

London Heathrow Airport Uses Real-Time Analytics for Improving Operations

Xiaojia Guo

UCL School of Management, University College London, London, United Kingdom, x.guo.11@ucl.ac.uk

Yael Grushka-Cockayne

Harvard Business School, Cambridge, Massachusetts 02163, ygc@hbs.edu
and Darden School of Business, University of Virginia, Charlottesville, Virginia 22903

Bert De Reyck

UCL School of Management, University College London, London, United Kingdom, bdereyck@ucl.ac.uk

Improving airport collaborative decision making is at the heart of airport operations centers (APOCs), recently established in several major European airports. In this paper, we describe a project commissioned by Eurocontrol, the organization in charge of the safety and seamless flow of European air traffic. The project's goal was to examine the opportunities offered by the colocation and real-time data sharing in the APOC at London's Heathrow airport, arguably the most advanced of its type in Europe. We developed and implemented a pilot study of a real-time data sharing and collaborative decision-making process, selected to improve the efficiency of Heathrow's operations. In Guo et al. (2019), we describe the system that we developed for Heathrow to better estimate passenger connection times, which takes advantage of the integration and data sharing that Heathrow's APOC offers. In this paper, we describe the process of how we chose the subject of the pilot, namely the improvement of transfer-passenger flows through the airport, and how we helped Heathrow move from its existing legacy system for managing passenger flows to an advanced machine learning-based approach using real-time inputs. The system, which is now in operation at Heathrow, can predict which passengers are likely to miss their connecting flights, reducing the likelihood that departures will incur delays while waiting for delayed passengers. This can be done by offloading passengers in advance, by expediting passengers through the airport, or by modifying the departure times of aircraft in advance. By aggregating estimated passenger arrival time at various points throughout the airport, the system also improves passenger experiences at immigration and security desks, by enabling modifications to staffing levels in advance of expected surges in arrivals. The nine-stage framework we present here can support the development and implementation of other real-time data-driven systems. To the best of our knowledge, the proposed system is the first to use machine learning to model passenger flows in an airport.

Key words: Data-driven prediction; collaborative decision making; machine learning; airport performance

Heathrow Airport, the busiest airport in Europe, carries more than 80 million passengers each year to 203 destinations in 84 countries (Heathrow Airport 2018). The airport employs over 6,500 people and operates collaboratively with over 200 stakeholder organizations. Heathrow's Airport Operations Center (APOC), established in 2014 and the first of its kind in Europe, brings together all airport stakeholders in a single room (see Figure 1) with a goal of enabling collaborative decision making. Five years on, Heathrow's APOC has become the nerve center of all operational decision-making processes among stakeholders, capturing the benefits of proximity via improved collaborations, planning, real-time monitoring, proactive decision making, and flow management. Stakeholders in the APOC integrate information from shared data sources, develop joint plans, and execute plans that are within their respective area of responsibility. Roles within the APOC include managers for the airport operations, aircraft flow, security flow, and passenger flow, operations lead coordinator, engineering help center advisor, airport control engineer, and baggage service manager. Heathrow APOC roles and interactions are shown in Figure 2.

The concept of an APOC was developed in the context of the European Commission's Single European Sky (SES) initiative, which seeks to reform the fragmented European air traffic management with the aim of satisfying future needs of the European airspace in terms of capacity, safety, efficiency, and environmental impact (De Reyck et al. 2016). Improving airport performance is at the heart of the SES APOC proposal. By providing access to real-time data from various data sources of different APOC stakeholders, airports can manage their operations more effectively and efficiently.

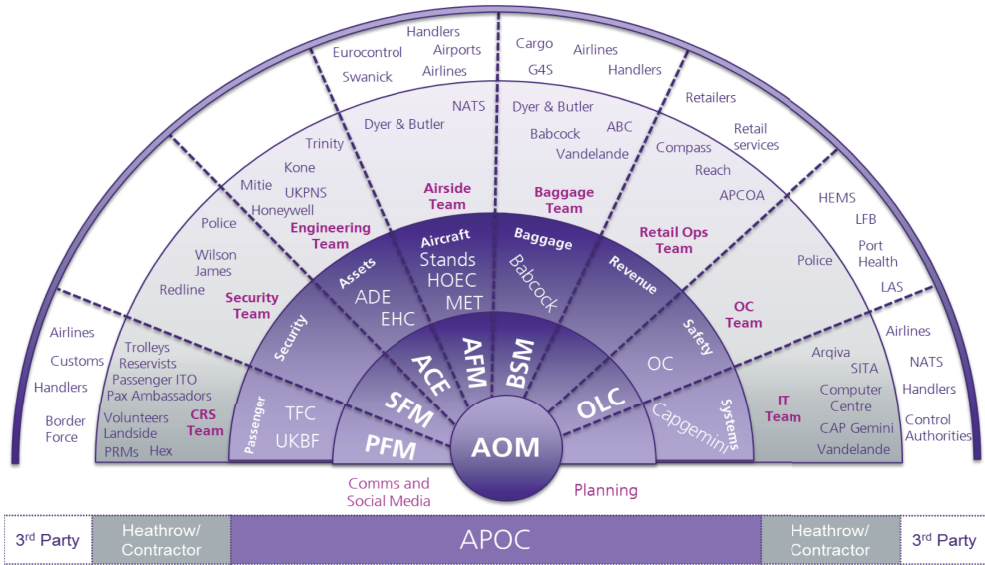
Although the colocation of stakeholders in an APOC has been moderately successful in achieving the collaborative decision making envisioned by the SES initiative, progress still needs to be made in terms of data sharing, the use of analytics, and genuine collaborative

Figure 1 Various Technologies Are Used to Give Stakeholders a View of All Operations at Heathrow Airport



Note. Information that the stakeholders can see includes traffic arriving at the airport, passengers checking in, security areas, immigration and border control areas, aircraft gates, and baggage operations.

Figure 2 Representatives of All Major Airport Stakeholders Are Present at the APOC, Thus Enabling Efficient Communications, Collaboration, and Joint Decision Making



decision making, to allow the stakeholders to evolve beyond their legacy systems, which were designed myopically for isolated stakeholder use. Therefore, the authors of this paper were commissioned by Eurocontrol to examine how Heathrow's APOC could be enhanced using real-time data sharing and advanced analytics.

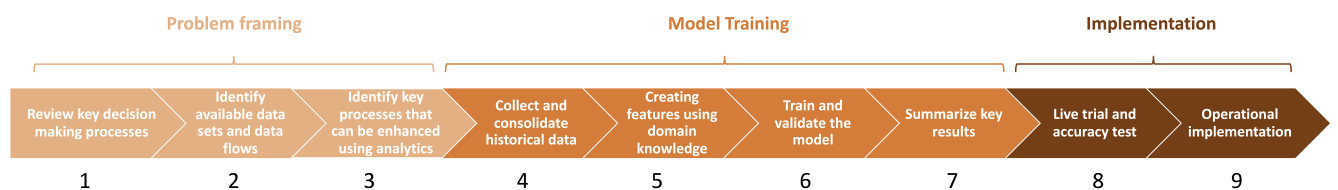
Developing and Implementing a Real-Time Data-Driven System

We reviewed the various APOC roles and responsibilities and identified the key APOC processes that could be enhanced by data-driven predictions and advanced analytics, including machine learning. We developed a case study that demonstrates how shared data and advanced analytics can improve passenger experience and flow through the airport. In particular, our work focused on improving the handling of transfer passengers, who make up a large proportion of Heathrow's passengers. Our goal was to identify ways to reduce waiting times in security and immigration areas, and minimize the number of passengers missing their connections or holding up aircraft departures. Throughout the project, we followed the nine-stage approach depicted in Figure 3.

Problem Framing: Steps 1-3

We began by reviewing the various APOC decision-making processes to identify potential processes that would be candidates for our study and would benefit from a decision support

Figure 3 We Used a Nine-stage Approach to Develop and Implement a Real-time Data-driven Collaborative Decision-making System for Improving the Connecting-passenger Flows Through Heathrow Airport



system based on real-time data and advanced analytics. In conjunction with our collaborators at Heathrow, we identified three such processes: (1) baggage flow management to reduce the number of misconnecting bags, (2) aircraft stand allocation, the allocation of a place where an aircraft can be parked, to reduce bottlenecks, and (3) passenger flow management to predict passengers' whereabouts throughout their time at Heathrow and thus improve the passenger experience and airline departure punctuality. After presenting these options to various Heathrow stakeholders, the third option was selected as the subject of the pilot study.

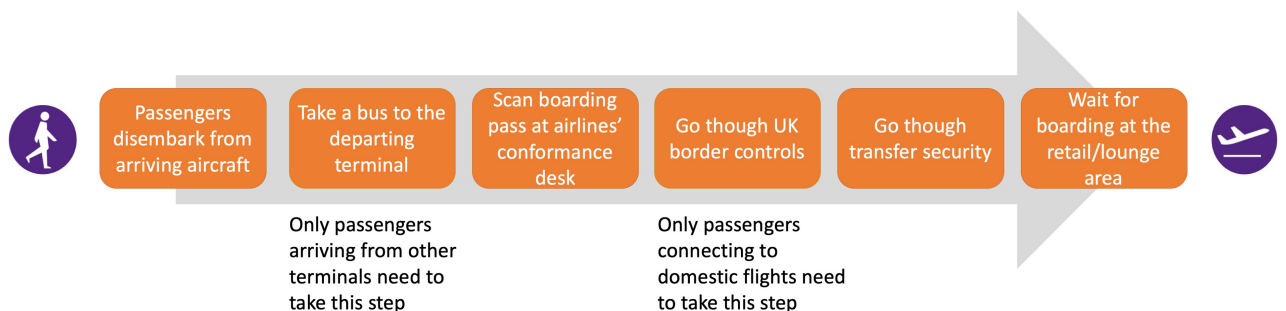
Passengers navigating an airport often encounter delays, especially at immigration and security. Although passenger arrivals at an airport can be predicted to some extent, there is significant uncertainty related to factors such as ground traffic delays, aircraft arrival delays, and volatility in the time passengers take to navigate through the airport. On the latter, we noted that once a passenger has arrived at the airport, whether as a departing, arriving, or connecting passenger, airports have little knowledge of that passenger's whereabouts in the airport. Improved passenger tracking in real time would enable airports to better serve their passengers, stabilize and predict departure times for outbound flights, and improve planning of resourcing needs.

Because Heathrow is a major international hub, transfer passengers make up a large proportion of its total passenger numbers, 32% or approximately 24 million, given annual traffic levels in 2015-2016 (De Reyck et al. 2016). Missed connections are the third leading reason for filing a complaint with an airline (Mott MacDonald 2016). Therefore, it is critical that the passengers' transfer journeys are optimized to ensure the airport can fulfil its mission to "give passengers the best airport service in the world" (Ferrovial 2016). Additionally, compared to departing or arriving passengers, connecting passengers have a more complex

journey, a greater interaction with Heathrow stakeholders, and a greater impact on the operations of the airport in terms of causing congestion and delayed departures. We choose the journeys of passengers arriving on international flights and connecting through Terminal 5 (approximately 3.7 million passengers annually) as the focus of this study. These passengers' outbound flights can be both domestic and international. Figure 4 offers an overview of this journey. As part of the problem-framing process, the project team participated in a physical walk through of this journey through Heathrow, experiencing the various possible routes that a connecting passenger can encounter.

Heathrow's APOC consists of a variety of airport teams, with the common goal of ensuring that passengers are happy and travel on time with their bags (De Reyck et al. 2016). The key decision makers in the context of connecting passengers are the managers of passenger flow, security flow, and aircraft flow, all of whom can influence the passenger journey. We interviewed the three managers to identify their roles in the handling of connecting passengers, their objectives, the decisions that they routinely make, and the information they have at hand to support their decisions.

Figure 4 The Transfer Journey for Passengers Arriving from International Flights Shows the Various Steps a Transfer Passenger Takes at Heathrow, Starting from Arrival, Being Bussed to the Terminal, Checking the Passenger's Boarding Pass at the Conformance Desk, Going Through Immigration and Security, and Departure



The passenger flow manager (PFM) monitors the flow of passengers through Heathrow's terminals. This manager works with terminal-based operations to deploy resources and optimize the passenger experience, with a main objective of reducing the number of passenger delay minutes while using resources efficiently. One delay minute is defined as one passenger waiting for one minute during his/her journey. Information used by the PFM includes: (1) passenger volume forecasts, generated the prior day; (2) CCTV feeds to monitor passenger flow (including automatic counting of passengers), and (3) queueing time estimates gathered via sampling. The PFM attends two conference calls and two meetings each day with other stakeholders in the APOC, where the anticipated impact of the passenger flows are communicated to relevant operational teams.

The security flow manager (SFM) manages resource planning in the security areas during the day through assessment of resource levels and forecasted passenger traffic. This process is complicated by unexpected peaks of arriving passengers, due to aircraft arriving early or late and passengers moving unexpectedly slowly or quickly through the airport, a journey that can take anywhere from 15 minutes to 2.5 hours. The main objective of the SFM is to minimize queue breaches, while using resources efficiently. A queue breach is defined as a passenger remaining in the queue longer than the service level for that area set by the UK Civil Aviation Authority; typically this is set at 15 minutes (De Reyck et al. 2016). Our interview with the SFM revealed that the main decisions the SFM makes include resourcing the security teams, based on predicted passenger traffic generated the day before. During the day, the SFM continually assesses the flow of the passengers and makes real-time decisions on the number of lanes are to be opened and whether staff should be redistributed between terminals, taking into account the inter-terminal travel time.

The aircraft flow manager (AFM) aims to enable a continuous stream of aircraft, taking into account global air traffic and weather conditions. The AFM can request that airlines

change their schedules based on airfield capacity and weather disruptions and can also amend the aircraft stand plan to assist airlines and handlers. Information used by the AFM includes local weather forecasts, air traffic conditions, and aircraft en route performance. The AFM interacts with the airlines, air traffic control, etc.

From our interviews and discussions with members of the APOC, it became apparent that the PFM, SFM, and AFM would all benefit from more timely and accurate predictions of passengers arriving at various points throughout the airport. Such a predictive system would: (1) enable the PFM to resource the immigration desks in advance of passenger surges, reducing the likelihood of queue breaches, (2) enable the SFM to ensure that security areas are properly prepared and resourced for incoming passengers, and (3) allow the AFM to warn airlines in advance that some passengers are unlikely to catch their connection, so that either the passengers could be offloaded or expedited, or the plane's scheduled departure time modified. This information, however, would need to be generated in real time, and made available to operations managers in time for them to be able to adjust their original plans to any newly generated forecast.

Following Steps 1-3, we identified the key target variable of our study as the time passengers take from the time their plane arrives at the gate to various points throughout the airport where unpredictable passenger arrivals can have significant consequences in terms of queues building up and aircraft being delayed. A point of significant interest at Heathrow's Terminal 5 is called the "conformance desk," which is both a checkpoint immediately before immigration and security and a station at which staff check whether a passenger has sufficient time to make it to the gate in time. If a passenger arrives at the conformance desk with less than a specific amount of time, typically 30 minutes, prior the departure time of his/her connecting flight, that passenger will be denied boarding and sent to a ticket desk

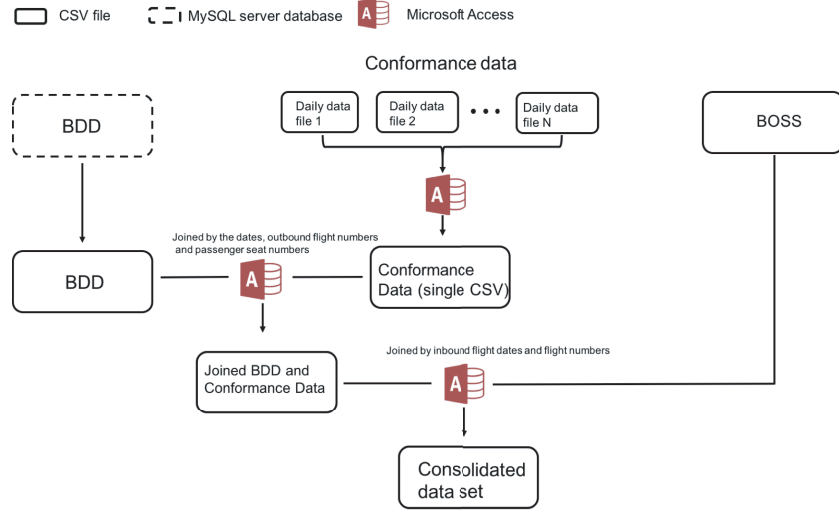
for rebooking onto a later flight. Therefore, the conformance desk is the critical point in a transfer journey, because it is the point at which a passenger either catches or misses a connecting flight. In addition, significant queues can build up at the conformance desk and the resulting delays can cause passengers to miss their flights. Therefore, we defined the passenger connection time, denoted by Δt , as the time difference between a passenger's arrival time at the airport and arrival time at the conformance desk.

Step 4: Collecting the Data

Variables containing the necessary information did not reside within a single database. Thus, we had to consolidate and combine data from various sources. We surveyed all available databases in the APOC. We identified three databases because they contain the necessary information for the prediction task at hand; the Business Objective Search System (BOSS) database, the Baggage Daily Download (BDD) database, and the conformance database. The BOSS database contains flight information, such as flight arrival times and aircraft capacity. The BDD database includes every piece of baggage of passengers connecting through Heathrow, as well as passenger information, including their inbound and outbound flight numbers. The conformance database, maintained by British Airways and shared daily with Heathrow, contains passengers' boarding pass scans and includes a time stamp showing the time at which each passenger arrived at the conformance desk. We collected data to capture an entire year to train and validate the model, picking up on any seasonal trends. Figure 5 shows how we consolidated the three databases.

We approximated arrival times at the conformance desk and at the airport by the "local conformance time" in the conformance database, and the "on-chock time" in the BDD data set, respectively. The "on-chock time" is a commonly used time stamp in aviation and is recorded when chocks are holding an aircraft in place, implying that the aircraft has landed and is parked at the gate.

Figure 5 We Set Up a Process to Consolidate and Merge the Existing Databases Carrying Information about Transferring Passengers



Approximately 1% of the data (just under 19,000 records) were found to be erroneous and were removed. Errors occurred because of rerouted passengers, whose information was not correctly modified in the system. We also discovered and removed records of passengers with a negative connection time, which is obviously impossible. We found that the main cause of negative Δt values was due to passengers changing seats after checking in, which resulted in mismatched records; for the same reason, we also removed excessively large connection times. The resulting data set contained information on 3,676,150 connecting passengers, with a mean connection time of 30.5 minutes (and a median of 27 minutes), and 30 variables. Among the 30 variables, only 11 of them were used as predictors in the model (Table 1). The others were eliminated because they were not available in real time, or they did not improve the accuracy of the model.

Step 5: Feature Design

Creating new variables, or feature engineering, is a key part of building an accurate machine learning predictive model (Domingos 2012). In this project, we worked with Heathrow experts to create several new variables based on the Heathrow team's knowledge of the aviation

domain and familiarity with the connecting passengers' journey. Table 2 shows an overview of the created new variables.

It is important to note that the process of engineering new variables, training the model, validating it (Steps 5-7) often requires several iterations, refining the model, and adding or removing variables to improve model performance, as we describe next.

Table 1 We Used 11 Existing Variables in Our Predictive Model to Help Explain the Variability in Passenger Connection Times Through the Airport

Variable	Rationale
Inbound/outbound flight capacity	Larger aircraft carry more passengers, making disembarking slower.
Inbound/outbound aircraft body type (wide/narrow)	The type of aircraft influences the time required for passengers to disembark, depending on passenger density and number of aisles.
Number of passengers on the inbound/outbound flight	Passengers arriving from a crowded aircraft may need more time to disembark.
Number of transfer passengers on the inbound/outbound flight	A larger number of passengers with connections can cause congestion; however, there is a smaller chance that a passenger will be lost if a group of people moves through the terminal together.
Passenger's inbound flight travel class (economy/business)	Business class travelers sit at the front and can disembark first. They may also move more quickly through the airport.
Passenger's arrival terminal	Passengers arriving from Terminals 2,3, and 4 need to take a bus transfer to Terminal 5, increasing the connection time.
Stand type (pier or remote)	Passengers arriving from an aircraft parked at a remote stand must take a shuttle bus to their arrival terminal.

Notes. The variables listed in the first four rows are for both inbound and outbound flights. We selected them based on real-time availability and their predictive power.

Table 2 We Used Eight Engineered Variables in Our Predictive Model to Help Explain the Variability in Passenger Connection Times Through the Airport

Feature	Insight
Inbound and outbound flight region (Europe, East Asia, North America, and rest of the world)	Flight region may be an indicator of a passenger's familiarity with connections through the airport.
Punctuality of inbound flight (time difference between on-chock time and scheduled arrival time)	Passengers arriving from late flights might be collectively affected by the lack of gate availability.
Hour of the day the arriving flight lands at the airport	Passengers who arrive during peak hours may need more time to make their connections because the airport is busy.
Perceived connection time (time difference between the arriving flight's on-chock time and the connecting flight's scheduled departure time)	Perceived connection time may be an indicator of the stress level of a passenger trying to make the connection, thereby impacting the speed with which that passenger move through the airport.

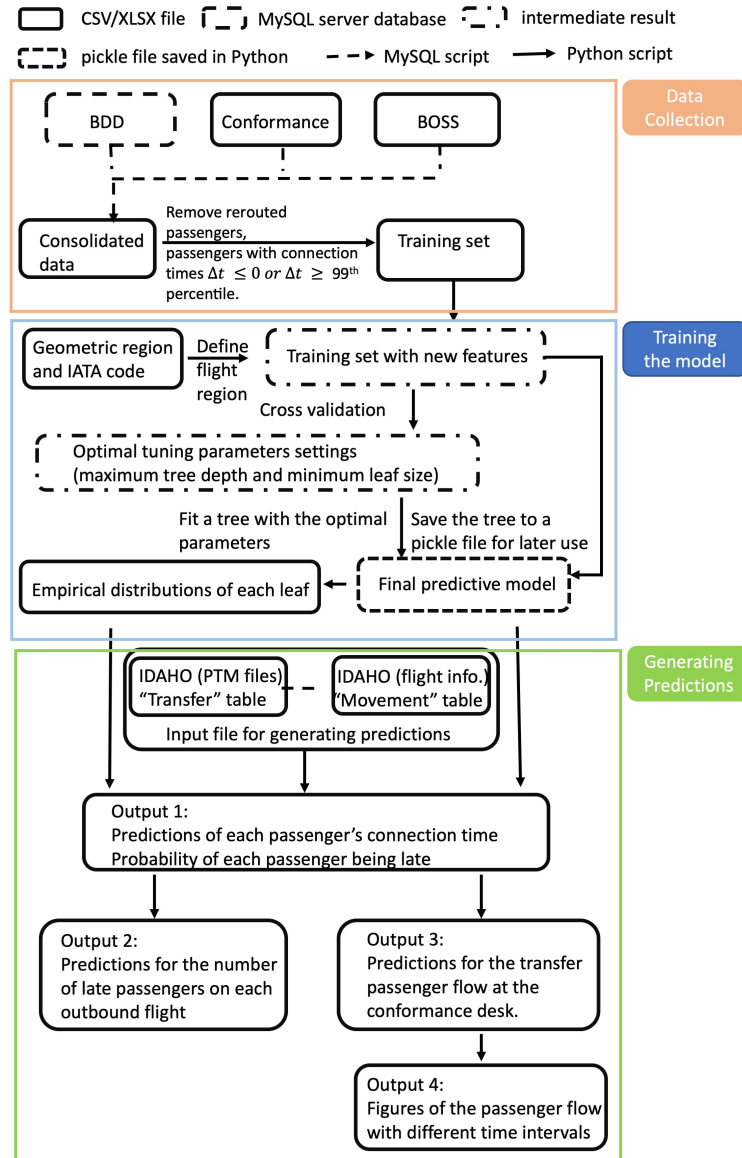
Notes. We created and experimented with a large number of engineered features (variables) that might help explain the variability in passengers' connection times through the airport. Only a subset of variables were retained for use in the model, based on the real-time availability of data and predictive power.

Step 6: Model Training and Validation

Next, we proceeded to build a model that would solve the two identified prediction problems: (1) predict the distribution of passengers' connection times Δt to identify potentially late passengers, and (2) predict the distribution of passenger flows, or number of passenger arrivals at the conformance desk, within specific time intervals. In Prediction Task 1, we also grouped late passengers based on their outbound flight, and predicted the number of passengers that will be late for each departing flight from the airport. Figure 6 gives an overview of the predictive system.

In the literature, these two prediction tasks are typically viewed as independent problems; classification models are used to identify late passengers (i.e., answer the question of whether a passenger will be late), and time series models are used to predict passenger flows (Wei and Chen 2012, Milenković et al. 2018). In our proposed system, however, we used the predictions

Figure 6 The Predictive System Consists of Three Parts: Data Collection, Training the Model, and Generating Predictions



Notes. The IDAHO system provides both flight-level and passenger-level information in real-time. Flight-level information includes a flight's actual and scheduled arrival time, aircraft body type, etc. Passenger-level information is extracted from the Passenger Transfer Message (PTM) files sent by the airline. The PTM files list passengers who are known to be transferring to another flight.

generated for Prediction Task 1 to calculate the number of arriving passengers (Prediction Task 2), an approach that yielded superior predictions than those generated using traditional benchmarks.

In order to avoid overfitting, we split the data consolidated in Steps 4 and 5 into an 80% training set and a 20% test set. We further split 80% of the training set into a smaller training set and 20% into a validation set. The smaller training set was used to train and tune the predictive model, with the validation set used later to obtain optimal distributional forecasts of the aggregate level predictions (i.e., the number of passengers arriving at the conformance desk).

Regression tree models are widely used for generating point forecasts (Breiman et al. 1984, James et al. 2013). Few studies, however, have applied these tree methods to produce distributional forecasts. We generalized the tree model to produce the entire distribution of the target variable. The regression tree partitions passengers into several leaves, or categories, and each category therefore contains a specific number of passengers' connection times. We approximated the distribution of the connection times within each leaf by applying kernel density estimation (Wand and Jones 1994). Given a new arriving passenger's information, the tree will first determine the leaf to which the passenger belongs (i.e., the customer segment, or category), and then utilize the distribution associated with this leaf to predict this passenger's connection time.

After experimentation, we decided to tune the tree to a maximum depth of 15 and the minimum leaf size of 700 using a 5-fold validation approach. The tuning parameters were selected by minimizing the pinball loss averaged over several quantiles and over five folds. We chose the pinball loss as a performance measure because it is a standard measure to test the accuracy of distributional forecasts (Jose and Winkler 2009, Grushka-Cockayne et al. 2017).

The formulation of a pinball loss function can be found in the appendix. In this project, we measured the 0.05, 0.25, 0.5, 0.75, and 0.95 quantiles that together roughly describe the entire distribution. For more detailed discussion on the quantile forecasts produced, see Guo et al. (2019). Our model uses the leaf distributions given by the regression tree to generate three sets of predictions as follows:

- *Quantiles of each passenger's connection time and the likelihood of being late.* Once the system receives a passenger's real-time information, the regression tree model predicts the category to which this passenger belongs. The quantiles of the distribution associated with this category are the predicted quantiles of this passenger's connection time. In our predictive system, passengers are considered to be late if they arrive at the conformance desk with less than 30 minutes before their connecting flights' scheduled departure time. The 30 minutes is often applied by airlines to identify late passengers at the conformance desk (De Reyck et al. 2016). Therefore, we then calculate a passenger's threshold connection time as his/her scheduled departure time, minus 30 minutes, and minus the passenger's arrival time at the airport. Finally, one minus the cumulative distribution function of this passenger's connection time evaluated at this threshold is used as the probability of the passenger being late.

- *Quantiles of the transfer-passenger flow.* The passenger flow, or the number of passengers arriving at the conformance desk during a specific time interval i , is obtained by running simulations using the distribution of passengers' connection times. Specifically, we simulate whether each passenger will arrive during time interval i . In the simulation, we assume the arrivals of passengers on the same inbound flight are correlated. To reduce the complexity of the predictive system, we also assume that each pair of the correlated passengers has the same correlation coefficient ρ . This ρ becomes a hyper parameter in the system and is selected

using the 5-fold cross-validation approach. Next, we count the number of passengers n_1 that are predicted to arrive during i . We repeat this process m times and the empirical distribution of the passenger flow is created using n_1, n_2, \dots, n_m . Finally, the quantiles of the number of passengers arriving at the conformance desk during time interval i can be calculated using the empirical distribution.

- *Number of late passengers for each outbound flight.* We group the passengers by their outbound flights, and calculate how many of them are likely to be late (i.e., their probability of being late is greater than 0.5). We also calculate the number of passengers that would be still at risk of being late if the airline delayed the departure time by 5, 10, 15, and 30 minutes.

The pseudo code of the model is shown in Algorithm 1 and Algorithm 2. In the next section, we report accuracy results of our predictive system compared to several leading benchmark machine learning methods.

Step 7: Key Results

We presented the most important features of the regression tree model to four groups of Heathrow stakeholders and received intuitive confirmation from these stakeholders. These features offer some insight into the connection times of passengers at the airport. Below we discuss the top five most important features and expand on why they are important in predicting passengers' connection times.

- *Whether a passenger arrives at Terminal 5.* Passengers arriving at Terminal 2, 3, or 4 need to take a shuttle bus to transfer to Terminal 5, and therefore they may need more time to make their connections.
- *Whether the passenger arrives from an EU country.* Passengers travelling from European countries might be more familiar with the connections through the airport, and the degree of familiarity may influence connection times.

- *Perceived connection time.* The perceived connection time indicates the stress level of a passenger trying to make a connection, thereby affecting the speed with which he/she moves through the airport.

- *Whether the passenger is traveling in Economy class.* Business class passengers sit in the front area on the aircraft, are more often single travelers, and often bring less hand luggage. As a result, they disembark faster and move more quickly through the airport.

- *Inbound hour.* If a passenger arrives during a busy hour, the airport might be crowded, and therefore that passenger may move more slowly through the airport.

To gauge the performance of the model, we compared the accuracy of our model with several leading benchmarks on the 20% test set. Table 3 shows the accuracy comparison between our approach and each of several benchmark models for Prediction Task 1 (predicting passengers' connection times) and Prediction Task 2 (predicting the number of passenger arrivals). In the table, “+” indicates the prediction generated from our approach is more accurate than the benchmark given in the column and under the accuracy measure given in the row, while “−” indicates the benchmark is more accurate. We applied the root mean squared error (RMSE) to test the accuracy of the point forecast, or the mean of a distribution. We also used pinball loss and hit rate, the percentage of realizations that fall within a central prediction interval, to measure quantile forecasts and prediction intervals. Based on the results shown in Table 3, the predictive system we describe in this paper performs better than the benchmarks in almost all cases. See Guo et al. (2019) for detailed accuracy results and descriptions of the benchmark models.

Figure 7 shows the 15-minute interval predictions of passenger flows on a randomly selected day from the test set. It is easy to observe the accuracy of the predictions generated by our approach compared to the best time series model, the Exponential Smoothing State Space

Table 3 A Comparison of the Accuracy of Our Approach on Predicting Connection Times and Number of Arrivals with Various Benchmarks Shows That Our Approach Outperforms Benchmark

Models for Almost All Accuracy Measures								
Prediction Task 1					Prediction Task 2			
Connection times					Number of passenger arrivals			
	LASSO				SARIMA			
	Linear	quantile	regression	boosting	with	LM	and	
	regression	regression	forest ^a	machine ^a	TBATS	covariates ^b	simulations ^c	
RMSE	+	+	+	+	+	+	+	
Q_{05}	+	+	+	+	+	+	+	
Q_{25}	+	+	+	+	+	+	+	
Pinball loss								
Q_{50}	+	+	+	+	+	+	+	
Q_{75}	+	+	+	+	+	+	+	
Q_{95}	+	+	+	—	+	+	+	
Avg.	+	+	+	+	+	+	+	
Hit rate (50%)	+	+	+	+	+	+	+	
Hit rate (90%)	+	+	+	+	+	+	—	

a. These two models are fit to a fraction of the data because they are extremely slow.

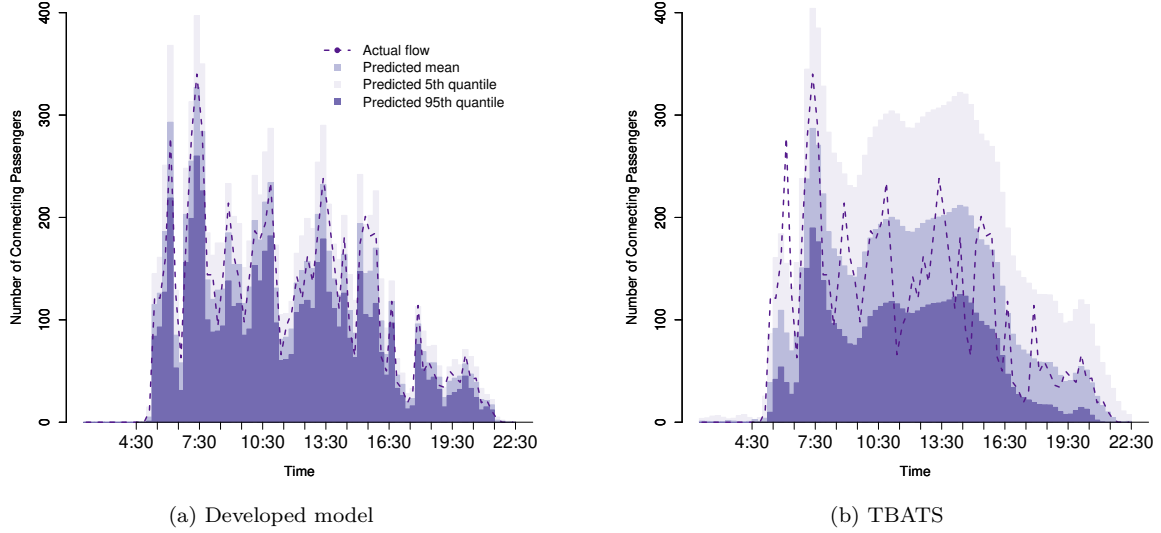
b. The covariates in the SARIMA model include the number of passengers arriving at the airport in the past 15 minutes, 30 minutes, and 45 minutes.

c. This model applies linear regression to predict the distribution of connection times. These individual distributions, together with simulations assuming correlated arrivals, are used to predict the number of passenger arrivals.

Model with Trigonometric Seasonality, Box-Cox Transformation, ARMA Errors, Trend And Seasonal Components (TBATS).

To test the impact of the new predictive system, we ran a simulation study to evaluate the optimal resourcing decisions made using the quantile forecasts of passenger flows generated from our predictive system compared to Heathrow's legacy system. The resource planning problem can be seen as a newsvendor problem, because both staffing and understaffing carry

Figure 7 Fifteen-Minute Interval Predictions of Passenger Flows on a Randomly Selected Day From the Test Set Show the Improved Accuracy of the Predictions Generated by Our Approach Compared to the Best Benchmark Model



costs at Heathrow. The optimal solution of the newsvendor-based staffing plan is a function of the quantile forecast of the number of passenger arrivals at the conformance desk. Which quantile to use depends on the ratio of staffing to understaffing cost. See Guo et al. (2019) for a more detailed discussion on the formulation. Based on the results of our backtesting study in which we evaluated our model across different quantiles, and compared it with Heathrow's legacy system, our model showed a 12% to 54% reduction in costs at the immigration and security areas.

Step 8 - 9: Implementation

Following several iterations of engineering additional features and training and testing the model, we conducted the first live trial of our model at Heathrow on July 19, 2016. During the trial, our model ran in parallel with existing systems for an entire day. Using real-time information that was sent from an aircraft flying into Heathrow, the model generated predictions and displayed the information to key decision makers in the APOC. The real-time

Figure 8 The Interface of the Connection Time Forecasting Application Allows Managers to Set the Forecasting Window, Number of Simulations, Granularity of the Passenger Flow, and the Update Frequency

The screenshot shows a window titled 'Heathrow_app'. Inside, there is a 'Parameters:' section with a table of input fields. The fields are: 'Forecast window (min)' with value '120', 'Number of simulations' with value '1500', 'Forecast resolutions (numbers separated with commas)' with value '1,5,15,60', 'Update frequency (min)' with value '5', 'Starting time (YYYY-MM-DD HH:MM:SS)' with value '2016-07-31 22:39:48', and 'Ending time (YYYY-MM-DD HH:MM:SS)' with value '2016-08-01 22:39:48'. Below the table is a 'START' button.

Parameters:	
Forecast window (min)	120
Number of simulations	1500
Forecast resolutions (numbers separated with commas)	1,5,15,60
Update frequency (min)	5
Starting time (YYYY-MM-DD HH:MM:SS)	2016-07-31 22:39:48
Ending time (YYYY-MM-DD HH:MM:SS)	2016-08-01 22:39:48

START

passenger information was sent via Passenger Transfer Messages (PTMs) from the IDAHO data system. These files are typically sent when a flight takes off from its origin, or sometimes when it is en route. The PTM also lists passengers who are known to be transferring to another flight.

We visually assessed our predictions by looking at the flight connections area on the screen in Terminal 5. There were no obviously large discrepancies between the predictions and the realizations assessed during the trial, and the trial was considered successful. The live trial was also useful for testing operational issues, such as whether the predictive algorithm was running fast enough and whether the data required by the model were readily available in real time when needed. We updated the predictions every five minutes and produced forecasts for the upcoming two hours. Predictions were generated on a rolling basis. Figure 8 shows the interface of the application we developed for the live trial.

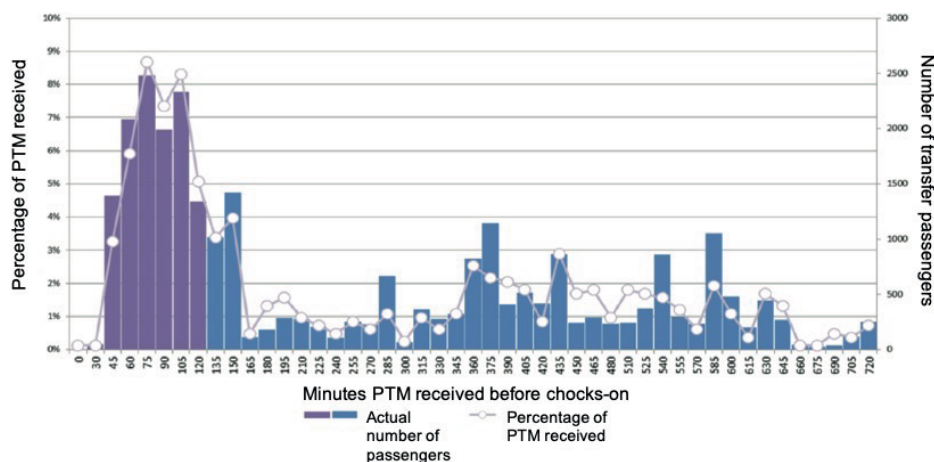
Note that the quality of the predictions depended on the quality of the available real-time data, which we obtained from the PTMs that the airlines provided. The timing of receiving these messages had a significant impact on the accuracy of our predictions. Figure 9 shows a histogram of the time between receiving the PTMs and the actual checks-on time of the

inbound flight on a particular day. The PTMs for 95% of the passengers were received more than 60 minutes ahead of their actual arrivals, while only 82% were received more than 90 minutes ahead. Thus, although the forecasting window was set to two hours during the live trial, 90 minutes turned out to be a better choice due to the availability of the data.

Following the live trial, the predictive system was adapted for stability prior to being fully implemented. An accuracy test of the expected number of passengers arriving at the conformance desk during peak hours (hourly passenger arrivals > 200) was conducted during July and August 2017. The RMSE of the predictions generated by our predictive system and by Heathrow's legacy system were 39.2 and 49.5, respectively.

Our predictive system has been fully implemented at APOC since 2017. The predictions generated from the system, as we show in Figure 10 in the appendix, are currently used to assist APOC managers in making real-time decisions. The PFM now identifies late passengers by looking at their probabilities of being late, as we show in Figure 10(a), and contacts ground staff to provide support to these passengers. The PFM also communicates with airlines and helps them adjust outbound flight departure times based on the predictions that

Figure 9 The PTM Timing Analysis on a Particular Day Shows 95% of Passengers' Information Was Received More Than 60 Minutes Ahead of Their Actual Arrivals, 82% More Than 90 Minutes Ahead, and 68% More Than 120 Minutes Ahead



we show in the example in Figure 10(b), which are the number of late passengers under the current and delayed departure times. The predictions of the passenger flow in the security and immigration area, which we show in the examples in Figures 10(c) and 10(d), are now applied by the SFM to adjust resourcing decisions in real time, for example, the number of lanes to be open and the number of personnel to move among terminals. These decisions are usually updated every 15 minutes at Heathrow.

Long-Term impact

Tom Garside, Head of Integrated Planning at Heathrow Airport, writes “The predictive system developed in this study has the potential to reduce misconnect rates (estimated saving to airline of \$100 per passenger), reduce headcount in transfer security through improved prediction of demand, and increase the reliability of the transfer product, which could enable passenger volume growth, benefitting both airport and airlines” (Garside 2019).

In addition to providing accurate forecasts, the system supports APOC decision-making processes in three ways. First, the system helps airport managers understand the key factors that influence passenger connection times. Second, the departure times for outbound flights can now be adjusted according to the predictions of the number of late passengers, facilitating greater predictability and stability of the entire European airspace. Third, the system allows the APOC to make better transfer security lane resourcing plans. Ultimately, this leads to increased passenger satisfaction because fewer passengers miss their connections and the delays resulting from having to offload their bags have decreased. In addition, our work has also highlighted the needs for a uniform data system among all stakeholders at APOC.

Robert Graham, the Head of Airport Research at Eurocontrol, writes “this study is ground breaking,” “has become a reference,” and “is used by a number of major European airports with a first deployment since 2017 in London Heathrow’s Airport Operations Centre.” He

adds that the study “demonstrates how decision making can be better informed by the flow of data and the use of predictive algorithms, and brings state-of-the-art thinking in machine learning, applied to a problem of crucial importance to airports around the world into the airport operations domain” (Graham 2018).

Our work has also attracted the attention of other airport operators around the world. Group ADP (formerly Aeroports de Paris) has contacted the Head of Airport Research at Eurocontrol, indicating an interest in the study, with a view to replicating it at Charles de Gaulle Airport, Orly Airport, and Le Bourget Airport. Fang Hao Foo, manager of Departure Experience at Changi airport in Singapore, wrote to the authors saying that “Being a major transfer hub, Changi also focuses on offering the best to our transfer passengers and would love to add strong data analytics capabilities to our arsenal. With this in mind, my team would like to seek your support to better understand what it would take to develop a similar predictive tool for Changi” (Foo 2018). In a similar vein, the authors are in discussion with LAX airport to help advise on designing a similar system. Finally, researchers at the Institute of Hygiene and Tropical Medicine in Lisbon, Portugal are using the model presented in this paper to create a model for predicting the spread of infectious diseases caused by people travelling by air and through airports, in order to develop an early warning surveillance system.

Acknowledgments

The authors are indebted to members of Heathrow’s APOC team, including Tom Garside, Neville Coss, and Frederick Tasker, for their close collaboration, tireless efforts, and dedication to seeing this project through to implementation. Special acknowledgment to Robert Graham from Eurocontrol for his vision and providing the opportunity for us to implement this project. Finally, the authors acknowledge the supporting work by colleagues at the

University of Virginia's Darden School of Business, Professor Casey Lichtendahl, Andrey Karasev, Missy Jenkins, and Nicole Puhl.

Appendix

The Pinball Loss Function

of the p -quantile (Q_p) of a distribution is given by

$$pl(Q_p, y) = \begin{cases} p(y - Q_p) & \text{for } Q_p \leq y \\ (1-p)(Q_p - y) & \text{for } Q_p > y \end{cases}$$

Where y is the realization.

Algorithm 1 - Train and Tune the Regression Tree Model to Predict Connection Times

Alg. 1 Train and tune the regression tree model.

Input:	x_{train} , Predictors in the training set
	t_{train} , Target variable (connection times) in the training set
	d_1, d_2, \dots, d_m , possible settings of the maximum depth
	k_1, k_2, \dots, k_n , possible settings of the minimum leaf size

Splitting the training set into five smaller sets x_1, x_2, x_3, x_4, x_5
 Splitting the target variable into five smaller vectors t_1, t_2, t_3, t_4, t_5

```

for  $i = 1$  to  $m$  do {for each setting of the maximum tree depth}
  for  $j = 1$  to  $n$  do {for each setting of the minimum node size}
    for  $s = 1$  to 5 do {for each smaller set  $x_s$  and  $t_s$ }
       $\delta \leftarrow \text{regression tree}(x_s, t_s, \text{maxdepth} = d_i, \text{minleaf} = k_j)$  {fit a regression tree}
       $x_{\text{valid}} \leftarrow x_{\text{train}} \setminus x_s; t_{\text{valid}} \leftarrow t_{\text{train}} \setminus t_s;$ 
       $l \leftarrow \delta(x_{\text{valid}})$ , {find the leaves that passengers in the validation set belong to}
       $\text{pred} \leftarrow \text{empirical distribution of connection times in } l$ 
      {pred are the predicted distribution of passengers' connection times in the validation set}
       $q_{05} \leftarrow \text{quantile}(\text{pred}, 0.05)$ 
       $q_{25} \leftarrow \text{quantile}(\text{pred}, 0.25)$ 
       $q_{50} \leftarrow \text{quantile}(\text{pred}, 0.50)$ 
       $q_{75} \leftarrow \text{quantile}(\text{pred}, 0.75)$ 
       $q_{95} \leftarrow \text{quantile}(\text{pred}, 0.95)$ 
      { $q_{05}, q_{25}, q_{50}, q_{75}$  and  $q_{95}$  are vectors that contain the quantile forecasts of passengers'
      connection times in the validation set}
       $\text{foldscore}_s \leftarrow \text{average}(\text{pinball loss}(q_{05}, t_{\text{valid}}), \text{pinball loss}(q_{25}, t_{\text{valid}}),$ 
         $\text{pinball loss}(q_{50}, t_{\text{valid}}), \text{pinball loss}(q_{75}, t_{\text{valid}}),$ 
         $\text{pinball loss}(q_{95}, t_{\text{valid}}))$ 
      { $\text{foldscore}_s$  is the average pinball loss across all five quantiles and all passengers in the
      validation set}
    end
     $\text{pinballscore}_{ij} \leftarrow \text{average}(\text{foldscore}, \text{across all } s \text{ with max depth } i \text{ and min leaf size } j)$ 
  end
end
 $d_{\text{opt}}, k_{\text{opt}} \leftarrow \text{which. min}(\text{pinballscore})$  {find optimal tuning parameters}
 $\delta \leftarrow \text{regression tree}(x_{\text{train}}, t_{\text{train}}, \text{maxdepth} = d_{\text{opt}}, \text{minleaf} = k_{\text{opt}})$ 
{fit the model using the optimal parameters}
Output: save the regression tree  $\delta$  to a pickle model file.

```


Algorithm 2 - Predict Quantiles of Passengers' Connection Times, Probabilities of Being Late, and Quantiles of Passenger Flows (Number of passengers arriving at the conformance desk in each time interval)

This algorithm is used first to determine the correlation coefficient ρ . To find the optimal value of ρ , we run the following algorithm in each of the five validation folds under different settings of ρ , and calculate the pinball loss of the quantile forecasts of the passenger flows. The setting ρ^* that gives the minimum pinball loss averaged over the five quantiles and five folds is determined as the optimal value. Finally, we run the following algorithm under ρ^* to generate predictions in the test set.

Alg. 2 Predict quantiles of passengers' connection times, probabilities of passengers being late, quantiles of the number of passengers arriving at the conformance desk from time h to time $h + k$ with every r minutes.

Input: h , Current time; k , Forecasting window; r , Resolution (length of time intervals)
 n , Number of passengers who arrived at the airport in the last 2.5 hours or will arrive in the next k minutes
 m , Number of simulations
 b , Number of inbound flights; w , Number of different outbound flights
 δ , The regression tree model
 x_1, x_2, \dots, x_n , Input of passengers' information
 t_1, \dots, t_n , On-chock time of each passenger's inbound flight
 T_1, \dots, T_w , Scheduled departure time of each outbound flight
 ρ , correlation coefficient of each pair of passengers' connection times

for $i = 1$ to n **do** {for each passenger}
 $l_i \leftarrow \delta(x_i)$ { l_i is the terminal leaf that passenger i belongs to}
 $f_i \leftarrow$ empirical distribution of leaf l_i
 $proplate_i \leftarrow 1 - CDF_{f_i}(T_i - 30\text{min} - t_i)$ { CDF_{f_i} is the CDF of the empirical distribution f_i }
end

for $i = 1$ to k/r **do** {find all the starting and ending points of the r minutes time interval}
 $hs_i \leftarrow h + (i - 1) * r$; $he_i \leftarrow h + i * r$
end

for $i = 1$ to m **do** {for each simulation}
for $j = 1$ to k/r **do** {for each time interval}
for $c = 1$ to b **do** {for each inbound flight}
 $x_{n_c} \leftarrow$ passengers arriving on the c flight {suppose we get a passengers}
 $R \leftarrow$ an $a \times a$ matrix with diagonal elements equal to 1 and off-diagonal equal to ρ
 $v \leftarrow$ sample a vector from a Gaussian copula with coefficient matrix R
 $u \leftarrow$ Probability Integral Transform (v) {get a uniformly distributed vector}
for $e = 1$ to a **do** {for each of the a passengers arriving from this flight}
 $p_e \leftarrow CDF_{f_e}(he_j - t_e) - CDF_{f_e}(hs_j - t_e)$
{calculate the probability of arriving in this time interval at the conformance desk}
 $S_{ce} \leftarrow$ quantile function of Bernoulli distribution(quantile = u_e , probability = p_e)
end
end
 $count_{ij} \leftarrow \text{sum}(S, \text{over all inbound flights } c \text{ and passengers } e \text{ on each flight } c)$
{ $count_{ij}$ is the number of passengers arriving within time interval j for simulation i }
for $j = 1$ to w **do** {for each outbound flight}
for d in $[0, 5, 10, 15, 30]$ **min do**
for $e = 1$ to g **do** {for each of the g passengers connecting to this flight}
 $p_{de} \leftarrow CDF_{f_e}(T_j + d - 30\text{min} - t_e)$
end
 $z_{ijd} \leftarrow$ count if $p_d > 0.5$ across all e
end
end
end

Output: mean and quantiles of each f_1, \dots, f_n {mean and quantiles of connection times}
 $proplate$ {probability of each passenger being late}
 $z_{1,0}, \dots, z_{w,30}$ {expected number of late passengers for each outbound flight averaged across all simulations}
mean and quantiles of each $count_1, \dots, count_{k/r}$ {mean and quantiles of the simulated number of passengers arriving at the conformance desk in each time interval}

Figure 10 The Output Files Generated by Running the Application Include a Table of Forecasts for Individual Passenger Connection Times, a Table of the Expected Number of Late Transfer Passengers for Each Outbound Flight, and Several Tables and Figures that Show the Predicted Passenger Flows with Different Granularities

A	B	C	D	E	F	G	H	I	J	K
passenger_id	on_chock_time	q0.05	q0.25	median	q0.75	q0.95	ib_flight_no	ob_flight_no	P(missing connecting flight)	
1	323698	01/07/2016 12:46	01/07/2016 13:09	01/07/2016 13:16	01/07/2016 13:22	01/07/2016 13:29	01/07/2016 13:46	BA293	0	
2	323723	01/07/2016 12:19	01/07/2016 12:43	01/07/2016 12:51	01/07/2016 12:56	01/07/2016 13:04	01/07/2016 13:19	BA069	0	
3	324028	01/07/2016 11:42	01/07/2016 11:52	01/07/2016 11:57	01/07/2016 12:01	01/07/2016 12:06	01/07/2016 12:14	BA115	0	
4	324213	01/07/2016 13:23	01/07/2016 13:34	01/07/2016 13:38	01/07/2016 13:41	01/07/2016 13:46	01/07/2016 13:57	BA279	0.11	
5	323846	01/07/2016 11:45	01/07/2016 11:56	01/07/2016 12:00	01/07/2016 12:04	01/07/2016 12:09	01/07/2016 12:18	BA287	0.01	
6	323561	01/07/2016 12:35	01/07/2016 12:58	01/07/2016 13:05	01/07/2016 13:11	01/07/2016 13:18	01/07/2016 13:31	BA269	0	
7	323767	01/07/2016 12:06	01/07/2016 12:18	01/07/2016 12:21	01/07/2016 12:25	01/07/2016 12:29	01/07/2016 12:38	BA1394	0.80	
8	323759	01/07/2016 12:28	01/07/2016 12:39	01/07/2016 12:43	01/07/2016 12:47	01/07/2016 12:51	01/07/2016 13:01	BA049	0	
9	323493	01/07/2016 11:31	01/07/2016 11:41	01/07/2016 11:46	01/07/2016 11:49	01/07/2016 11:54	01/07/2016 12:03	BA1484	0.96	

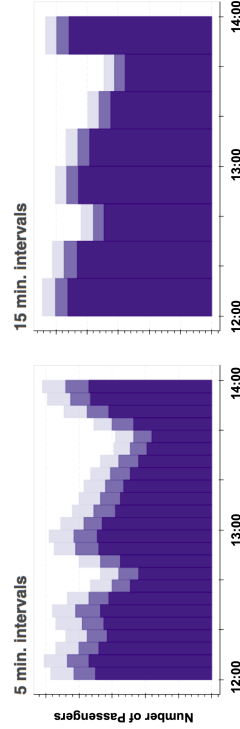
(a) Connection time for each passenger. Column A and B show passenger IDs and on-chock times. Column C - G are the five quantiles of passengers' connection times. Column H and I show passengers' inbound and outbound flight number. Column J shows the probability of a passenger being late for her connecting flight.

A	B	C	D	E	F
ob_flight_no	current SDT	+5 min	+10 min	+20 min	+30 min
1	BA005	3	2	2	1
2	BA007	22	20	17	11
3	BA009	1	1	0	0
4	BA011	0	0	0	0
5	BA017	2	2	1	1

(b) Expected number of late passengers for each outbound flight. Column A shows flight number. Column B, C, D, and E present the expected number of late passengers with current scheduled departure time, current scheduled departure time + 5min, +10 min, and +15 min.

A	B	C	D	E	F	G
inter_begin	inter_end	q0.05	q0.25	median	q0.75	q0.95
1	01/07/2016 12:00	01/07/2016 12:15	233	244	253	260
2	01/07/2016 12:15	01/07/2016 12:30	217	227	235	244
3	01/07/2016 12:30	01/07/2016 12:45	172	183	190	197
4	01/07/2016 12:45	01/07/2016 13:00	214	225	233	240
5	01/07/2016 13:00	01/07/2016 13:15	198	207	216	224
6	01/07/2016 13:15	01/07/2016 13:30	163	173	179	189
7	01/07/2016 13:30					199

(c) Plots of the number of passengers arriving at the conformance desk during every 5 minutes and 15 minutes time interval. Both the mean and 90% prediction interval are shown in the figure.



(d) Transfer passenger flows at the conformance desk. Column A and B represent the beginning and ending time of the 15 minutes time interval. Column C, D, E, F, and G show the five quantiles of the number of passengers arriving at the conformance desk during every 15 minutes time interval.

References

- Breiman L, Friedman J, Stone C, Olshen R (1984) *Classification and Regression Trees* (Chapman and Hall/CRC, Boca Ratón, FL).
- De Reyck B, Guo X, Grushka-Cockayne Y, Lichtendahl Jr K, Karasev A, Garside T, Coss N, Tasker F (2016) Apoc business process reengineering big data study. Technical report, Eurocontrol, Brussels.
- Domingos P (2012) A few useful things to know about machine learning. *Comm. ACM* 55(10):78–87.
- Ferrovial (2016) Heathrow recognized for excellent passenger service by Skytrax. Accessed April 2, 2020, <https://newsroom.ferrovial.com/en/news/heathrow-recognized-for-excellent-passenger-service-by-skytrax/>.
- Foo FH (2018) Email to Bert De Reyck, Xiaojia Guo, and Yael Grushka-Cockayne, November 21, 2018.
- Garside T (2019) Email to Bert De Reyck, Xiaojia Guo, and Yael Grushka-Cockayne, March 8, 2019.
- Graham R (2018) Letter of endorsement, August 24, 2018.
- Grushka-Cockayne Y, Lichtendahl Jr KC, Jose VRR, Winkler RL (2017) Quantile evaluation, sensitivity to bracketing, and sharing business payoffs. *Oper. Res.* 65(3):712–728.
- Guo X, Grushka-Cockayne Y, De Reyck B (2019) Forecasting airport transfer passenger flow using real-time data and machine learning. Accessed April 2, 2020, https://www.hbs.edu/faculty/Publication%20Files/19-040_89360426-c7a9-4aac-95e1-a3a3db276dc8.pdf.
- Heathrow Airport (2018) Company facts and figures. Accessed April 2, 2020, <https://www.heathrow.com/company/company-news-and-information/company-information/facts-and-figures>.
- James G, Witten D, Hastie T, Tibshirani R (2013) *An Introduction to Statistical Learning* (Springer, New York).
- Jose V, Winkler R (2009) Evaluating quantile assessments. *Oper. Res.* 57(5):1287–1297.
- Milenković M, Švadlenka L, Melichar V, Bojović N, Avramović Z (2018) Sarima modelling approach for railway passenger flow forecasting. *Transport* 33(5):1113–1120.
- Mott MacDonald (2016) Annual analyses of the EU air transport market. Technical report, European Commission, Brussels.

Wand M, Jones M (1994) *Kernel Smoothing* (Chapman and Hall/CRC, Boca Ratón, FL).

Wei Y, Chen M (2012) Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. *Transportation Res. Part C: Emerging Tech.* 21(1):148–162.