

Prescriptive Analytics in Airline Operations: Arrival time prediction and cost index optimization for short-haul flights

A APPENDIX: DETAILED LIST OF FEATURES

Table 1: Time and date features

Time dependent	Description	Structure	Transformation	Source
Hour	Block on/ block off-Date and Time	integer (1-23)	none	derived
Time of day	Block on- Date and Time & clustering of hours for one day	factor (6 levels)	none	derived
Month	Block on- Date and Time	numeric (1-12)	none	derived
Season	Block on-Date and Time & clustering of month	factor (4 levels)	none	derived
Weekday	Block on- Date and Time	numeric (1-7)	none	derived
Days to holiday	minimum days until the next major holiday	numeric	none	derived
Weekend	Weekend yes or no variable (yes=1)	binary (0/1)	none	derived

Table 2: Weather features

Weather	Description	Structure	Transformation	Source
Cloud base	TAF	numeric	none	existing
Vertical visibility	TAF	numeric	none	existing
Horizontal visibility	TAF	numeric	none	existing
Wind speed	TAF	numeric	none	existing
Wind gust	TAF	numeric	none	existing
Wind direction	TAF	numeric	none	existing
Weather description	TAF	numeric	Factor level conversion	existing

Table 3: Aircraft and flight features

Flight	Description	Structure	Transformation	Source
Aircraft type	combination of letters and number	factor (3 levels)	none	existing
Fuel type	numbers	factor (3 levels)	none	existing
Air distance	Total flight air distance	numeric (km)	none	existing
Great circle distance	Total flight ground distance	numeric (km)	none	existing
Maximum altitude	Maximum altitude reached during the flight	numeric	none	existing
Total weight	Take-off weight of aircraft	numeric (kg)	none	existing
Actual total trip time	Actual block off time until actual block on time	numeric (min)	none	existing
Scheduled total trip time	Scheduled block off time until scheduled block on time	numeric (min)	none	existing
Taxi-in Time	Touchdown until block on time	numeric (min)	none	existing
Taxi-out Time	Block off until take-off time	numeric (min)	none	existing
ETA Time	Scheduled block off time until scheduled touchdown time	numeric (min)	none	derived
Flight Time	Scheduled take-off time until scheduled touch down time	numeric (min)	none	derived

Table 4: Departure features

Departure Airport	Description	Structure	Transformation	Source
Departure airport	IATA code	numeric	Factor level	existing
Departure runway	combination of letters and number	numeric	conversion Factor level	existing
Departure stand	combination of letters and number	numeric	conversion Factor level	existing
Country	Name	numeric	conversion Factor level	derived
City	Name	numeric	conversion Factor level	derived
Block off delay	Scheduled block off Time - Actual block off time	numeric (min)	conversion none	derived
Frequency of destination	number of times destination is visited within a year	numeric	none	derived

Table 5: Arrival features

Arrival Airport			Description	Structure	Transformation	Source
Number of arriving flights			Planned landing time used as the reference to calculate number of arriving flights Steps of 5 min (from 30 min prior to landing until 30 min after landing)	numeric	none	derived
Number of departing flights			Planned landing time used as the reference to calculate number of arriving/ departing flights Steps of 5 min (from 30 min prior to landing until 30 min after landing)	numeric	none	derived
Arrival airport			IATA code	factor (2 levels)	none	existing
Arrival runway			combination of letters and number	factor (12 levels)	none	existing

B APPENDIX: ENSEMBLE MODEL

S_0 = new variable subset

Stage 1 - Linear regression

S_1 = $S_0 \cup$ (Actual total trip time, scheduled total trip time, block off delay, ETA time, flight time, hour, time of day, season, days to holiday, arrival runway, air distance, great circle distance, total weight)

Algorithm: Linear Regression

Parameters: Default settings

Input: S_1 features

Output: Arrival time prediction (min.)

Stage 2 - Gradient boosting machine

S_2 = $S_0 \cup$ (S_1 , Departure stand, departure runway, departure airport, arrival airport, aircraft type, fuel type, maximum altitude, month, weekday, taxi-out time, taxi-in time, frequency of destination, weekend)

Algorithm: Gradient Boosting

Parameter settings: target variable = residual of previous stage, n.trees = 100, shrinkage = 0.1, interaction depth = 40, n.minobsinnode = 100

Input: S_2

Output: Residual of predicted arrival time (min.)

Stage 3 - Gradient boosting machine

S_3 = $S_0 \cup$ (S_1 , Weather description, cloud base, vertical visibility, horizontal visibility, wind speed, wind gust, wind direction)

Algorithm: Gradient Boosting

Parameter settings: target variable = residual of previous stage, n.trees = 100, shrinkage = 0.15, interaction depth = 40, n.minobsinnode = 100

Input: S_3

Output: Residual of predicted arrival time (min.)

Stage 4 - Gradient boosting machine

S_4 = $S_0 \cup$ (S_1 , Number of arriving flights, number of departing flights)

Algorithm: Gradient Boosting

Parameter settings: target variable = residual of previous stage, n.trees = 100, shrinkage = 0.15, interaction depth = 40, n.minobsinnode = 100

Input: S_4

Output: Residual of predicted arrival time (min.)

C APPENDIX: EXAMPLE RESULTS

Table 6: Example results for 20 flights

Route		Total trip time (min.)		
Departure	Arrival	Scheduled	Actual	Predicted
MUC	HUB	63	58.95	57
OSL	HUB	132	128.79	126
KBP	HUB	154	140.44	137
BIO	HUB	137	126.68	133
BCN	HUB	130	119.65	117
MAN	HUB	98	93.13	97
TLV	HUB	267	256.46	262
TXL	HUB	70	63.97	65
DUS	HUB	69	61.72	64
CPH	HUB	89	84.59	86
GOT	HUB	97	97.85	95
LED	HUB	168	167.33	169
MAD	HUB	143	142.85	132
MXP	HUB	86	77.02	92
ARN	HUB	138	122.09	122
HAM	HUB	66	64.04	72
HAJ	HUB	70	63.36	59
FCO	HUB	98	90.43	87
KBP	HUB	186	176.19	168
MUC	HUB	58	51.90	47
TUN	HUB	142	137.05	134
OPO	HUB	151	145.96	138

D APPENDIX: AIRCRAFT PERFORMANCE DATA FOR OD PAIRS

Table 7: HAM - HUB 1 (285 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	49.66	1.66	0.00%	1950.51	-57.49
10	0.35%	49.48	1.48	0.61%	1962.46	-45.54
20	1.82%	48.75	0.75	1.64%	1982.50	-25.50
30	3.45%	48.00	0.00	2.86%	2008.00	0.00
40	4.73%	47.31	-0.69	4.13%	2031.15	23.15
50	5.49%	46.93	-1.07	5.36%	2055.08	47.08
60	5.80%	46.78	-1.22	6.49%	2077.06	69.06
70	5.86%	46.75	-1.25	7.48%	2096.40	88.40
80	5.85%	46.75	-1.25	8.30%	2112.50	104.50
90	5.88%	46.74	-1.26	8.92%	2124.48	116.48
99	5.84%	46.75	-1.25	9.23%	2130.55	122.55
100	5.82%	46.76	-1.24	9.25%	2130.88	122.88

Table 8: MXP - HUB 2 (294 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	51.21	1.71	0.00%	1950.51	-57.49
10	0.35%	51.03	1.53	0.61%	1962.46	-45.54
20	1.82%	50.27	0.77	1.64%	1982.50	-25.50
30	3.45%	49.50	0.00	2.86%	2008.00	0.00
40	4.73%	48.79	-0.71	4.13%	2031.15	23.15
50	5.49%	48.40	-1.10	5.36%	2055.08	47.08
60	5.80%	48.24	-1.26	6.49%	2077.06	69.06
70	5.86%	48.21	-1.29	7.48%	2096.40	88.40
80	5.85%	48.21	-1.29	8.30%	2112.50	104.50
90	5.88%	48.20	-1.30	8.92%	2124.48	116.48
99	5.84%	48.22	-1.28	9.23%	2130.55	122.55
100	5.82%	48.23	-1.27	9.25%	2130.88	122.88

Table 9: DUS - HUB 2 (307 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	52.55	1.75	0.00%	2056.39	-60.61
10	0.35%	52.37	1.57	0.35%	2068.99	-48.01
20	1.82%	51.59	0.79	1.02%	2090.11	-26.89
30	3.45%	50.80	0.00	1.93%	2117.00	0.00
40	4.73%	50.07	-0.73	2.97%	2141.41	24.41
50	5.49%	49.67	-1.13	4.02%	2166.64	49.64
60	5.80%	49.50	-1.30	5.00%	2189.81	72.81
70	5.86%	49.48	-1.32	5.81%	2210.20	93.20
80	5.85%	49.48	-1.32	6.42%	2227.17	110.17
90	5.88%	49.46	-1.34	6.84%	2239.80	122.80
99	5.84%	49.48	-1.32	7.10%	2246.20	129.20
100	5.82%	49.49	-1.31	7.12%	2246.55	129.55

Table 10: MAN - HUB 1 (509 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	79.50	2.00	0.00%	3123.50	-61.50
10	0.11%	79.41	1.91	0.61%	3134.41	-50.59
20	1.20%	78.55	1.05	1.64%	3155.44	-29.56
30	2.58%	77.50	0.00	2.86%	3185.00	0.00
40	3.86%	76.44	-1.06	4.13%	3216.21	31.21
50	4.85%	75.64	-1.86	5.36%	3249.21	64.21
60	5.56%	75.08	-2.42	6.49%	3279.64	94.64
70	6.02%	74.71	-2.79	7.48%	3305.07	120.07
80	6.33%	74.47	-3.03	8.30%	3324.16	139.16
90	6.47%	74.36	-3.14	8.92%	3337.09	152.09
99	6.35%	74.45	-3.05	9.23%	3345.18	160.18
100	6.31%	74.48	-3.02	9.25%	3345.99	160.99

Table 11: LHR - HUB 2 (528 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	82.07	2.07	0.00%	3070.06	-60.44
10	0.11%	81.98	1.98	0.35%	3080.78	-49.72
20	1.20%	81.08	1.08	1.02%	3101.45	-29.05
30	2.58%	80.00	0.00	1.93%	3130.50	0.00
40	3.86%	78.90	-1.10	2.97%	3161.18	30.68
50	4.85%	78.08	-1.92	4.02%	3193.61	63.11
60	5.56%	77.51	-2.49	5.00%	3223.52	93.02
70	6.02%	77.12	-2.88	5.81%	3248.52	118.02
80	6.33%	76.87	-3.13	6.42%	3267.28	136.78
90	6.47%	76.75	-3.25	6.84%	3279.99	149.49
99	6.35%	76.85	-3.15	7.10%	3287.94	157.44
100	6.31%	76.88	-3.12	7.12%	3288.74	158.24

Table 12: FCO - HUB 1 (609 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	93.35	2.85	0.00%	3788.68	-63.82
10	0.24%	93.13	2.63	0.28%	3799.44	-53.06
20	1.53%	91.92	1.42	0.85%	3820.85	-31.65
30	3.15%	90.50	0.00	1.66%	3852.50	0.00
40	4.63%	89.03	-1.47	2.62%	3888.01	35.51
50	5.72%	88.01	-2.49	3.64%	3926.44	73.94
60	6.34%	87.43	-3.07	4.59%	3962.47	109.97
70	6.55%	87.24	-3.26	5.38%	3992.54	140.04
80	6.51%	87.28	-3.22	5.96%	4014.51	162.01
90	6.42%	87.36	-3.14	6.33%	4028.51	176.01
99	6.47%	87.31	-3.19	6.55%	4036.84	184.34
100	6.50%	87.29	-3.21	6.57%	4037.71	185.21

Table 13: ARN - HUB 2 (798 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	120.42	4.42	0.00%	4762.16	-64.84
10	2.20%	117.77	1.77	0.21%	4772.04	-54.96
20	3.22%	116.54	0.54	0.65%	4793.19	-33.81
30	3.81%	116.00	0.00	1.34%	4827.00	0.00
40	4.35%	115.18	-0.82	2.22%	4867.88	40.88
50	4.95%	114.45	-1.55	3.17%	4913.33	86.33
60	5.58%	113.70	-2.30	4.08%	4956.65	129.65
70	6.09%	113.08	-2.92	4.84%	4992.66	165.66
80	6.39%	112.72	-3.28	5.38%	5018.22	191.22
90	6.48%	112.62	-3.38	5.70%	5033.62	206.62
99	6.55%	112.53	-3.47	5.90%	5042.92	215.92
100	6.57%	112.50	-3.50	5.92%	5044.00	217.00

Table 14: DME - HUB 1 (1222 NM)

CI	Time Diff in %	Time (min.)	Deviation	Fuel Diff in %	Fuel (kg)	Deviation
0	0.00%	177.87	4.37	0.00%	6549.52	-61.98
10	0.67%	176.68	3.18	0.08%	6554.82	-56.68
20	1.63%	174.97	1.47	0.39%	6574.76	-36.74
30	2.52%	173.50	0.00	0.94%	6611.50	0.00
40	3.17%	172.23	-1.27	1.69%	6660.09	48.59
50	3.56%	171.54	-1.96	2.54%	6716.02	104.52
60	3.74%	171.21	-2.29	3.38%	6771.14	159.64
70	3.84%	171.04	-2.46	4.10%	6818.36	206.86
80	3.95%	170.84	-2.66	4.63%	6852.76	241.26
90	4.14%	170.51	-2.99	4.95%	6873.39	261.89
99	4.32%	170.18	-3.32	5.11%	6883.98	272.48
100	4.34%	170.15	-3.35	5.12%	6885.02	273.52