A data-driven approach to the Air Traffic Flow Management problem

Luigi De Giovanni

DIPARTIMENTO MATEMATICA

UNIVERSITÀ DEGLI STUDI DI PADOVA

CFR CONSORZIO FUTURO IN RICERCA

with Carlo Lancia[1] and Guglielmo Lulli[2]

[1]*Mathematical Institute, Leiden University*

[2]*Department of Management Science, University of Lancaster*

Outline

- Air Traffic Flow Management
  - Airspace capacity, user preferences, conflicts, strategies
  - Problem statement
- Classical approaches
  - Literature review and IP models principles
  - Limits for real-world application
- Data-driven approach
  - IP model based on trajectory selection
  - Learning feasible trajectories from data
  - Learning user preference from data
- Conclusions and future work

Air Traffic Management

«*Air traffic management (ATM) considers the **trajectory** of a manned or unmanned vehicle during all phases of **flight** and manages the **interaction** of that trajectory with other trajectories or hazards to achieve the optimum system outcome, with minimal deviation from the **user-requested** flight trajectory, whenever possible.*» (ICAO Doc. 9854, §1.9.2)

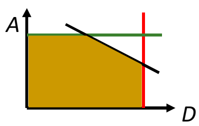
Decision levels towards effective flight plans:

- **Strategic** (months to week before): airspace *capacity*
- **Tactical** (days to hours before): up-to-date capacity, *regulations*
- **Operational** (day of flight): collision avoidance

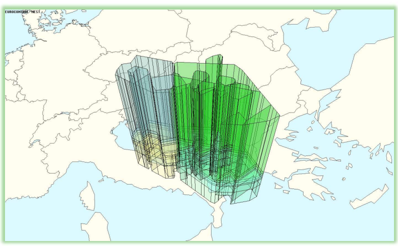
Air Traffic Flow Management Problem - ATFM

Airspace capacity restrictions

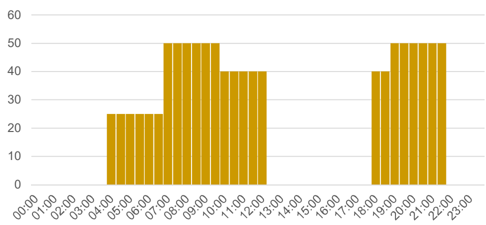
- **Airport capacity**
  - Arrival / departure



- **En-route sectors capacity**
  - Num. flights that can enter (ECAC) [cross (NAS)] per time unit
  - Depends on geometry (e.g. size) and ATC resources

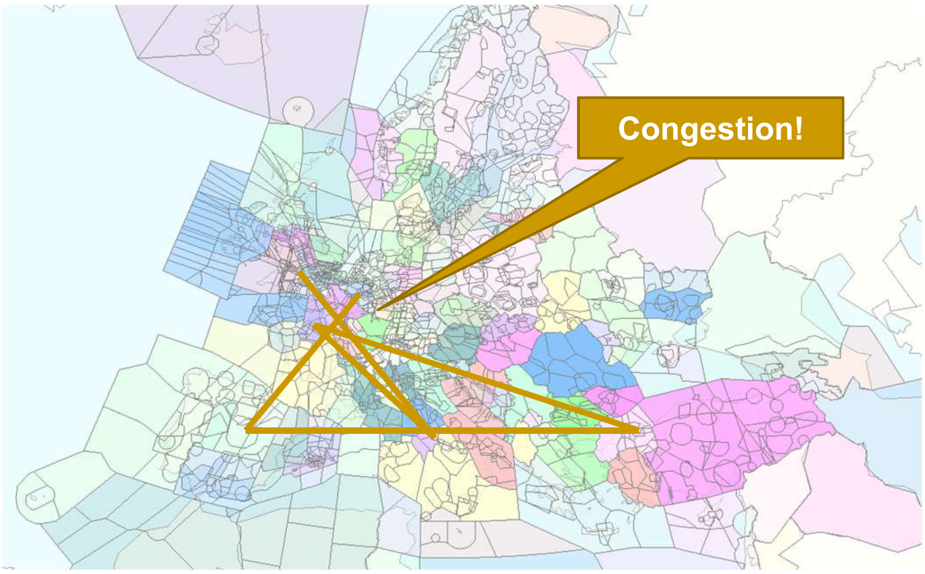


Control Area (CTA) with 4 sectors [Eurocontrol - DDR2 + NEST]



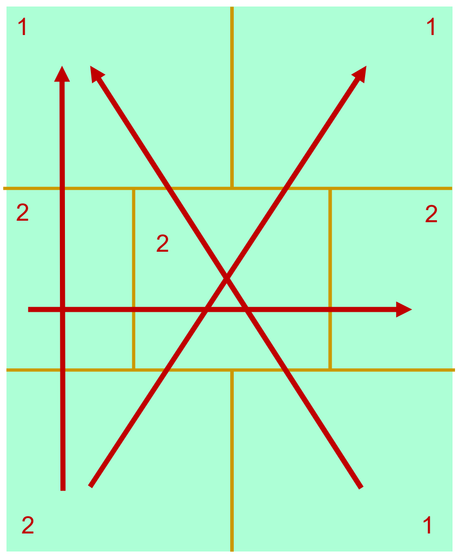
Dynamic capacity (4D sectors)

An example



[Eurocontrol - DDR2 + NEST]

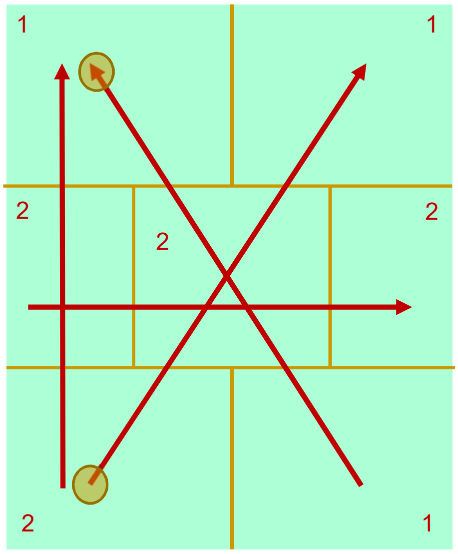
Possible strategies: a toy example



4 flights
Sector capacity: 1 or 2
Cross time 1 t.u.:
STD: time 0

User requested unfeasible

Possible strategies: a toy example

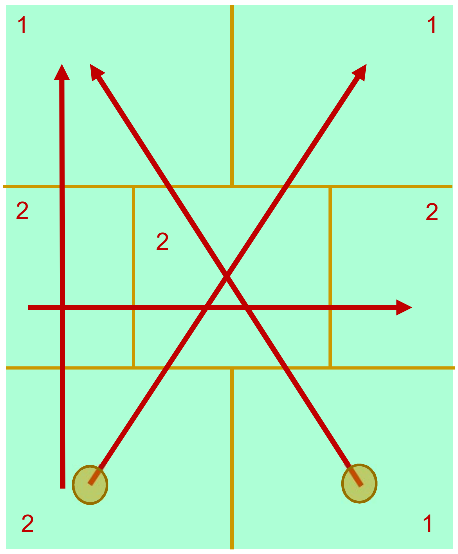


4 flights
Sector capacity: 1 or 2
Cross time 1 t.u.:
STD: time 0

User requested unfeasible

a) 1 t.u. of Ground Delay
1 t.u. of Airborne Delay

Possible strategies: a toy example



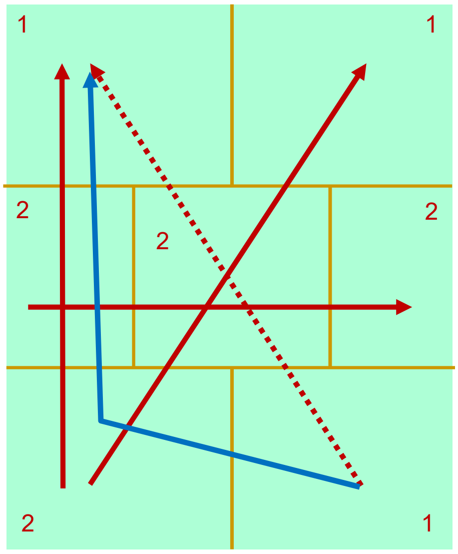
4 flights
Sector capacity: 1 or 2
Cross time 1 t.u.:
STD: time 0

User requested unfeasible

a) 1 t.u. of Ground Delay
1 t.u. of Airborne Delay

b) 2 t.u. of Ground Delay

Possible strategies: a toy example



4 flights
Sector capacity: 1 or 2
Cross time 1 t.u.:
STD: time 0

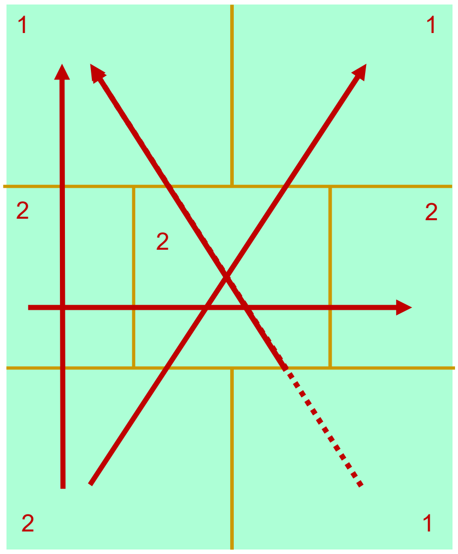
User requested unfeasible

a) 1 t.u. of Ground Delay
   1 t.u. of Airborne Delay

b) 2 t.u. of Ground Delay

c) Deviation (1 t.u., cost+)

Possible strategies: a toy example



4 flights
Sector capacity: 1 or 2
Cross time 1 t.u.:
STD: time 0

User requested unfeasible

a) 1 t.u. of Ground Delay
   1 t.u. of Airborne Delay

b) 2 t.u. of Ground Delay

c) Deviation (1 t.u., cost+)

d) Speed control (1 t.u.)

Air Traffic Flow Management Problem (ATFM): general statement

Given

- A set of flights with *initial* 4D trajectories
- Airspace configuration and capacity restriction

determine

- A set of *modified* trajectories

such that

- Capacity restrictions are satisfied
- System «efficiency» maximized (e.g. minimum delays, minimum deviation, airspace users' preferences)

Some Integer Programming models for ATFM

Helme (1992): ground holding, airborne delay

Bertsimas & Stock-Patterson (1998): + speed control

Bertsimas & Stock-Patterson (2000): + rerouting (small instances)

Bertsimas, Lulli & Odoni (2011): + rerouting, fairness

Augustin, Alonso-Ayuso, Escudero (2012): + waypoints

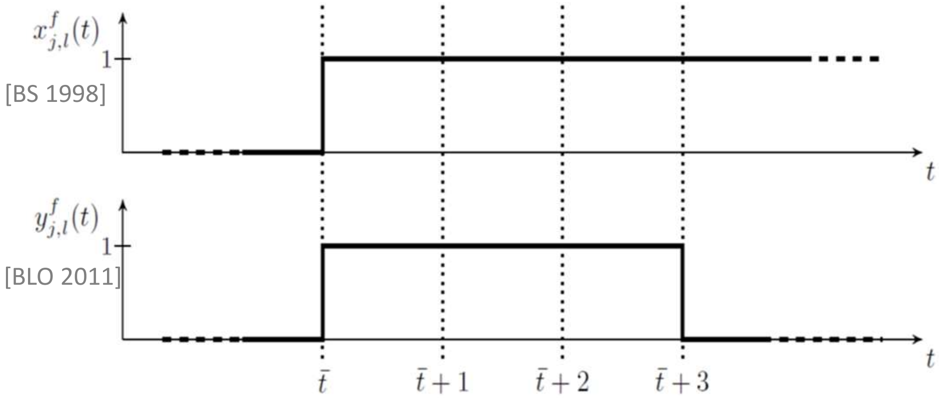
Djemou, Lulli & Zografos (2017): + flight levels, TBO

Akgunduz, Jaumard & Moeini (2017): collision avoidance

Diao & Chen (2018): collision avoidance

"Classic" IP models: decision variables

- $x(f,j,l,t)$: 1 if flight f reaches sector j at flight level l **by** time t
- $y(f,j,l,t)$: 1 if flight f is in sector j at flight level l **at** time t



"Classic" IP models: objective function (simplified)

- E.G.: weighted sum of ground (g_f) and airborne delays (a_f) w.r.t. nominal initial flight routes (d_f and r_f)

$$\min \sum_{f \in \mathcal{F}} [c^f_g g_f + c^f_a a_f]$$

$$g_f = \sum_{t \in T^f_k, k=P(f,1)} t(x^f_{k,l}(t) - x^f_{k,l}(t-1)) - d_f$$

$$a_f = \sum_{t \in T^f_k, k=P(f,N_f)} t(x^f_{k,l}(t) - x^f_{k,l}(t-1)) - r_f - g_f$$

t is the time *f* departs/arrives

“Classic” IP models: connectivity (simplified)

$$\sum_{l \in \mathcal{L}_j^f} x_{j,l}^f(t) \leqslant 1$$

one flight level (FL) per sector

$$\sum_{l \in \mathcal{L}_j^f} x_{j,l}^f(t) - \sum_{l' \in \mathcal{L}_{j'}^f} x_{j',l'}^f(t + I_j^f) = 0$$

I_j^f t.u. from sector j to j’, the next in the route (here, fixed route and speed)

$$x_{j,l}^f(t) - \sum_{l' \in \mathcal{L}_{j'}^f \cap [l - \delta_j^f, l + \delta_j^f]} x_{j',l'}^f(t + I_j^f) \leqslant 0$$

maximum δ_j^f (FL variation) from j to j’

$$\sum_{l \in \mathcal{L}_j^f \cap [l^0, l^0 + \delta_j^f]} x_{j,l}^f(t) - x_{k,l^0}^f(t + I_j^f) \geqslant 0$$

I_j^f and δ_j^f at the arrival airport

“Classic” IP models: capacity (simplified)

- Dynamic capacity at time *t* for
  - airport *k* **D**epartures and **A**rrivals
  - flights crossing en-route sector *j*

$$\sum_{\substack{f \,:\, P(f,1)=k, \\ t \in T_k^f}} (x_{k,l^0}^f(t) - x_{k,l^0}^f(t-1)) \leqslant D_k(t)$$

$$\sum_{\substack{f \,:\, P(f,N_f)=k, \\ t \in T_k^f}} (x_{k,l^0}^f(t) - x_{k,l^0}^f(t-1)) \leqslant A_k(t)$$

$$\sum_{\substack{f \,:\, P(f,i)=j,\ P(f,i+1)=j', \\ i<N_f,\ l \in \mathcal{L}_j^f,\ l' \in \mathcal{L}_{j'}^f}} (x_{j,l}^f(t) - x_{j',l'}^f(t)) \leqslant S_j(t)$$

At time *t*, *f* is entered (at any level *l*) and not yet exited (at any level *l’*, towards any sector *j’*) from sector j

Other constraints

- Variable $\boldsymbol{x}$ consistency (= 1 from some t on)

$$x^f_{j,l}(t) - x^f_{j,l}(t-1) \geqslant 0$$

$$x^f_{P(f,N_f),0}(t^f_{\text{last}}) = 1$$

- Variables $\boldsymbol{x}$ determine variables $\boldsymbol{y}$

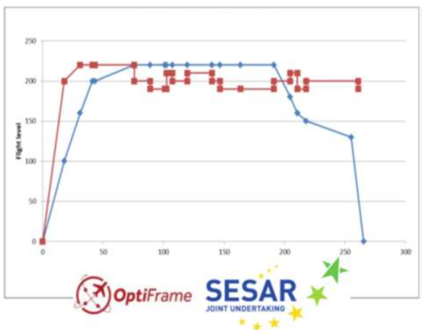
$$y^f_{j,l}(t) \;\geqslant\; x^f_{j,l}(t) \;-\; \sum_{\substack{l' \in \mathcal{L}^f_{j'},\\ j' = P(f,i+1)}} x^f_{j',l'}(t)$$

- Domains

$$x^f_{j,l}(t),\; y^f_{j,l}(t),\; u_{a,l}(t) \in \{0,1\}$$

Research motivations

- "We **cannot** fly that trajectory" (too many vertical / horizontal deviation, «technical» limitations etc.)
  - Add constraints, but lose model structure



- "We **prefer** not to fly that trajectory - sorry, cannot say why" (business model is hardly revealed)

Basic idea: consider deviating flights on ***historical*** **trajectories** and ***learn*** **users' preferences** from data repositories

IP model based on trajectory selection

- ***Given**** sets $\boldsymbol{P(f)}$ of "possible" 4D trajectories for each flight $\boldsymbol{f}$
- ***Given**** parameters $\boldsymbol{G(p,f)}$: preference of flight $\boldsymbol{f}$ to fly trajectory $\boldsymbol{p \in P(f)}$
- *Variables* directly model the selection of one trajectory:
  $\boldsymbol{y_p^f}$: 1 if flight $\boldsymbol{f}$ flies trajectory $\boldsymbol{p \in P(f)}$, 0 otherwise
- *Objective*: minimize $\boldsymbol{delays}$ / maximize $\sum_{f,p} \boldsymbol{G_p^f\, y_p^f}$
  - Combine objectives or multi-objective approach
- *Constraints*: assign one trajectory to each flight, sector capacities

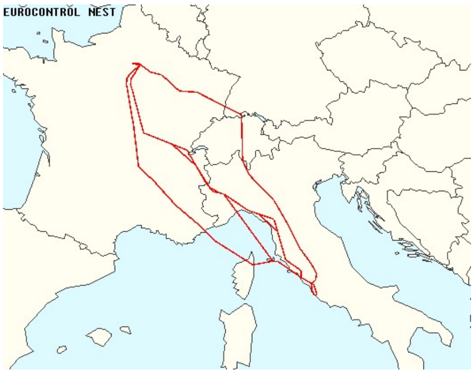
*** Learn $\boldsymbol{P(f)}$ and $\boldsymbol{G_p^f}$**

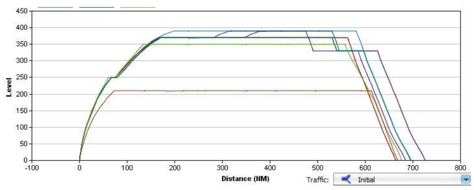
from historical data on actually flown trajectories

Learning $\boldsymbol{P(f)}$: methodology

1. Extract all the DDR2 trajectories between (almost) same origin-destination pair as f
2. Determine (subsets of «similar») possible trajectories
   - Use **clustering** to filter noise (and keep the number of variables limited)
3. Extract reference trajectories to feed ATFM models
   - $\boldsymbol{P(f)}$ contains the clustered trajectories (or some representatives, e.g. the 1-center of each cluster)
   - For each trajectory determine: airborne delay, airspace capacity utilization $A_{ps}(t)$ … (from DDR2)

Learning $P(f)$: DDR2 trajectories



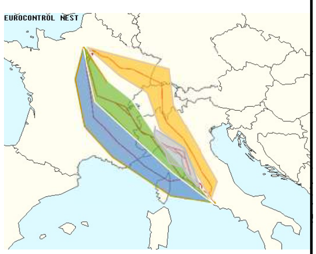


- Sequence of points in 4D: WPx, WPy, FL, time
- Different 3D route structures
- Different speeds
- Other available info: callsign, A/C type, STD, cost, etc.

Learning $P(f)$: trajectory clustering

- **Resample** each trajectory: from differently many WPs to a **same number** *n* of *equidistant* 4D points
  - each trajectory is represented as a vector of *n* x 4 elements $(x_1,y_1,FL_1,t_1 , x_2,y_2,FL_2,t_2 , x_3,y_3,FL_3,t_3 , \ldots)$
- Run a **Principal Component Analysis** to reduce dimensionality (keep most variance)
- Run a **clustering algorithm** (DBSCAN: flexibility, outliers detection)

[*similar to Gariel et al. 2011, Liu et al. 2017*]

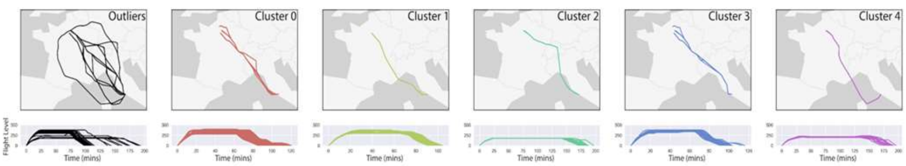


Example: DDR2 data for Rome-Paris

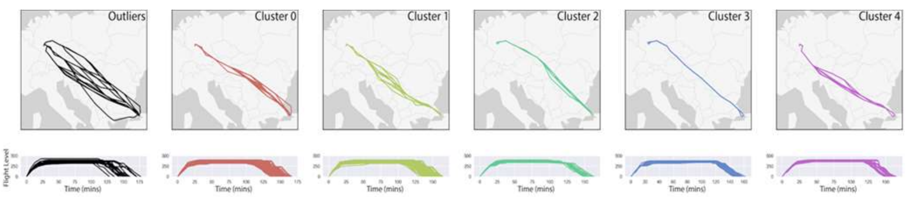
- From June 15 to September 15, 2016
- ~ 2000 flights (LIRF, LIRA) → (LFPG, LFPO, LFOB)
- 16 to 38 WPs, 76 resampled hits

Example: clustering of trajectories

Rome - Paris



Istanbul - Frankfurt



	Outliers	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Rome-Paris	31	1616	223	32	39	16
Istanbul-Frankfurt	22	519	310	37	25	22

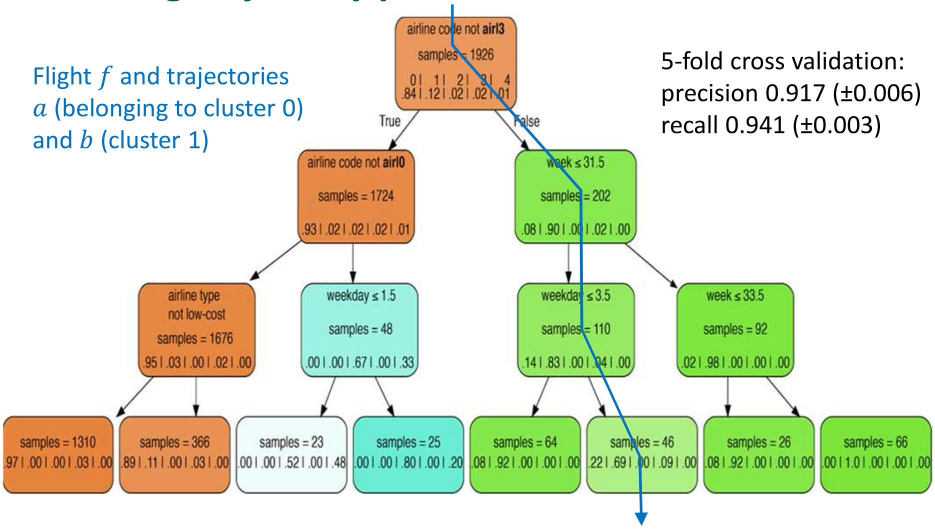
Learning trajectory preferences G_p^f

1. **Learn a tree classifier** to predict the cluster flown by a flight based on some flight features
   - day of the week, week number (seasonal effects), part of the day (morning, afternoon, evening, night), airline code, airline type (legacy/low-cost), aircraft model
2. **Use the tree classifier** and count, for each leaf l and cluster c, the number $n[l, c]$ of flights in l flying a trajectory in c
3. G_p^f is obtained by **normalizing** $n[l(f), c(p)]$, where $l(f)$ is the leaf reached by f and $c(p)$ is the cluster of p

Learning trajectory preferences: Rome-Paris

Flight f and trajectories a (belonging to cluster 0) and b (cluster 1)

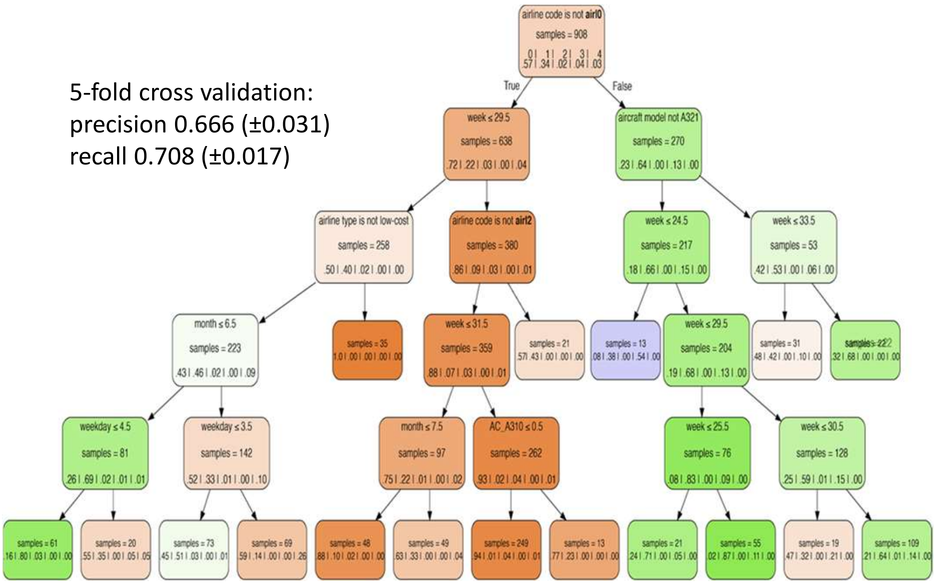
5-fold cross validation:
precision 0.917 (±0.006)
recall 0.941 (±0.003)



$G_a^f = 0.22$ $G_b^f = 0.69$

Learning trajectory preferences: Istanbul-Frankfurt

5-fold cross validation:
precision 0.666 (±0.031)
recall 0.708 (±0.017)



Learning trajectory preferences: interrelation strength between flight features

Cramer's V index (Bergsman's bias correction) between cluster and flight features

	airline	legacy / low cost	aircraft model	day part	weekday	week	month
Rome - Paris	0.57	0.62	0.45	0.18	0.07	0.05	0.01
Istanbul - Frankfurt	0.28	0.17	0.22	0.11	0.10	0.20	0.18

- Preference model performance and interrelation depend on O/D pair
- Towards determining trajectory determinants
  - Include further flight features and avoid “airline”
  - May provide a better trajectory-preference model

Conclusions and perspectives

- ILP formulation for ATFM based on 4D trajectory selection
  - More realistic solutions
  - Take user preference into account
- Data analytics to determine model parameters
  - Identify typical trajectories via clustering
  - Learn clusters and related preferences via tree classifiers
- Future work
  - Plug results from DDR2 into the ILP model
  - Evaluate performance and assess possible benefits
  - Improve the trajectory preference model
  - Allow further trajectories by column generation

Some references

- V. Dal Sasso, F. Djeumou Fomeni, G. Lulli, K.G. Zografos: Incorporating Stakeholders' Priorities and Preferences in 4D Trajectory Optimization. Transportation Research B 117, pp. 594-609, 2018
- V. Dal Sasso, F. Djeumou Fomeni, G. Lulli, K.G. Zografos: Planning efficient 4D trajectories in Air Traffic Flow Management, To appear in European Journal of Operational Research
- M. Gariel, A.N. Srivastava, E. Feron: Trajectory clustering and an application to airspace monitoring. IEEE Transactions on Intelligent Transportation Systems 12(4), pp. 1511-1524, 2011
- C. Lancia, L. De Giovanni, G. Lulli: Data analytics for trajectory selection and preference-model extrapolation in the European airspace. To appear in: M. Labbé and B. Fortz, (eds), The OR 2018 Proceedings, Springer Verlag, Brussels, 2018
- C. Lancia, G. Lulli: Predictive modeling of inbound demand at major European airports with Poisson and Pre-Scheduled Random Arrivals. Submitted to European Journal of Operational Research (preliminary version available at https://arxiv.org/pdf/1708.02486.pdf)