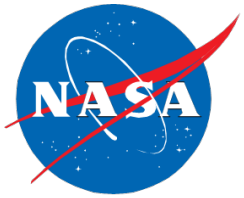

ATD-2 Integrated Arrival/ Departure/Surface (IADS) System Machine Learning Services

Estimated ON Model (EON)

Dan Wesely, Andrew Churchill, John Slough
Mosaic ATM Inc.
NASA Ames Research Center, Moffett Field, CA

William J. Coupe
NASA Ames Research Center, Moffett Field, CA

August 2021

Table of Contents

Model Name

Estimated ON Model (EON)

Problem Statement

Aircraft landing times and predictions thereof are of interest to many stakeholders and for many reasons. The physics of aircraft flight is well-understood, so physics-based models can be very accurate when configured appropriately and provided with accurate input data. The aim of the EON model service is to use machine learning (ML) to better predict the landing time of a flight than physics based approaches (i.e. best SWIM ETA).

Technical Approach

The ML algorithm used in the EON model is Random Forest Regression, an ensemble learning method. The ML models are not formulated to predict the time until the flight lands, but rather the error or residual in the mediated physics-based landing time prediction for a flight.

Model Features

Feature	Sample Value
Timestamp: time at which prediction is being made	4/16/2020 0:04
Mediated physics-based landing time prediction: combination of SWIM ETA values	4/16/2020 1:44
First airborne surveillance timestamp	4/15/2020 22:28
Actual take-off time	4/15/2020 22:27
Actual departure gate push-back time	4/15/2020 22:18
Airline-scheduled gate push-back time	4/15/2020 22:25
Airline-scheduled take-off time	4/15/2020 22:35
Airline-scheduled landing time	4/16/2020 1:42
TBFM arrival stream class	DAL_FINGR_JETS
Predicted arrival runway: from arrival runway model	31R
TBFM arrival meter fix	FINGR
Arrival runway assigned by TBFM	31R
Major air carrier	UAL
Aircraft type	WB

Model Inputs & Outputs

See OpenAPI specification in the appendix.

Data Sets

The model training and test dataset was extracted from a fused dataset of the Traffic Flow Management System (TFMS), Time-Based Flow Management (TBFM), and System Wide Information Management (SWIM) Terminal Data Distribution System (STDDS) Surface Movement Event Service (SMES) feeds. Features were derived from all three of these data sources.

To break the data into train and test sets, we randomly selected 30% of dates in the data set to be used for testing; all samples from flights that landed on those dates became test samples. This train-test split approach was used to avoid “leakage” in which samples from the same flight or flights that operated nearby in time end up in both the training and testing data sets.

Airport	Start Date	End Date	Training Samples	Test Samples
KCLT	01-01-2020	08-01-2020	197465	42316
KDAL	01-01-2020	08-01-2020	125740	23962
KIAH	01-01-2020	08-01-2020	215399	44751
KDFW	01-01-2020	08-01-2020	268301	59999
KJFK	01-01-2020	08-01-2020	251334	59865

Model Results / Evaluation

Errors

Airport	Mean Absolute Error: Training Dataset	Mean Absolute Error: Test Dataset	Mean Absolute Percent Error: Training Dataset	Mean Absolute Percent Error: Test Dataset
KCLT	159.56	179.47	39.93	46.55
KDAL	136.30	138.00	14.78	14.92
KIAH	231.55	257.28	23.99	22.20
KDFW	142.83	182.82	12.20	13.96
KJFK	219.40	455.70	10.10	23.18

Percent within X Seconds

Airport	Percent within 10 Seconds: Training Dataset	Percent within 10 Seconds: Test Dataset	Percent within 30 Seconds: Training Dataset	Percent within 10 Seconds: Test Dataset	Percent within 60 Seconds: Training Dataset	Percent within 60 Seconds: Test Dataset
KCLT	5.02	4.20	14.89	12.48	29.26	24.76
KDAL	7.56	7.10	22.28	20.88	41.17	39.44
KIAH	4.61	4.42	13.63	12.76	26.04	24.65
KDFW	8.08	6.81	22.07	18.91	39.41	34.24
KJFK	6.27	3.10	18.11	8.74	33.90	17.04

Comparison to Best SWIM Estimated Time of Arrival (ETA) Baseline

Airport	Approach	Mean Absolute Error: Training Dataset	Mean Absolute Error: Test Dataset	Percent within 60 Seconds: Training Dataset	Percent within 60 Seconds: Test Dataset
KCLT	Baseline	211.45	206.59	24.43	24.69
KCLT	ML Model	159.56	179.47	29.26	24.76
KDAL	Baseline	149.71	143.93	39.93	40.46
KDAL	ML Model	136.30	138.00	41.17	39.44
KDFW	Baseline	213.13	205.04	31.36	32.10
KDFW	ML Model	142.83	182.82	39.41	34.24
KIAH	Baseline	295.70	290.72	22.54	21.24
KIAH	ML Model	231.55	257.28	26.04	24.65
KJFK	Baseline	469.40	507.30	15.86	14.84
KJFK	ML Model	219.40	455.70	33.90	17.04

Open Source Repository

https://github.com/nasa/ML-airport-estimated-ON

Reference Documentation

Wesely, D., Churchill, A., Slough, J., Coupe, W., “A Machine Learning Approach to Predict Aircraft Landing Times using Mediated Predictions from Existing Systems,” AIAA AVIATION Forum, Washington, DC, USA, 2021.

Appendix: OpenAPI Specification

Explore

Estimated On Time Service Client

1.0.0 OAS3

Estimated On Time Service Client

Servers

http://localhost:9105/ - Estimated On Time Service Client

Computed URL: `http://localhost:9105/`

Server variables

port 9105

default

POST /eon-service

Schemas

EstimatedOnTimeRequest {

airport* string
example: KDFW
airport name

flights* [...]

}

EstimatedOnTimeRequestFlight {

aircraftType	**string** *example: E170* *nullable: true* **aircraft type**
arrivalMeterFixTbfm	**string** *example: VKTRY* *nullable: true* **TBFM meter fix**
arrivalRunwayBestTime	**string**($date-time) *example: 2020-07-06T13:46:01Z* *nullable: true* **best SWIM ETA**
arrivalRunwayScheduledTime	**string**($date-time) *example: 2020-07-06T13:42:54Z* *nullable: true* **airline-scheduled ON time**
arrivalRunwayTbfm	**string** *example: 13R* *nullable: true* **TBFM runway**
arrivalStreamClassTbfm	**string** *example: DFW_KLAWW_JETS* *nullable: true* **TBFM stream class**
departureRunwayActualTime	**string**($date-time) *example: 2020-07-06T13:10:00Z* *nullable: true* **actual OFF time**
departureRunwayScheduledTime	**string**($date-time) *example: 2020-07-06T13:09:00Z* *nullable: true* **airline-scheduled OFF time**
departureStandActualTime	**string**($date-time) *example: 2020-07-06T12:55:00Z* *nullable: true* **actual gate pushback time**
departureStandInitialTime	**string**($date-time) *example: 2020-07-06T13:00:00Z* *nullable: true* **first gate out time received for the flight**
gufi	**string** *example: ENY3575.OKC.DFW.200705.1300.0096.TFM* *nullable: true* **globally unique flight identifier**
majorCarrier	**string** *example: AAL* *nullable: true* **airline carrier**
predictedArrivalRunway	**string**

```
                                example: 13R
                                nullable: true

                                The predicted arrival runway

    timeFirstTracked            string($date-time)
                                example: 2020-07-06T13:10:33Z
                                nullable: true

                                time of first TFMS or TBFM surveillance data

    timestamp                   string($date-time)
                                example: 2020-07-06T13:20:00Z
                                nullable: true

                                time now (at prediction time)

}
```

EstimatedOnTimeResponse {

```
    airport*                    string
                                example: KDFW

                                airport name

    flights*
                                    [...]
}
```

EstimatedOnTimeResponseFlight {

```
    error*                      string
                                example: ABC123.CLT.MEX

                                error description

    gufi*                       string
                                example: ABC123.IAH.MEX

                                globally unique flight identifier

    pred*                       string
                                example: 2019-08-16 1:17:25

                                estimated on time without timezone

}
```