Traffic Forecasts, Delay and Costs
A Backcasting Exercise

6th InterFAB Expert Talks: ATM performance data - can we do better?

08 September 2021

Thomas Standfuß
Institute of Logistics and Aviation, Chair of Air Transport Technology and Logistics

Ibon Galarraga
Basque Centre for Climate Change / Metroeconomica

Agenda

Starting Point: Accuracy of Forecasts

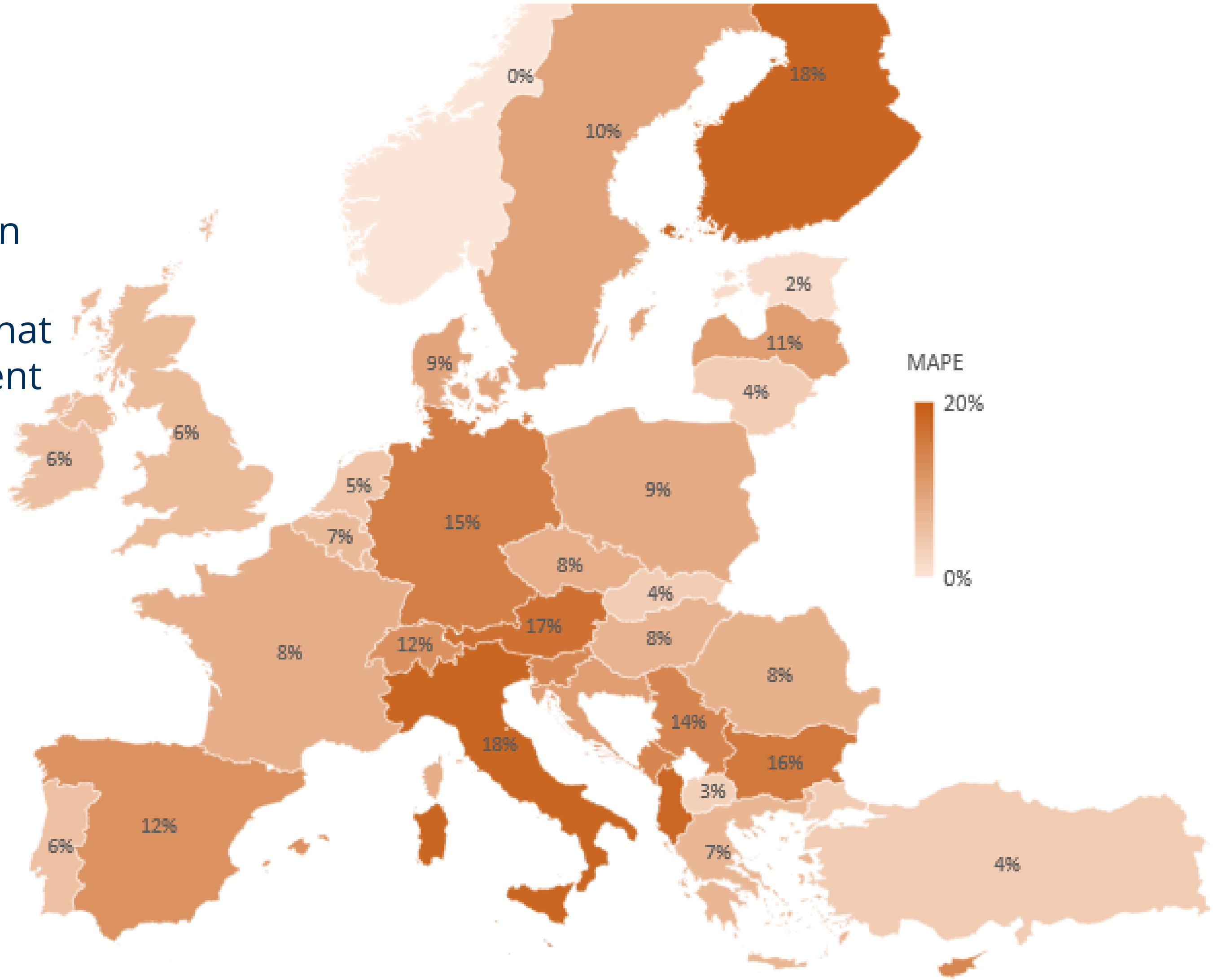
Why to apply a backcasting?

Forecast Quality and Consequences

- Studies showed that forecast quality of European air traffic is limited:
  - Applying metrics (e.g., MAPE-Score) shows that a majority of ANSPs are faced with insufficient forecasts.
  - Despite a wide confidence interval, most of the prediction do not match.
- In case of under-estimations, the probability of delay increases.

What is a "backcasting"

- Backcasting is an exercise to predict the past.
- It enables the evaluation of a forcasting model (ex-post).



Data and Method

How to apply the backcasting method?

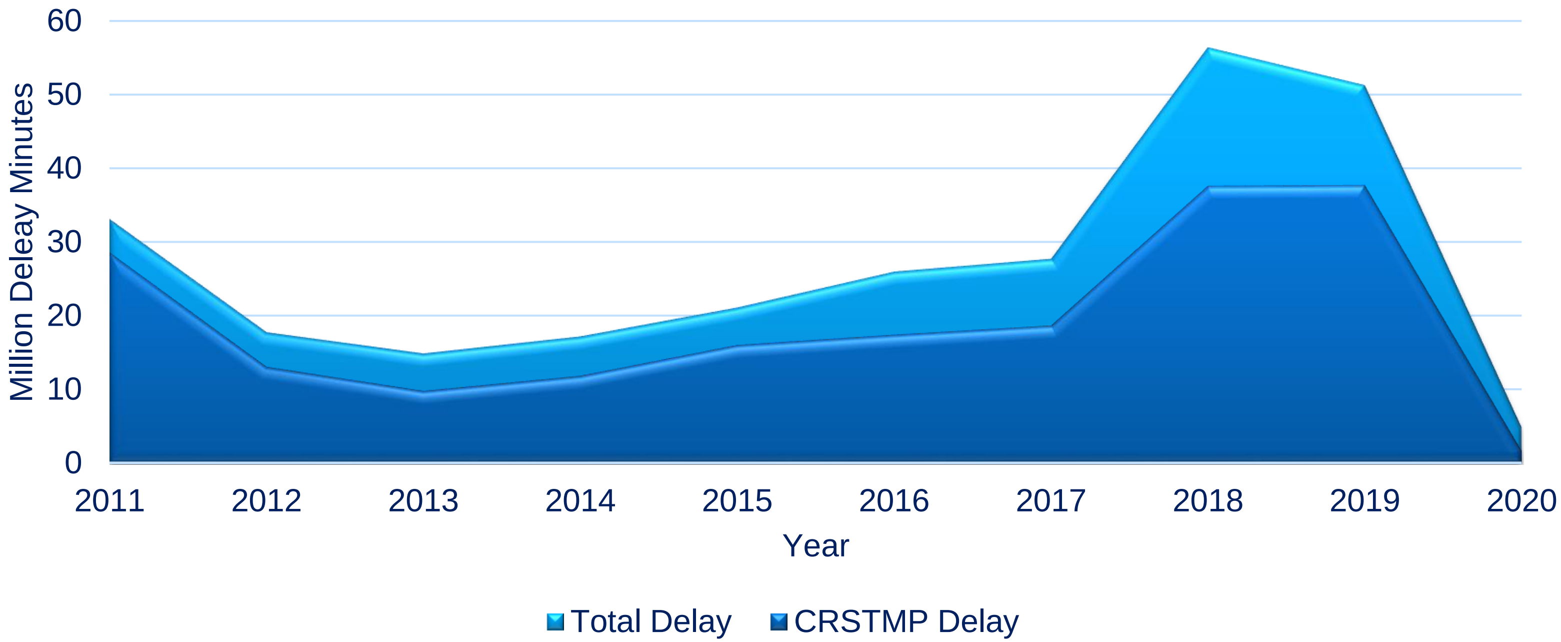
Analysis Steps

1. Data collection:
    1. Forecasted and Actual flights → Source: STATFOR Spring Reports, PRU data
    2. CRSTMP Delay → Source: PRU data
    3. Delay Targets → Source: Performance Plans
    4. Planned Scenario (Low, Base, High) → Source: Performance Plans
2. Identfying saturated airspaces in Europe.
3. Check, in which years an under-estimation occured.
4. Clustering of flights in to classes. Calculation of Average Delay Minutes and Probability of delay target missmatch.
5. Backcasting
    1. Reduction of flights and reclaculation of delay.
    2. Calculation of difference between actual delay and backcasted delay as well as subsequent costs.

Data and Method
Which delay to be used?

Source and Types of Delay:

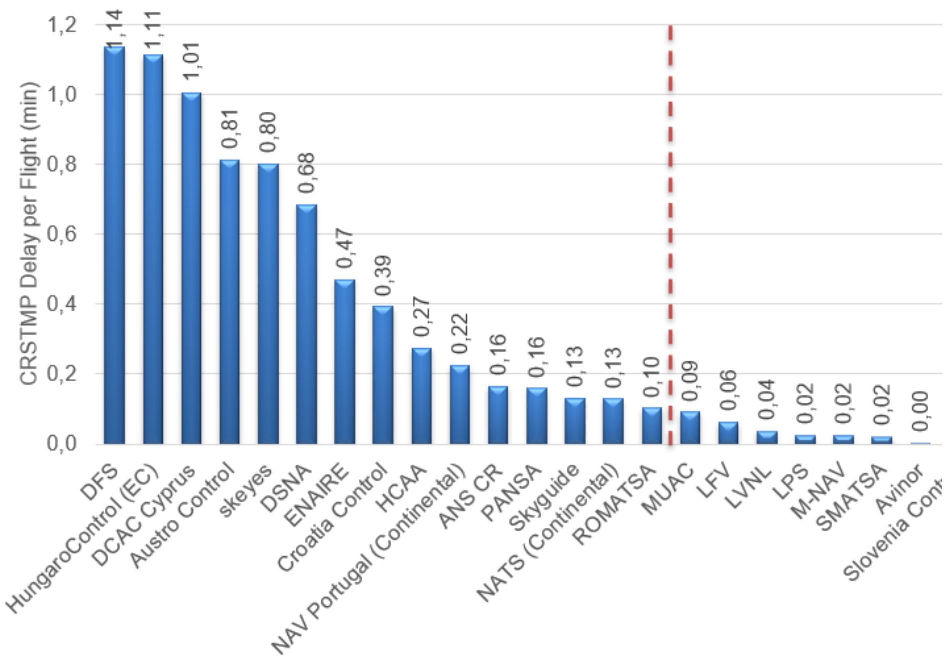
- The analysis is based on PRU data
- The database contains Flights, Delay (Minutes), Delay Causes and Number of delayed Flights.
- Delay is distinguished into several causes (figure on the right, bottom). For investigation, only CRSTMP-delay causes will be included:
  - C - ATC Capacity
  - R - ATC Routeing
  - S - ATC Staffing
  - T - Equipment (ATC)
  - M - Airspace Management
  - P - Special Event

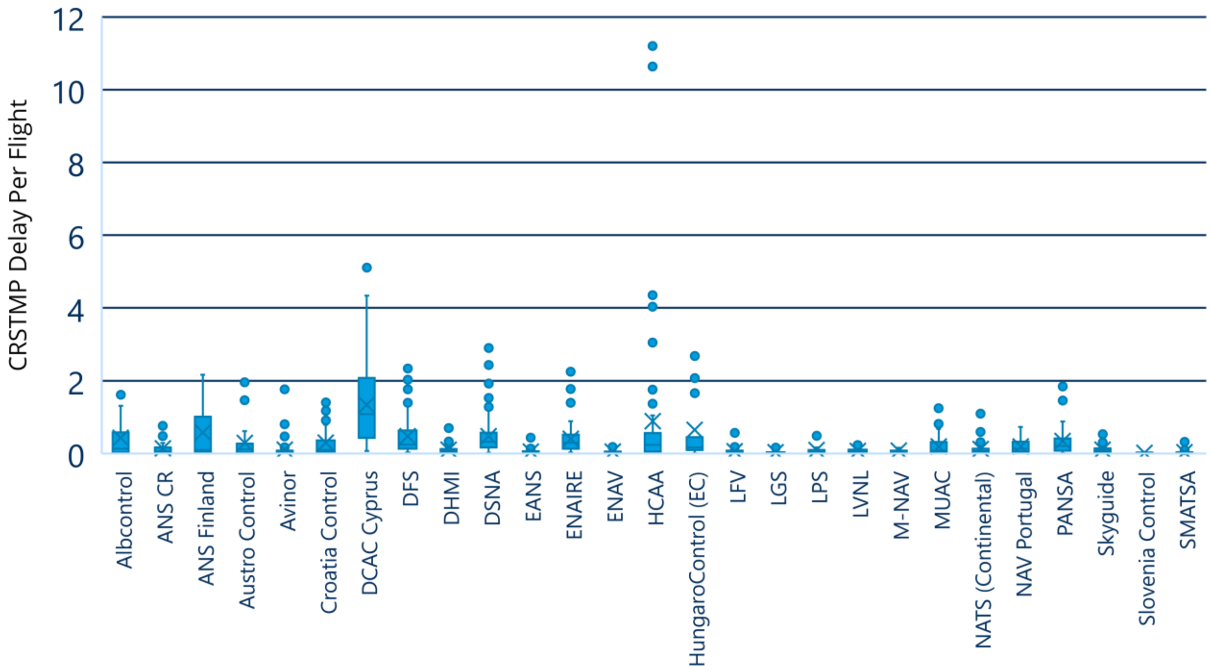


Data and Method
Which airspaces are saturated?

Saturated Airspaces

- Airspaces considered as saturated match two criteria:
  - Average delay minutes higher (or equal) to 0.1 minutes per flight,
  - A high value of average delay minutes is not caused by extreme values.





Application Part 1: Backcasting the Delay

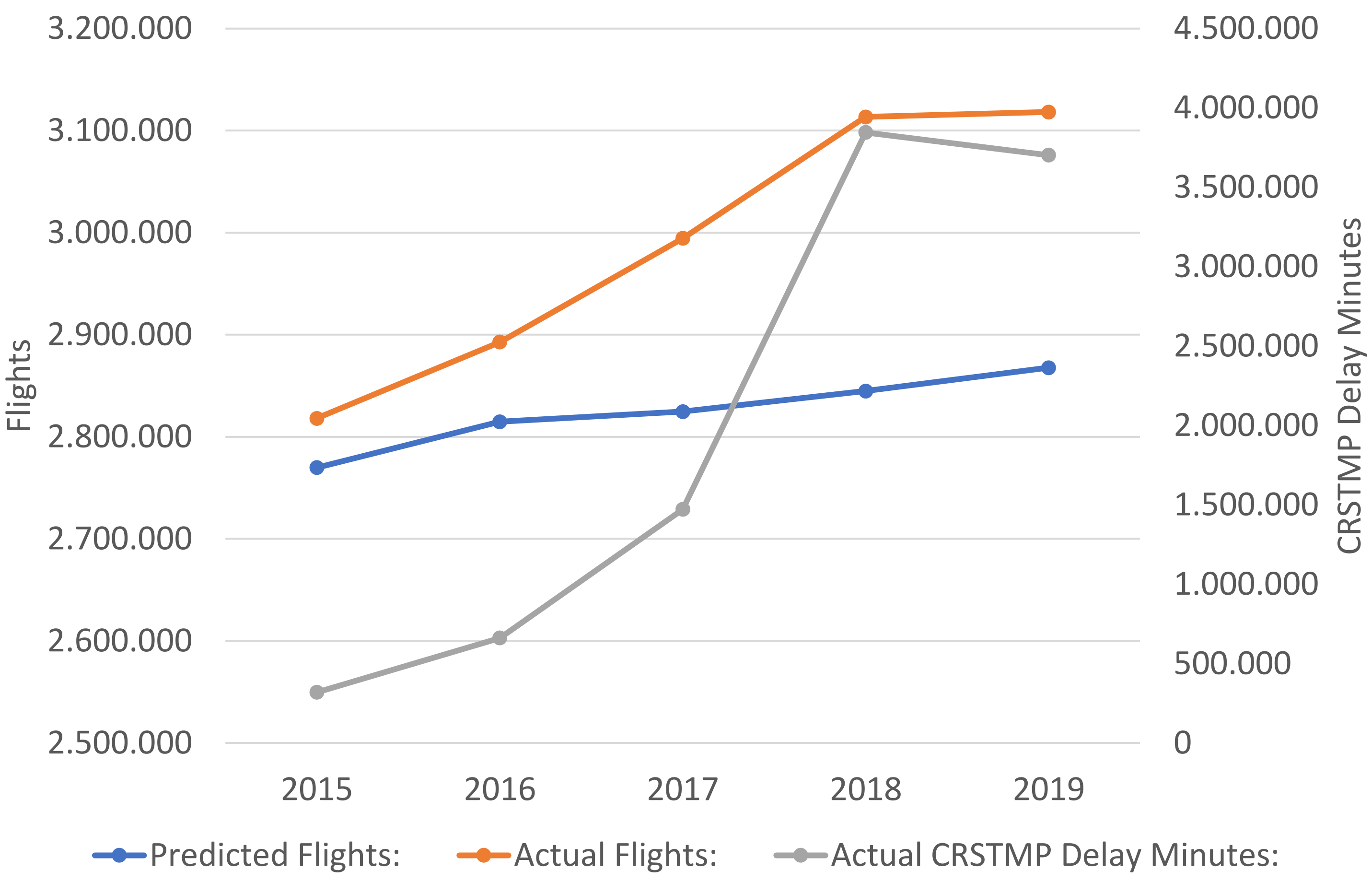
How to approach the Backcasting?

Situation 2015, spring report:

Predicted 5-year growth:	3.4%
Flights 2014:	2.772.617
Forecasted Flights 2019:	2.867.783

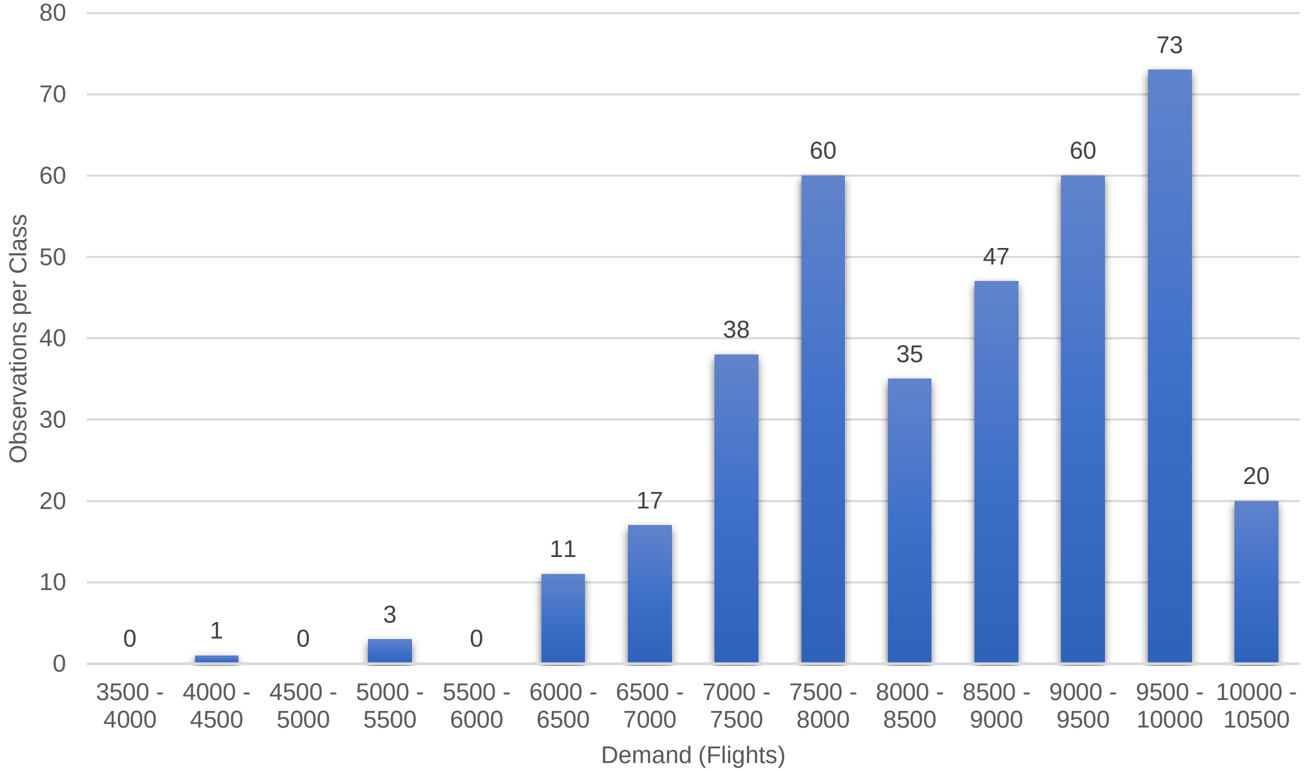
Actual Situation in 2019:

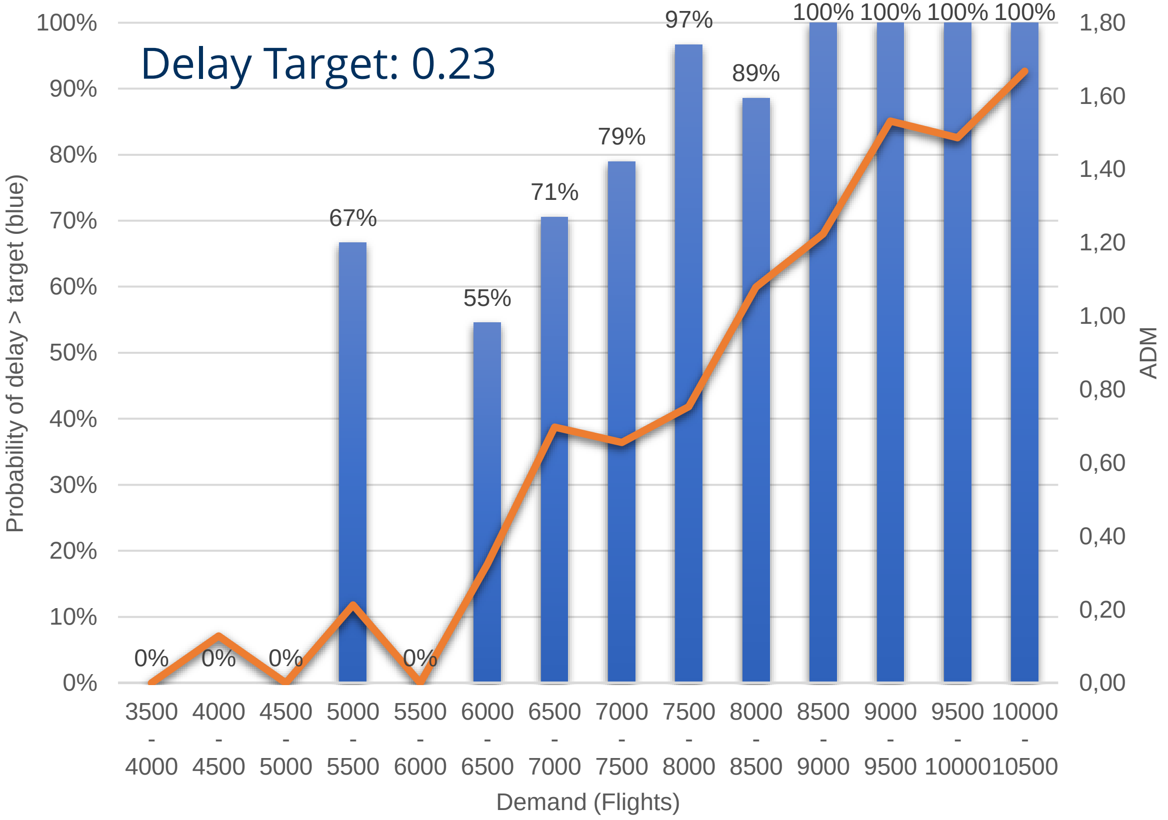
Flights:	3.118.176
Difference :	250.393 (8%)
CRSTMP Delay Minutes:	3.702.239
Average Delay per Flight:	1.19



Question: How does the deviation affect the delay? What if actual = predicted?

Application Part 1: Backcasting the Delay
How does it look like?





→ Each class has a different number of observations – Identification of extreme values (ADM) possible.

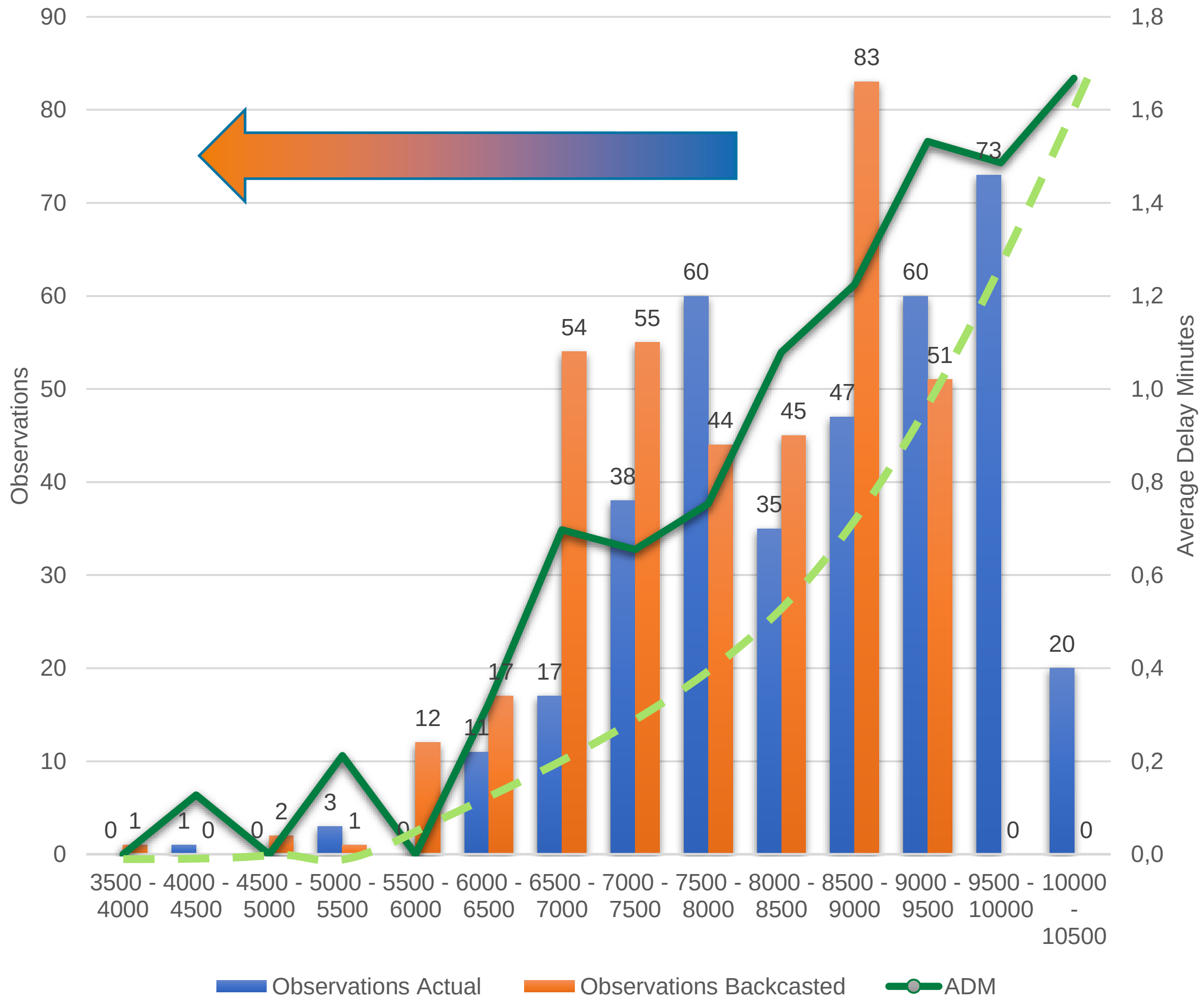
→ The more flights, the higher the probability of mismatch the delay target.

→ ADM-curve might be seen as an approximation for the Interdependency between demand and delay.

Application Part 1: Backcasting the Delay
How to calculate the backcasted delay?

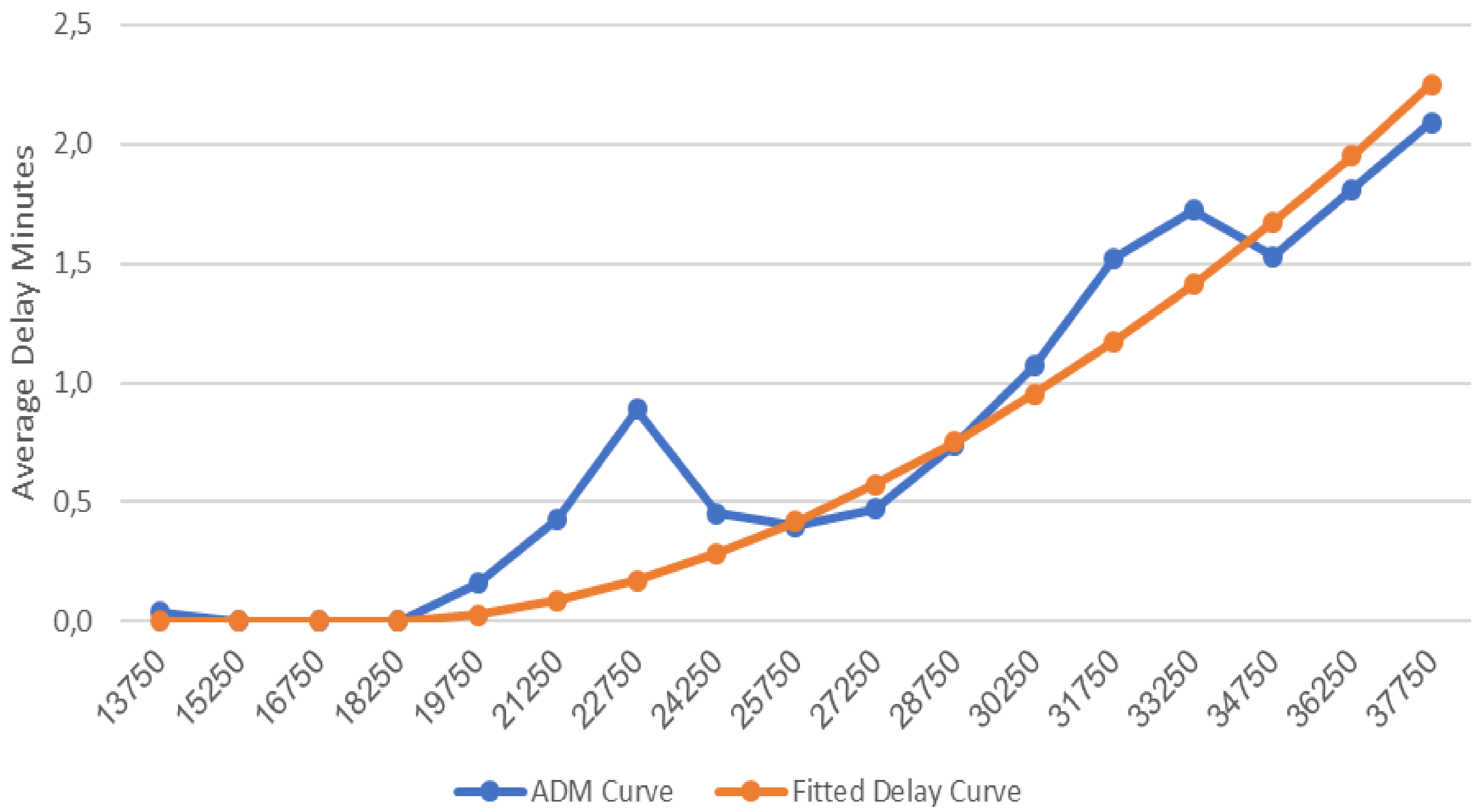
Assumptions:

- The reduction of flights is equally distributed over the year. Subsequently, the number of flights of each day are reduced by 8%.
- The average delay minutes (green curve) is constant for each class. It might be seen as a proxy for capacity.
- However, by reducing the flights, the number of observations per class shifts to the left (arrow).
- Since delay minutes are usually lower for classes with a lower number of flights, the overall delay will decrease as well.
- Problem: In case of peaks in the ADM curve, backcasted delay might be biased. Thus, there is a need for correction (dashed green curve).



Application Part 1: Backcasting the Delay

How to fit the delay curve



Fitted curve:
y=ax^b

with a=0.8
b=1.8

R^2=0.88

y… delay
x… flights

Application Part 2: Declustering the Data

Why to get rid of the clustering?

Mainly to improve quality of:

- The mathematical function to relate delays and flights.
- To extend the analysis.

How is the analysis done?

1. Daily flights and delay minutes are used.
2. Clustered delay analysis is omitted, and
3. an exponential flattening is applied to the non-clustered data.

- Then, **backcasted daily delays** (in minutes) (BD) for all ANSPs and years are estimated,
- and divided by the number of delayed flights to obtain the **backcasted mean time of delay** (in minutes).
- This way, backcasted delays are explored with a **different function for each year** following an exponential flattening.

Application Part 2: Declustering the Data
Economic cost of delays

Information on costs per minute from University of Westminster's work for several ranges and apply it to the number of flights delayed due to forecast quality.

This is done in 2 steps:

1. **Estimate cost functions**: using the 10 cost ranges from University of Westminster in a similar approach to what we did in earlier paper.

For delays ≤ 74.5 minutes $\alpha_1 D + \beta_1 D^2$ [1]

For delays > 74.5 minutes $\alpha_2 D^{\beta_2} + \gamma_2$ [2]

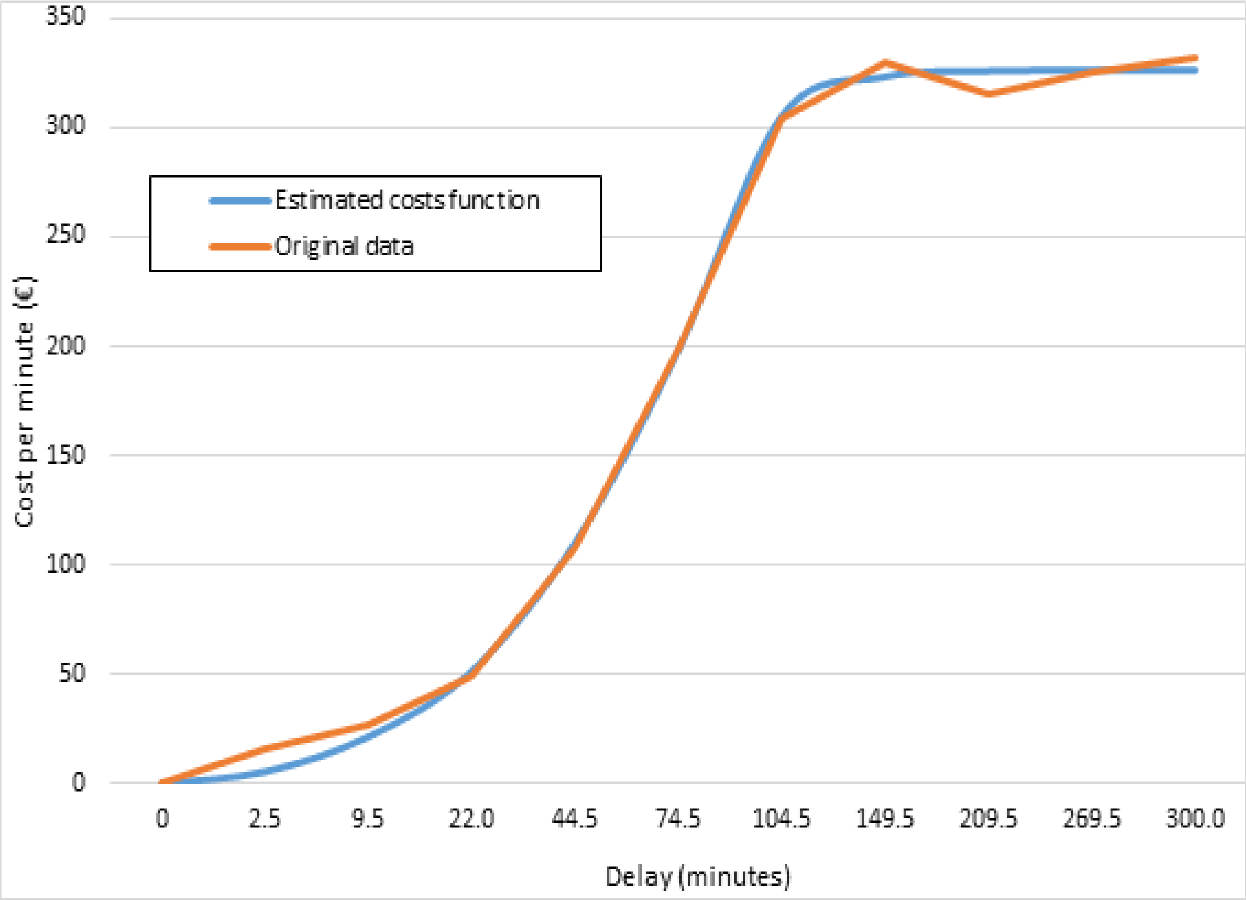
Where D stands for delays and the annual parameters are α_1, β_1,α_2, β_2 and γ_2.

See: University of Westminster (2003). Successive revisions in 2011 and 2015 and Ruiz-Gauna, I. et al (2020).

Application Part 2: Declustering the Data
Economic cost of delays

2. **Apply cost functions**:

- The second step is to applying these cost functions to the number of minutes of actual and backcasted delays.
- Very good fit!



Results

1. **Flights:**
   - **Actual flights** (i.e. the number of real flights occurring each year) **were higher than expected for all ANSP** (except for Croatia Control in 2015 and 2016, Austro Control in 2016 and DCAC in 2016).
   - **The difference between Actual and Backcasted increases with time**. That is, the further is the year the poorer is the forecast.

Difference between actual and backcasted flights, in percentage (%)

Flights Back % Actual	2015	2016	2017	2018	2019
DFS	-1.71	-2.70	-5.67	-8.63	-8.03
DSNA	-1.60	-4.22	-7.83	-9.67	-10.22
ANSCR	-1.04	-3.96	-3.32	-7.48	-3.89
AustroControl	-2.39	0.17	-5.15	-9.12	-10.29
CroatiaControl	0.25	3.03	-2.46	-9.10	-14.95
DCAC	-2.43	0.66	-7.84	-13.01	-14.36
MUAC	-1.19	-3.84	-6.93	-7.29	-5.94
NATS	-1.42	-4.98	-6.96	-6.16	-5.26
NAV Portugal	-2.07	-9.67	-16.17	-17.13	-17.66
skeyes	-3.54	-2.56	-7.43	-9.56	-7.33
skyguide	-1.48	-1.89	-4.53	-8.23	-7.84

Actual Flights, 2015-2019

Actual Flights	2015	2016	2017	2018	2019
DFS	2,818,110	2,892,868	2,994,472	3,113,468	3,118,176
DSNA	2,918,465	3,051,154	3,173,063	3,257,894	3,302,045
ANSCR	730,979	781,615	797,657	856,742	850,179
AustroControl	915,007	918,769	994,781	1,063,825	1,108,735
CroatiaControl	530,607	533,791	581,327	640,384	707,995
DCAC	319,091	322,214	359,540	393,558	411,460
MUAC	1,702,263	1,779,969	1,848,581	1,872,690	1,862,754
NATS	2,268,666	2,399,723	2,490,666	2,514,044	2,536,427
NAV Portugal	501,873	556,204	610,028	630,321	647,617
skeyes	591,480	595,248	628,705	649,574	639,865
skyguide	1,184,665	1,205,751	1,242,610	1,303,816	1,310,481

Backcasted Flights, 2015-2019

Backcasted Flights	2015	2016	2017	2018	2019
DFS	2,769,872	2,814,708	2,824,775	2,844,906	2,867,783
DSNA	2,871,869	2,922,527	2,924,473	2,942,983	2,964,414
ANSCR	723,356	750,690	771,190	792,666	817,070
AustroControl	893,176	920,289	943,528	966,768	994,655
CroatiaControl	531,908	549,972	567,034	582,088	602,160
DCAC	311,335	324,350	331,357	342,369	352,380
MUAC	1,681,924	1,711,533	1,720,417	1,736,235	1,752,114
NATS	2,236,383	2,280,253	2,317,298	2,359,217	2,403,087
NAV Portugal	491,479	502,424	511,377	522,321	533,265
skeyes	570,567	580,010	581,999	587,466	592,933
skyguide	1,167,171	1,182,959	1,186,342	1,196,491	1,207,768

Results

2. Delays:

— **Backcasted CRSTMP Delay Minutes are lower than actual CRSTMP Delay Minutes** for both clustered and non clustered data.

Actual and Backcasted CRSTMP Delay Minutes, in aggregate terms, by applying the clustered analysis and the non-clustered daily data

	Aggregate						
	Actual	Clustered	Diff		Non-clustered	Diff	
			Absolute	%		Absolute	%
DFS	10,000,801	7,292,994	2,707,807	▼37.13	7,013,793	2,987,008	▼42.59
DSNA	12,065,354	7,482,961	4,582,393	▼61.24	6,560,528	5,504,826	▼83.91
ANSCR	501,157	385,155	116,002	▼30.12	378,808	122,349	▼32.30
AustroControl	1,491,419	770,489	720,930	▼93.57	924,079	567,340	▼61.40
CroatiaControl	856,971	271,062	585,909	▼216.15	640,629	216,342	▼33.77
DCAC	2,274,296	1,267,836	1,006,460	▼79.38	1,497,656	776,640	▼51.86
MUAC	2,841,932	2,068,618	773,314	▼37.38	1,928,653	913,279	▼47.35
NATS	1,744,993	1,050,367	694,626	▼66.13	1,207,161	537,832	▼44.55
NAVPORTUGAL	715,843	469,577	246,266	▼52.44	398,796	317,047	▼79.50
skeyes	1,020,458	796,628	223,830	▼28.10	822,949	197,509	▼24.00
skyguide	791,310	503,002	288,308	▼57.32	455,815	335,495	▼73.60
FABEC	**26,719,85[illegible]**	**18,144,204**	**8,575,651**	**▼47.26**	**16,781,738**	**9,938,117**	**▼59.22**
Non-FABEC ANSPs	**7,584,679**	**4,214,486**	**3,370,193**	**▼79.97**	**5,047,130**	**2,537,549**	**▼50.28**

For example FABEC:

— **Actual delays 26 mill minutes.**

— **Backcasted between 16-18 mill minutes.**

Results

3. Costs:

- **The cost of delay (€) is much lower for Backcasted delays than the Actual ones**. (Or much higher for Actual delays than for Backcasted delays).
- Estimated figures **for the non-clustered daily data are significantly lower** than estimated by the clustered analysis.

For example FABEC:

- **Actual cost between 1-2,7€ bill.**
- **Backcasted 0,5-1,9 € bill.**
- **Between 30-50% lower.**

Actual and backcasted delay costs for the clustered analysis and the non-clustered daily data

Backcasted CRSTMP Delay Costs	Actual			Backcasted		
	Clust.	Non-clustered	Non-clustered /clustered	Clust.	Non-clustered	Non-clustered /clustered
DFS	1,040,618,988	324,941,956	31.23%	758,700,509	208,031,038	27.42%
DSNA	1,239,562,918	518,996,108	41.87%	764,456,390	237,571,828	31.08%
ANSCR	52,358,206	16,503,929	31.52%	40,238,200	8,124,515	20.19%
Austro Co	156,904,810	62,793,788	40.02%	81,071,572	25,339,552	31.26%
Croatia Co	88,991,924	30,398,476	34.16%	28,381,494	14,552,025	51.27%
DCAC	232,735,060	117,408,200	50.45%	128,928,025	58,189,060	45.13%
MUAC	290,616,096	92,984,630	32.00%	210,960,012	50,448,776	23.91%
NATS	179,323,380	79,111,374	44.12%	107,502,119	35,104,761	32.65%
NAV Port	73,148,906	33,091,754	45.24%	47,740,483	6,452,561	13.52%
skeyes	105,730,424	46,499,315	43.98%	82,416,297	18,038,325	21.89%
skyguide	81,630,248	31,809,108	38.96%	51,602,706	14,353,437	27.82%
FABEC	**2,758,178,674**	**1,015,231,116**	**36.81%**	**1,868,135,914**	**528,443,404**	**28.29%**
Non-FABEC ANSPs	**783,462,286**	**339,307,320**	**43.31%**	**433,861,893**	**147,762,474**	**34.06%**

TECHNISCHE UNIVERSITÄT DRESDEN

Metroeconomica

Traffic Forecasts, Delay and Costs - InterFAB Expert Talks
TU Dresden & Metroeconomica / Thomas Standfuß & Ibon Galarraga
ATM performance data - can we do better? // 08.09.2021

DRESDEN concept

Concluding Remarks

- Poor forecast is very likely to generate significant delays due to problems to adequately manage the real demand. **Actual delays were higher than Backcasted delays**.
- Delays generate very significant additional costs to the system. As much as 200% increase in cost across the saturated airspace.
- **The cost of Actual delays is much higher than for Backcasted delays**. And the non-clustered daily data are significantly lower than for the clustered analysis. As argued, this is mainly due to the improvements in estimation accuracy.
- Being able **to forecast adequately** expected number of flights, expected delays as well as corresponding costs is **very valuable for a good management** of the air navigation system.
- **Backcasting** exercises are **very useful** efforts to:
  - Predict what may have happened if things were planned differently.
  - Test the goodness of forecasting techniques.
  - Give a taste with respect of the resilience of the air navigation services.
- **Investing in improvements of forecasting capacities is extremely likely to be very cost-efficient**.
- **Evaluations and adjustments of forecast on traffic, delay and emissions should be executed on a regular basis** to improve the method and reduce the total economic costs of the ATM system.