Plan](image)

Figure 1: Historic vs. predicted 5 minute departure punctuality averaged per day, per operational plan
To conclude, we will, in the full paper, describe our forecasting methodology which is showing very promising early results. This promise will be explored further and evaluated in the full paper. The full paper will also highlight some guidelines which can be used to take schedule robustness into account when generating airline schedules.

Figure 2: Historic and predicted 5 minute departure punctuality averaged per operational plan

We will also outline some future research directions which include developing more *explanatory variables* in the initial subset. It is likely that the inclusion of additional - perhaps more relevant - parameters might result in a better and more accurate model.

REFERENCES

[8] Leus R. (2003), The generation of stable project plans, Complexity and exact algorithms, PhD thesis at the university of Leuven, Faculty of applied economics.