Forecasting airport transfer passenger flow using realtime data and machine learning

Xiaojia GUO
University College London

Yael GRUSHKA-COCKAYNE
Harvard University

Bert DE REYCK
Singapore Management University, bdereyck@smu.edu.sg

Forecasting Airport Transfer Passenger Flow Using Real-Time Data and Machine Learning

Xiaojia Guo
UCL School of Management, University College London, London, UK, x.guo.11@ucl.ac.uk

Yael Grushka-Cockayne
Harvard Business School, Cambridge, MA 02163, ygrushkacockayne@hbs.edu
and Darden School of Business, University of Virginia, Charlottesville, VA 22903

Bert De Reyck
UCL School of Management, University College London, London, UK, bdereyck@ucl.ac.uk

Problem definition: In collaboration with Heathrow airport, we develop a predictive system that generates quantile forecasts of transfer passengers' connection times. Sampling from the distribution of individual passengers' connection times, the system also produces quantile forecasts for the number of passengers arriving at the immigration and security areas.

Academic/Practical relevance: Airports and airlines have been challenged to improve decision-making by producing accurate forecasts in real time. Our work is the first to apply machine learning for predicting real-time quantile forecasts in the airport. We focus on passengers' connecting journeys, which have only been studied by few researchers. Better forecasts of these journeys can help optimize passenger experience and improve airport resource deployment.

Methodology: The predictive model developed is based on a regression tree combined with copula-based simulations. We generalize the tree method to predict complete distributions, moving beyond point forecasts. To derive insights from the tree, we introduce the concept of a *stable tree* that can be summarized by its key variables' splits.

Results: We identify seven key factors that impact passengers' connection times, dividing passengers into 16 passenger segments. We find that adding correlations among the connection times of passengers arriving on the same flight can improve the forecasts of arrivals at the immigration and security areas. When compared to several benchmarks, our model is shown to be more accurate in both point forecasting and quantile forecasting.

Managerial implications: Our predictive system can produce accurate forecasts, frequently, and in real-time. With these forecasts, an airport's operating team can make data-driven decisions, identify late connecting passengers and assist them to make their connections. The airport can also update its resourcing plans based on the prediction of passenger arrivals. Our approach can be generalized to other domains, such as rail or hospital passenger flow.

Key words: quantile forecasts; regression tree; copula; passenger flow management; data-driven operations.

1. Introduction

Passengers arriving at an airport often experience delays, especially at immigration and security. These delays are caused in large part by the volatility and uncertainty in arrival patterns of passengers at immigration desks and security lanes. Airports and airlines have long invested in optimizing and controlling aircrafts' arrivals and departures (Barnhart and Cohn 2004, Lohatepanont and Barnhart 2004, Lan et al. 2006, Atkinson et al. 2016). Once passengers have disembarked, however, airports have little knowledge of passengers' whereabouts in the airport. Improved passenger tracking in real time, in particularly transfer passengers, would enable airports to better serve their passengers, stabilize and predict departure times, and plan resourcing needs.

Flights landing or departing from international hubs often carry a high proportion of connecting passengers. For example, the vast majority of the flights traveling through Heathrow, the busiest airport in Europe with more than 75 million passengers each year (Heathrow 2016), have at least 25% connecting passengers. Missed connections are the third leading reason for filing a complaint with an airline (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).

We take a passenger-centric approach to studying transfer passenger flows in the airport. Motivated by our collaboration with Heathrow, we develop a system that predicts the distributions of the connection times of international arriving passengers. With these distributions, airport decision makers can identify passengers with high chance of missing their outbound flights, and take proactive action to reduce missed connections. Here, the connection time is defined as the time difference between a passenger's arrival at the airport, i.e. when their plane arrives at its gate, and their arrival at airlines' conformance desk, which is the last check point before passengers progress to immigration and security. The system also samples from the distribution of each passenger's connection time to predict the passenger flows, or the number of arrivals at the conference desk. These forecasts are updated every 15 minutes based on real-time data. In addition to the average number of arrivals, we also predict quantiles to enable robust resource planning for peak and trough passenger flows. The system is currently in use at Heathrow airport, informing the resourcing of immigration desks and security lanes, informing airlines of late passengers, and facilitating improved collaborative decision-making.

Passenger delays and passenger flow management have already received some attention in the literature. Wei and Hansen (2006) build an aggregate demand model for air passenger traffic in a hub-and-spoke network, and find that airlines can attract more connecting passengers by increasing service frequency. Barnhart et al. (2014) develop a methodology for modeling delays of the domestic passengers in the United States, and study the key factors that affect the performance of the National Air Transportation System. In their paper, they mention that lack of passenger travel data has made it difficult for researchers to explore passenger centric problems. To the best of our knowledge, no studies relating to passengers' transfer journeys exist that are based on actual data of passengers' travel information.

Our paper is the first to study passengers' transfer journeys using airport data, and provide decision support in real time. Real-time data has been used for decision-making in other contexts (Bertsimas and Patterson 2000, Mukherjee and Hansen 2009). Most of these studies, however, focus on air traffic problems. Additionally, the integration of machine learning, big data, and real-time decision making has received only limited attention in the literature (Shang et al. 2017). Our study is also the first of its kind to exploit large data sets of flight and passenger information using customized machine-learning algorithms.

A predictive system that can be implemented at the airport must meet three requirements. First, the approach must produce both accurate point forecasts and well-calibrated quantile forecasts. Second, the operational plan needs to be updated frequently; and thus, any predictive system developed must be capable of generating forecasts rapidly. Third, the model needs to be intuitive in the sense that it enables airport managers to understand the key factors that influence passengers' connection times. This third requirement was essential in order to secure the buy-in of Heathrow teams for scaling up implementation.

The predictive model is based on the regression tree method (Breiman et al. 1984). The interpretation of the results summarized in a regression tree is intuitive compared to other advanced machine learning methods (Friedman et al. 2001). Although the regression tree method has been widely used to make point forecasts in business (Eliashberg et al. 2007, Ferreira et al. 2015, Xue et al. 2015), few have applied it to make prediction intervals. In this study, we first train and tune a regression tree using Heathrow's historical data. Next we fit a Gamma distribution to each leaf of the regression tree in order to generate quantiles of the passengers' connection time. Other machine learning algorithms, such as quantile regression forests (Meinshausen 2006), can also generate prediction intervals. These methods, however, are often time-consuming and are typically treated as a black box.

Given real-time information, we categorize passengers based on the regression tree, and compute the corresponding distribution of their connection times. We also simulate connection times from each of the passengers' distributions, and calculate the number of passengers arriving at immigration and security areas within 15 minute windows. The connection times of passengers travelling on the same flight are assumed to be correlated. Such correlations are incorporated using Gaussian copulas. We find that the prediction intervals become wider as dependencies increase, and the copula-based simulations make our forecasts more accurate.

Several results are derived from the regression tree model. First, we report key findings on significant factors that affect passengers' connection times, based on a reduced version of the tree. Second, we compare the performance of our regression tree model in forecasting individual connection times against a naïve model, and four other methods that are widely used in the machine learning community: linear regression, quantile regression, quantile regression forest and gradient boosting machine with a quantile regression objective. Finally, we compare the performance of our regression tree model in predicting number of arrivals at immigration and security areas against a naïve model and the linear regression. Given the simplicity of the regression tree, we anticipated trading off accuracy for improved interpretability and run time. Surprisingly, the single regression tree performs favorably in both point forecasting and quantile forecasting.

Our work offers the following contributions. First, to our knowledge, we are the first to develop a predictive system using real-time data to forecast passenger movements at an airport. Second, we make use of a popular machine learning algorithm to produce distributional forecasts instead of point estimates. In order to derive insights from the model, we define a *stable tree* such that the tree is easy to interpret, and the findings from the tree do not vary much when retraining the model. Third, our system provides individual and aggregate level predictions simultaneously. Finally, we identify key factors that influence passengers' connection times. When combined, these contributions assist practitioners and academics looking to use machine learning to solve practical problems.

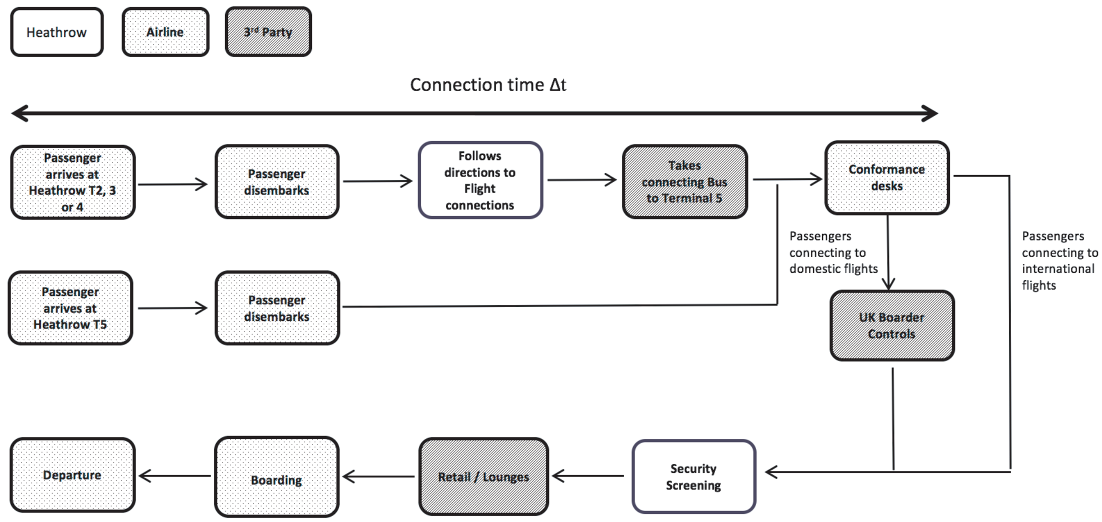
The system was first tested in a live trial at Heathrow in July 2016. Following this, it was adapted for stability, prior to being fully implemented in 2017. Accuracy tests over July and August 2017 suggested that the average accuracy of the model outperformed Heathrow's existing methods. Following the positive feedback and results so far, Heathrow is planning to extend the system to enhance other airport processes. We are currently in discussion with Aéroports de Paris, who manage the three biggest international airports in Paris, to implement a similar system.

2. Problem Description

In 2015, a third of Heathrow's passenger traffic, or approximately 24 million arriving passengers, were connecting passengers, moving through Heathrow to other destinations. When compared with transfer passengers arriving on domestic flights, those arriving on international flights have more complex journeys and an increased level of interaction with Heathrow stakeholders. For instance, international arrival passengers need to go through security at Heathrow, while domestic arrival passengers can go directly to their boarding gate after landing. The transfer journey of passengers arriving on international flights is the focus in this study.

The journey for Heathrow connecting passengers is not dissimilar to other international hubs. Typically, once an international flight lands at Heathrow and the passengers disembark, connecting passengers follow signs for flight connections. Passengers arriving at Heathrow Terminals 2, 3, or 4 transfer to Terminal 5 using a shuttle bus. Upon arrival at Terminal 5, passengers check in at the

Figure 1 The Connecting Journey of International Arriving Passengers Departing Through Terminal 5



airport conformance desks. Staff at the conformance desks check the boarding pass to ensure that the passengers are in the right place with enough time to catch their onward flight. If a passenger at the conformance desk is unlikely to reach their outbound flight, they are redirected to a ticket desk for assistance. This transfer journey is depicted in Figure 1.

Passengers connecting to domestic destinations enrol at UK Border Control. After that point, passengers progress to security screening. Enrolment at UK Border Control is not required for passengers connecting to other international destinations. After passing through the conformance desk, these passengers progress directly to the security screening. Finally, after passing airport security, passengers enter the departure lounge, and proceed toward their boarding gates.

Heathrow's Passenger Flow Manager, together with terminal-based operations, is in charge of resource deployment and optimizing passengers' experience. In cases where the Passenger Flow Manager realizes that there is likely to be passenger congestion, she informs relevant teams, such as the Border Force and airlines, and recommends actions to resolve the issue. Heathrow's Security Flow Manager generates initial resource plans prior to the day of operation. During the day, the Security Flow Manager continually assesses the flow of passengers, and identifies any deviations from the plan. The Security Flow Manager then makes decisions, in real time, on how many lanes are required to be open and whether staff could be better distributed between terminals (De Reyck et al. 2016).

Both the Passenger Flow Manager and the Security Flow Manager operate from Heathrow's Airport Operation Centre (APOC), which consolidates all airport operations including gate management, security, baggage, passenger processes and crisis management in a centre in order to improve data sharing and collaborations across teams (Eurocontrol 2010). The APOC provides different stakeholders access to the historical and real-time data from various data sources. However, the data sets available to each stakeholder group are often in different formats and stored in different systems, making them difficult to cleanse and consolidate.

Prior to the current study, the decision-making process in the center was primarily manual. Decisions relating to the transfer passengers were not all made with complete information, preventing proactive and truly collaborative decision-making. Without individual level forecasts of connection times, the Passenger Flow Manager could not precisely identify passengers with high chances of missing their outbound flight and assist them to make their connections. With accurate real-time individual predictions, early off-loading and re-booking can be facilitated. On an aggregate level, real-time information would enable the APOC to alert the airlines of the expected number of late passengers. In addition, lack of real-time predictions of passenger flows into immigration and security areas limits the Security Flow Manager's ability to adjust necessary resources. The Security Flow Manager would only monitor passengers through CCTV cameras. It was often too late to reallocate staffs by the time the manager realized there was potential congestion.

The focus of our study is on forecasting connection times between passengers' arrivals at the airport and their arrivals at Terminal 5's conformance desks. These forecasts of the arrival time at the conformance desks will allow the airlines to predict which passengers are at risk of missing their outbound flights. Given these predictions, the airlines, with the support of the Passenger Flow Manager, will be able to expedite late passengers. The forecasts of the number of passengers arriving at the conformance desks will also support making decisions on allocating security resources. Moreover, with the predictive model, point and quantile forecasts can be generated in real time, allowing both managers to plan for peak and trough passenger flows.

In this study, we utilize data on 3.7 million passengers arriving from international flights and departing from Terminal 5 to train the model. While our study focuses on Terminal 5 departures, the system has been extended to include other terminals. There are two major constraints to consider: the availability of real-time data and the time needed to adjust plans. For example, it usually takes a Security Flow Manager more than half an hour to move staff between terminals. To guarantee that a manager has enough time to make adjustments, the forecasts should be provided as soon as new data comes in. This requires an effective procedure that can produce accurate forecasts ahead of time.

3. The Predictive Model

In this section, we begin by describing the data used to train and test the predictive model. We follow with details about the model itself and the process of selecting tuning parameters. Insights from the model and accuracies of its predictions are presented in Section 4. All the code of this study is available from the authors.

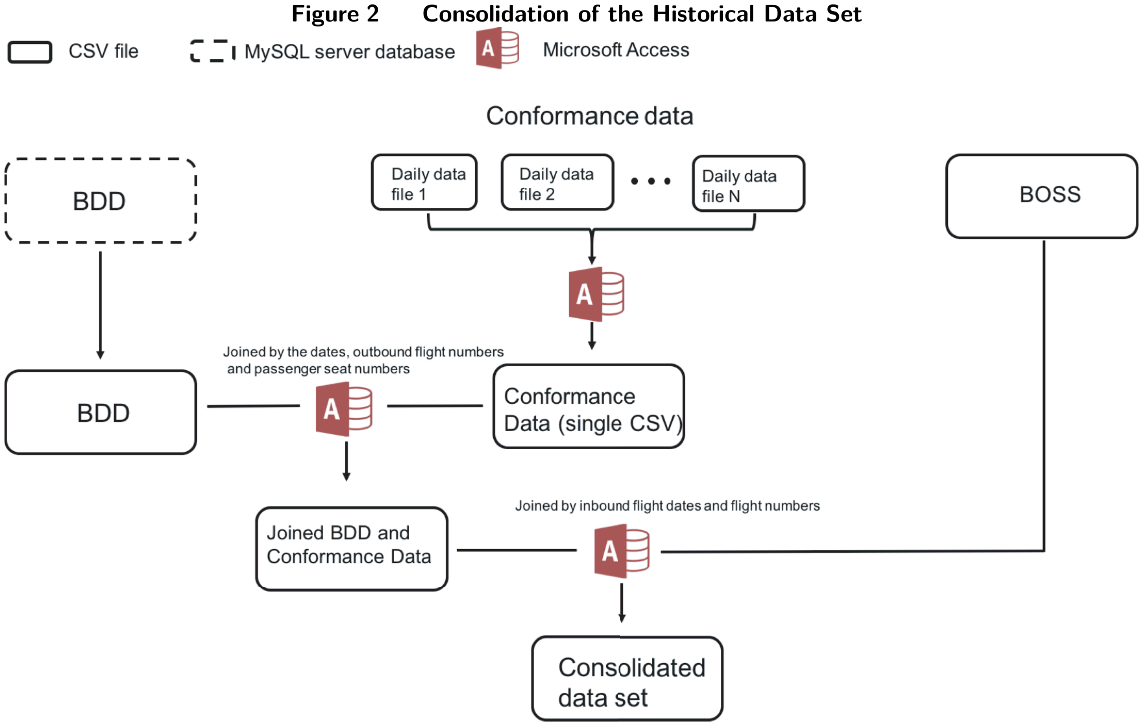
3.1. Data Processing

The data sources available at Heathrow can be classified into two categories: flight level data and passenger level data. The flight level data provides detailed information on departures, arrivals, and aircraft features (e.g., aircraft body type) for individual flights. The passenger level data records each passenger's travel information, such as their travel class.

Historical data for all of 2015 was used to train and test the model. This data was collected from three data sets: the Business Objective Search System (BOSS), the Baggage Daily Download (BDD), and the Conformance data sets. The BOSS data contains all flight information data. The BDD data records every piece of connecting baggage through Heathrow on the previous day. Each of the records also contains passenger information, such as passengers' arriving and connecting flight number. The Conformance data, provided on a daily basis, stores the records of passengers' boarding pass scans at the conformance desks.

The data sets were consolidated in Microsoft Access. As shown in Figure 2, the BDD data was exported from a MySQL server, and the Conformance data was merged from daily CSVs into a single CSV. Next the BDD and the Conformance data sets were joined, resulting in a new table containing all rows from both data sets as long as there was a match between the dates, outbound flight numbers, and passenger seat numbers. If there was no match, the data was discarded. Finally, flight-level information was added to the new table by mapping inbound flight dates and flight

Figure 2 **Consolidation of the Historical Data Set**



numbers with those in BOSS. The resulting data file represented a total of 3,762,690 records, each record representing a transfer passenger, over the course of 2015.

The target variable of our study — passengers' connection times Δt — was calculated as the time difference between the arrival time at the airport and the arrival time at the immigration and security areas. Here, the arrival time at the airport and at the immigration and security areas are approximated by the "on-chock time" in the BDD data set, and the "local conformance time" in the Conformance data set, respectively. The "on-chock time" is recorded when the chocks of an aircraft are in place, meaning the aircraft has landed and is parked at the gate. The "local conformance time" is recorded when the boarding pass is scanned at a conformance desk, and the passenger progresses to the immigration and security areas. Although Δt is the target variable in our predictive system, in practice we report arrival times at immigration and security for managers at APOC to use in their daily decision-making.

The data was cleansed for rerouted passengers, negative connection times, and connection times greater than the 99% quantile. The resulting data set contains approximately 3.7 million passenger records and 30 variables. The median and mean of Δt are 27.0 minutes and 30.5 minutes, respectively.

3.2. Feature Engineering

It is well known that feature engineering, or creating new input features for machine learning models, is a key part of building an accurate predictive model (Domingos 2012). Every problem is domain-specific so that better features are often the deciding factor of a model's performance. In this study, we create seven features relying on Heathrow experts' knowledge of the aviation domain and the connecting passengers' journey. The features which are engineered are:

Inbound flight region and outbound flight region. Airports in the data set are grouped into four categories based on their retail market regions provided by Heathrow: Europe, East Asia, North America, and rest of the world. According to the data we collected from 2015, the majority of international aircrafts landing at Heathrow were from Europe (39%) and North America (37%). Similarly, the destinations of most flights departing from Terminal 5 were also in these two regions (33% for Europe and 30% for North America).

Punctuality of the arriving flight. Punctuality is defined as the time difference between the actual on-chock time of an inbound flight and its scheduled arrival time. Passengers on late flights might collectively be affected, e.g. in cases of lack of gate availability. In general, 74% of the flights arriving at Heathrow over 2015 are reported to land within 15 minutes of their scheduled arrival times (Heathrow 2015). According to our data, we find that there is a long variability in the punctuality of international arrival flights, with a mean of 42 minutes and standard deviation of 55 minutes.

Hour of the day the arriving flight lands at the airport. A passenger can move faster through the airport during hours when the airport is not busy. The busiest hours in 2015 were 6:00

— 7:00 and 12:00 — 13:00, both with an average of 28 international flights landing at the airport. Compared to other time periods during a day, the airport was relatively calm after 16:00 and before midnight. On average, only 16 flights arrived at the airport per hour during this period.

Perceived connection time. This feature is calculated as the time difference between the arriving flight's on-chock time and the connecting flight's scheduled departure time. The perceived connection time can indicate the stress level of a passenger trying to make the connection. The mean value of this feature is 132 minutes.

Arriving flight load factor and connecting flight load factor. The load factor is calculated as the ratio of the actual number of passengers to the capacity of the flight. Passengers arriving on a flight that is relatively full may need more time to disembark from the aircraft. The mean load factor of the flights in our data set is 0.79, which is slightly higher than the overall average load factor at Heathrow (Heathrow 2016).

A full list of the 37 variables is shown in Table A.1 in the Appendix. Note that we use only 19 of them as predictors because: (i) seven of them are not available in real time (the runway number and stand number of the arriving and connecting flights, local conformance time, conformance location code, and conformance location description); (ii) four variables are removed because they provide information similar to the created variables (arriving and connecting flights' on-chocks time and scheduled time); (iii) the arrival dates of the passengers' arriving and connecting flights are only used to join and split the data; and (iv) five variables are categorical variables with too many levels and cannot be clustered easily (the flight number and aircraft type of the arriving and connecting flights, and the passenger's seat number on the arriving flight). In addition, if we include flight numbers as predictors, the model cannot provide predictions for passengers landing or connecting to a new flight that is not included in the training set. Thus, we try to use generic features, such as aircraft body type, instead of detailed flight information. Summary statistics for the 19 variables that are used as predictors can be found in Table A.2 and Table A.3 in the Appendix.

3.3. The Regression Tree Model for Connection Times

We first select a random 80% of the days in the data as the training set, and the remaining 20% as the test set. We randomly select days instead of passengers' records because of the convenience of predicting passenger flows, or the number of passengers arriving at the conformance desk within certain time windows. The training set is used to tune two parameters of the tree — the maximum depth of the tree and the minimum size of the leaf node.

The scoring rule we use to evaluate the forecasts is the popular pinball loss function (Jose and Winkler 2009, Hong et al. 2016, Grushka-Cockayne et al. 2017). Given a realization y, the pinball loss of the p-quantile (Q_p) is

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

In this study, we score the median, or the 0.50 quantile, and the 0.05, 0.25, 0.75, and 0.95 quantiles that describe the central 50% and 90% prediction intervals. The point forecast we score is the median, and we use the mean absolute error (MAE) to measure its accuracy. The pinball loss function of the 0.50 quantile is equivalent to one-half times the mean absolute error.

The predictive model built in this study is based on the regression tree method. In most cases, the interpretation of the results summarised in a regression tree is simple. This simplicity is not only useful for the purpose of rapid prediction, but can also yield intuitive explanations on why observations are predicted in a particular manner.

The regression tree method is widely used in generating point forecasts; however, relatively few studies have applied it to produce quantiles. We obtain distributional forecasts from a regression tree. After the regression tree is fit, each of the leaves represents a segment of the passengers, and thus contains a number of observations. We find that the majority of the leaves have instances that follow right-skewed distributions. We fit several distributions that can capture right skewness, such as the Gamma and Gumbel distributions. The Gamma distribution stands out because of its out-of-sample accuracy based on the average pinball score over five quantiles. Given a connecting passenger's information, the regression tree will first determine which leaf (or segment) this passenger belongs to; the Gamma distribution attached to this leaf will then provide the quantile forecasts of this passenger's connection time, and the probability of this passenger being late for his onward flight. Another way to model the distribution is to directly use the empirical distribution of the observations within each leaf. This method, however, requires loading in all 3.7 million observations when running simulations in the next session. We choose the parametric method due to its convenience and efficiency.

To avoid overfitting to the training set with a too complex tree, the maximum depth of the tree, d_{max}, and the minimum number of observations, l_{min}, in each leaf are tuned using a five-fold cross validation approach. The tree is fit — for a range of values of the two parameters ($5 < d_{max} < 25$ with an increment of 1; $100 < l_{min} < 1500$ with an increment of 100) — to 80% of the data in the training set, and the pinball loss of the 0.05, 0.25, 0.50, 0.75 and 0.95 quantiles are computed and then averaged in the remaining 20%. This is done in turn for each fold, and the five pinball scores are averaged. The results of the grid search suggest the optimal d_{max} and l_{min} should be 15 and 700, respectively. The tree with these optimal tuning parameters partitions the transfer passengers in our training set into 2,591 segments.

3.4. Quantile Forecasts for Arrivals at Immigration and Security

We use the leaf distributions from the regression tree to generate predictions for the number of passengers arriving at immigration and security. Since Heathrow's security resources are planned in

15 minutes intervals, we use similar windows to construct passenger flow patterns. It is difficult to derive analytical expressions for the distribution of the number of passengers arriving at immigration and security, due to our choice of the Gamma distribution in Section 3.3 and the need to assume passengers' connection times are correlated. Therefore, we generate quantile forecasts by running simulations rather than directly applying analytical results.

We simulate connection times for each of the passengers arriving at Heathrow. For each simulation and each 15-minute interval, we calculate the number of passengers arriving at the immigration and security areas. The distribution of the passenger flow in a 15-minute interval is then approximated by the empirical distribution constructed by the number of arrivals obtained from the simulations. For each 15-minute interval, we only consider passengers arriving in the previous 150 minutes, which is the longest connection time we find in the training set.

Not all passengers travel independently. Many passengers travel in groups, and thus it is important to consider correlations between passengers' connection times. Technically, information is only available for assigning passengers to the same flight. Therefore, in the simulation, we apply a copula approach to incorporate correlations of passengers on the same flight (Nelsen 2007). Similar to the tuning parameters of the regression tree, the copula parameter is selected using the five-fold cross validation approach.

Different copulas can capture different dependence structures. For example, the Gumbel copula exhibits greater dependence in the upper tail than in the lower, perhaps assuming that slower passengers with longer connection times are more correlated. The Clayton copula, on the other hand, exhibits the opposite, or greater dependence in the lower tail. In our study, however, we select the symmetric Gaussian copula, as the differences of the out-of-sample accuracies among using the three copulas were found to be insignificant.

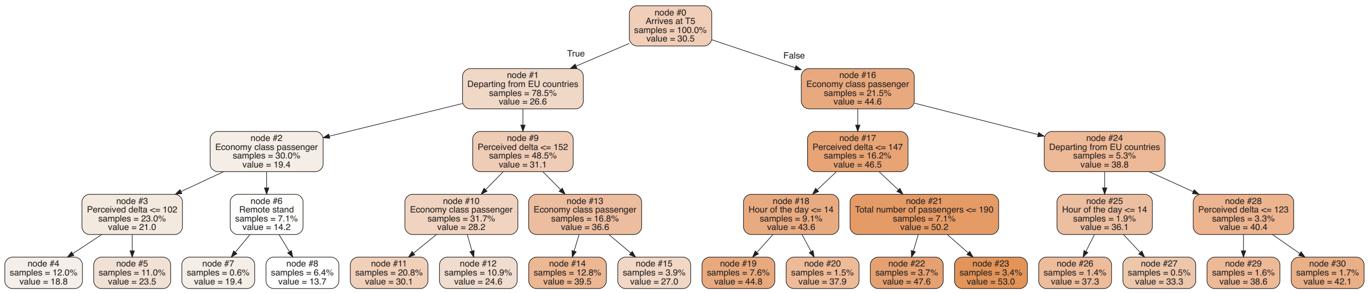
4. Results

Next we present several results derived from our predictive model. We first introduce a new way to define a stable tree. We then highlight a few key findings from the stable tree. Finally, the accuracy of our model in predicting connection times and passenger flows is compared against several benchmarks.

4.1. Key Findings from the Model

Although the tuned tree has more than two thousand leaves, we progress with a reduced version of the tree. The reason of using a reduced tree is twofold. First, a tree with over a thousand leaves is too complex to interpret. Second, the regression tree method is known to be unstable, and small changes in the training set or different settings of the tuning parameters may cause dramatic changes in the leaves. The reduced version of the tree partitions the passengers into only a few segments, making it

Figure 3 The First Four Levels of the Tree Trained to the Entire Training Set



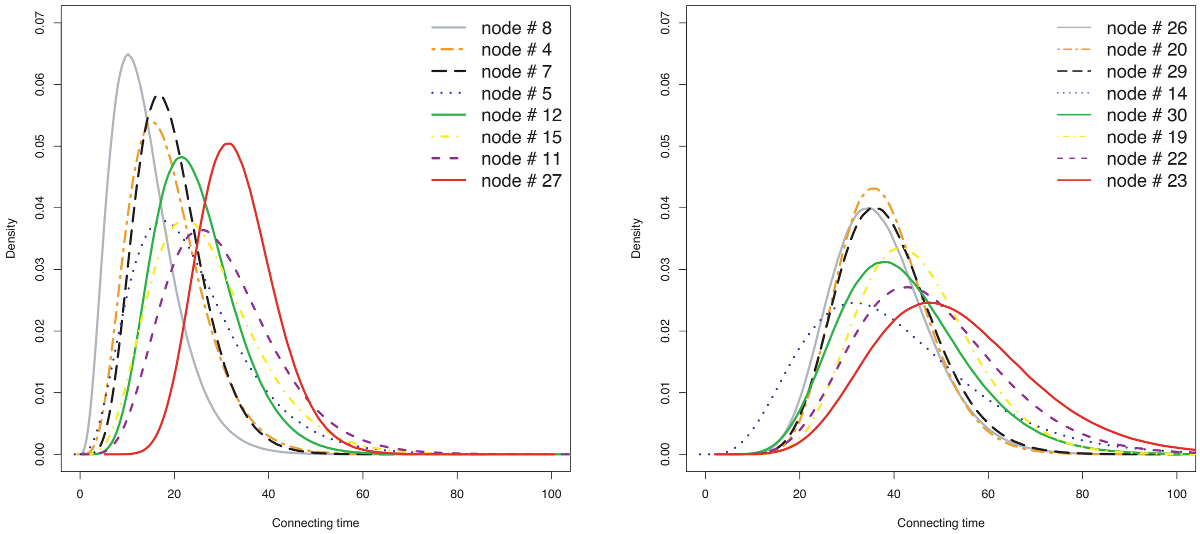
easy to explain why a passenger's connection time is predicted in a particular manner. This tree is also relatively stable, which gives us more confidence to generate insights from it.

We define a tree to be *stable* if it follows two conditions. First, the structure of the tree stays the same when it is fit to different subsamples of the data, and with different settings of the tuning parameters. Here, same tree structure means the splits rely on the same features, and the sequences of these features chosen to be split are the same. Second, the deviations of variables' cutting points across the trees fit to different subsamples and tuning parameters need to be within an acceptable level. To obtain such a stable tree, we first visualize the trees trained using the five folds and with different tuning parameters. We find that the structure of the first four levels of splits remains stable. The coefficients of variation of the variables' cutting points that appear in the first four levels are within 6%. This reduced version of the tree is also visually manageable so that we can easily derive insights from it.

The first four levels of the tree have partitioned the passengers into 16 segments. Figure 3 visualizes the stable tree trained to the entire training set. Each node of the tree presented in the figure gives information on the name of the feature split in this node, percentage of passengers falling in this node, and the average connection time of these passengers. A summary of the 16 segments, represented by the 16 leaf nodes in Figure 3, is provided in Table A.4 in the Appendix. The Gamma distributions fitted to the leaves are shown in Figure 4. Using the results summarized in the table and the figures, we highlight four key findings regarding passengers' connection times.

Key Finding 1: Key Factors The key factors that impact passengers' connection times are (1) whether or not the passenger arrives at Terminal 5, (2) whether or not the passenger arrives from a European Union (EU) country, (3) perceived connection time, (4) whether or not the passenger is in economy class, (5) Hour of the day the arriving flight lands at the airport, (6) Arriving flight's stand type (pier serviced or remote stand), and (7) the total number of passengers on the arriving flight. Although our predictive model contains 19 predictors, the first four levels of the tree only use seven

Figure 4 Gamma Distributions of the Connection Times in the 16 Segments[a]



(a) Segments with low average connection times (b) Segments with high average connection times

[a] The numbered nodes in the figures are the leaf nodes shown in Figure 3.

of them. Therefore, we treat only these seven features listed above as the key factors that impact passengers' connection times.

Key Finding 2: Factor Interactions The regression tree allows us to quickly identify complex interactions among the factors. Arriving terminal and travel class dictate all 16 segments. According to Figure 3, these two variables appear in every branch of the tree. Whether or not the passenger arrives from a EU country and the perceived connection time are determining features for 12 segments. A passenger's connection time does not seem to depend on the departing region of their arriving flight if they are economy passengers and arrive at Terminal 2, 3 or 4. Moreover, passengers' connection time does not seem to be influenced by their perceived connection time if they are business or first class passengers and travel from EU countries.

Key Finding 3: Expected Connection Times Passengers arriving at Terminal 5, from EU countries, in business or first class, whose arriving flight parks at a pier served stand take the shortest amount of time to make their connections. We also find that passengers arriving at Terminal 5 and departing from EU countries have shorter connection times compared with other passengers. This is regardless of the value of the rest predictors. According to Table A.4, these passengers form the first four segments that have the lowest connection times.

Passengers arriving at Terminal 2, 3 or 4, in economy class, with perceived connection time greater than 147 minutes, whose arriving flight has a total number of passengers greater than 190 take the longest time to make their connections.

Key Finding 4: Uncertainty of Connection Times In general, the uncertainty and average value of passengers' connection times within a segment are positively correlated. The uncertainty in the segment with the shortest average connection time is also the lowest, as the standard deviation is the smallest across all 16 passenger segments. For most segments shown in Figure 4 (b), not only are the average connection times longer, but the variances among the passengers are also larger compared to those in Figure 4 (a). The coefficient of variation within a segment, defined as the ratio of the standard deviation to the average connection time, is however negatively correlated with the average connection times. Therefore, when taking the average connection time into account, the higher the connection time, the more predictable the journey. Moreover, although the segment with the shortest average connection time has the lowest standard deviation, it is the least predictable when considering the magnitude of the mean.

4.2. Accuracy of Forecasts for Individuals' Connection Times

Among the methods that can generate quantile forecasts, four are widely used by the machine learning community: linear regression, quantile regression, quantile regression forest, and gradient boosting machine (Hong et al. 2016). Quantile regression and the gradient boosting machine estimate different quantiles independently, possibly resulting in lack of monotonicity in the estimated quantile function. This longstanding problem is also known as the quantile crossing problem (Bassett and Koenker 1982). Chernozhukov et al. (2010) propose a method of rearranging the curve into a monotone curve. However, this rearranging method requires the entire quantile regression function, and thus the estimates of thousands of quantile regressions, making the method computationally expensive.

Linear regression and quantile regression forest are able to produce monotone quantiles. Linear regression is also easy to fit and usually runs fast. However, it does not perform well with non-linear relationships and complex interactions. Quantile regression forest is a generalization of the random forest and is competitive in terms of predictive power. However, it is time consuming and typically treated as a black box.

We compare our regression tree model with these four methods and a naïve forecast on the test set. The naïve model of predicting passengers' connection times is based on their arrival terminals, the most important predictor given by our regression tree. Specifically, given a new passenger's arrival terminal, the quantiles of her connection time are predicted as the quantiles of the connection times of passengers arriving at the same terminal in the training set. The linear regression and the quantile regression are fit to all 19 variables that were selected as predictors in Section 3.3 and their 252 pairs of interactions.

The other two methods, quantile regression forest and gradient boosting machine, have tuning parameters to avoid overfitting. Quantile regression forest fits independent trees and constructs conditional distributions from these trees. It has several tuning parameters, such as the number of trees

Table 1 Accuracy of Forecasts on Connection Times in the Test Set

	MAE	Pinball losses					
		$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$	average
Naïve model	10.20	0.96	3.54	5.10	4.86	2.44	3.38
Linear regression	8.52	1.18	3.14	4.26	4.31	2.15	3.01
Quantile regression	8.18	**0.78***	2.81	4.09	**3.99**	2.03	2.74
Quantile regression forest	8.24	0.81	2.84	4.12	4.01	2.06	2.77
Gradient Boosting Machine	8.38	0.81	2.90	4.19	4.07	2.07	2.81
Regression tree	**8.16**	0.80	**2.77***	**4.08**	4.01	**1.98***	**2.73***

The symbol * indicates significance at the 0.1% level.
Values in bold indicate the lowest errors.

in the forest and the minimum node size. Gradient boosting is an ensemble technique in which the regression trees are not fit independently, but sequentially, learning from one tree to the next. Since these two methods are extremely time consuming and require large memory size, we are limited in our ability to fit all 19 variables. Instead, we fit the models to the seven key factors identified by our regression tree. We tune the number of trees and the learning rate of the gradient boosting machine. For the quantile regression forest, we set the number of trees and the minimum node size to 100 and 1000, respectively. It takes more than six hours to train the quantile regression forest on a machine with 16 cores.

Among all six models shown in Table 1, our regression tree has the lowest average pinball loss (2.73) and is best at three of the five quantiles. Not surprisingly, the naïve model performs the worst with an average pinball loss of 3.38. The quantile regression forest and the gradient boosting machine, which are often considered as advanced machine learning methods with high accuracies, perform worse than the second best model, the quantile regression. The weak performance of these two advanced machine learning methods is likely due to the fact that only seven variables were used to train the models. In terms of point forecasting, our regression tree model has the lowest MAE. The difference between the MAE of our regression tree model and the second best model, however, is not significant.

While our regression tree approach is not significantly more accurate in generating point forecasts of passengers' connection times, it can be easily applied to forecast the number of passengers arriving at immigration and security. The current implementation of quantile regression forests in leading machine learning tools, such as `R` or `Python`, does not provide complete leaf-distributions as outputs. As mentioned above, the gradient boosting machine and the quantile regression fit quantiles independently. To construct an empirical distribution and simulate from it, thousands of models need to be fit and stored, making the model-training process inefficient. The crossing problem inherent to these two models may also make it difficult to construct empirical distributions.

Table 2 Accuracy of Forecasts on Passenger Flows in the Test Set

	MAE	Pinball losses					
		$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$	average
Naïve model	45.62	5.37	17.66	22.81	18.53	6.14	14.10
Linear regression	27.84	3.44	10.79	13.92	11.56	4.11	8.76
Regression tree	**19.88***	**2.99***	**9.27***	**11.98***	**9.94***	**3.49***	**7.53***

The symbol * indicates significance at the 0.1% level.
Values in bold indicate the lowest errors.

4.3. Accuracy of Aggregate Forecasts for Arrivals at Immigration and Security

We use the five-fold cross validation approach to search for the copula parameter giving the most accurate predictions for passenger arrivals at immigration and security areas. The average pinball score over five quantiles calculated over peak hours is used to select the copula parameter. If the sum of the actual number of passengers arriving in a given 15-minute interval and those of the previous three 15 minutes is above 200, this 15 minutes interval can be regarded as part of a peak hour (De Reyck et al. 2016).

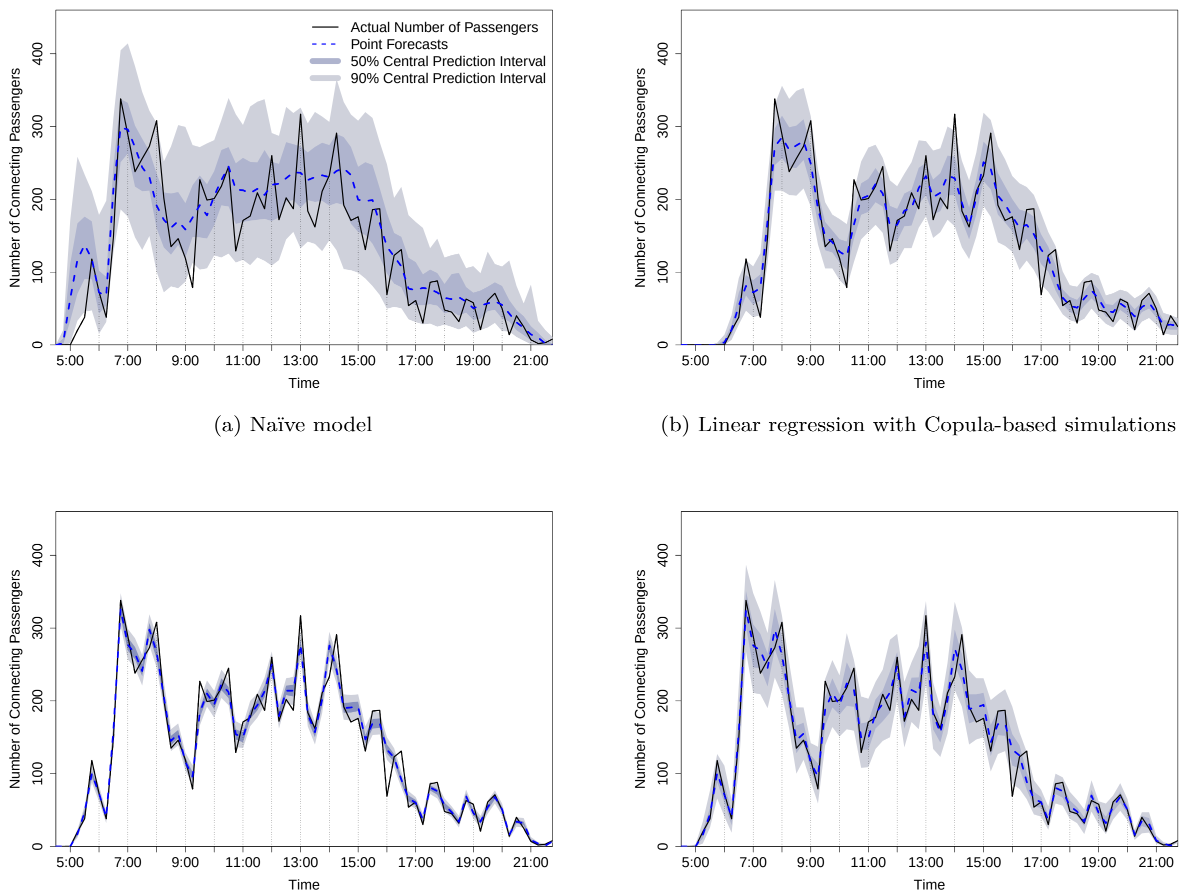
Based on the grid search on the five folds, setting the correlation to 0.5 provides well-calibrated forecasts whereby the correct percentage of realizations falls within the 50% and 90% prediction intervals. This correlation provides the lowest average pinball loss as well. We also notice that the prediction interval becomes wider as the dependencies between passengers' connection times increase.

We construct a naïve model as a benchmark to predict passenger flows. For each day in the training set, we calculate the number of passengers arriving at immigration and security areas during each 15-minute interval. These numbers of passengers across all days in the training set are then used to construct the distribution of the passenger flows. Since linear regression has an assumption that the error terms follow a normal distribution, we can easily simulate from the model and generate quantile forecasts for passenger flows. Similar to our regression tree model, the optimal Gaussian copula parameter is obtained using the five-fold cross validation.

Next we measure the accuracy of the models in predicting passenger flows over peak hours. The pinball losses and the MAE of the forecasts are presented in Table 2. Our regression tree model outperforms the other two models in all five quantiles. The MAE of the point forecasts generated by our model is also the lowest. The differences in the pinball losses and the MAE between the regression tree and the linear regression model are statistically significant at the 0.1% level.

Figure 5 presents the prediction intervals for a randomly selected day from our test set. These prediction intervals are generated by the naïve model (Figure 5a), the linear regression (Figure 5b), and the regression tree model using independent and copula-based simulations, Figure 5c and Figure 5d, respectively. It is easy to observe the accuracy of the point forecasts from the regression tree model. The prediction intervals produced when assuming independence (Figure 5c) are clearly much

Figure 5 Number of Passenger Arrivals at Immigration and Security Areas on Test Day August 31, 2015



(a) Naïve model

(b) Linear regression with Copula-based simulations

(c) Regression tree with independent simulations

(d) Regression tree with Copula-based simulations

narrower than when using copula-based simulations (Figure 5d). Such narrow intervals could result in overconfidence and misjudged decisions regarding peak activity.

5. Real-Time Implementation

With the predictive model at hand, we worked with Heathrow's APOC to develop an application for forecasting individual connection times, and the flow of transfer passengers into the immigration and security areas in real time. We use these predictions to also calculate the probability of a passenger missing her connection, and the expected number of late passengers for each outbound flight. The application generates forecasts on a rolling basis and updated every 15 minutes. A prototype of the predictive system was first tested at Heathrow in 2016.

5.1. Real-Time Input Data

Real-time data was exported from the Airport Flight Operations System, also known as IDAHO. IDAHO contains all the required flight level information. Real-time passenger level information was obtained from Passenger Transfer Message (PTM) files. These files are sent by the airline when a flight takes off from an origin airport. PTM files are currently the only real-time data source that contains passenger level information. How far in advance of the flight's arrival these files get sent by the airline has a significant impact on the accuracy of our predictions; for instance, missing PTMs can cause forecasts to underestimate the passenger flow. According to Heathrow, PTMs for 95% of the passengers are received more than 60 minutes prior to arrival, 88% more than 90 minutes, and 68% more than 120 minutes prior to arrival. Therefore, to ensure that there is sufficient data to enable the predictive model to provide accurate forecasts, the forecasting window chosen for the application was 90 minutes.

It should also be noted that the actual on-chock time will not be available if an aircraft is en-route. In this case, we used estimated on-chock time to calculate the punctuality of the arriving flights and the passengers' arrival times at the conformance desk. If we had neither of the actual or estimated on-chock time, we used the actual time of operation, the estimated time of operation, or the scheduled time of operation. If none of these five fields were presented in IDAHO, we dropped the record.

5.2. Forecasting Application

A prototype for real-time forecasting of passenger connection times was developed using a `Python` GUI scripting interface. Forecasts were generated every 15 minutes, on a rolling basis, for the next 90-minute time window. At the start of each iteration, the application collected real-time information of passengers who have arrived in the previous 150 minutes or will arrive in the next 90 minutes. The application ran for approximately three minutes on a typical Heathrow machine, and generated as output several CSV files containing individual level and aggregate level forecasts. Plots were also generated to visualize the forecasts of transfer passenger flows.

In the application, users can set the granularity of the passenger flow and the forecasted time-horizon. Forecasts with different granularities are generated for different purposes. For example, the forecasts of passengers arriving in every 5 minutes provide detailed flow profiles, while the forecasts of passengers arriving in every 15 minutes can be used to adjust resourcing plans.

Figure 6 shows an example of the output for the forecasts of individual connection times generated at 12:00 on July 1, 2016. Each row in the file represents the forecast for one passenger. In addition to the quantile forecasts of passengers' arrival time at immigration and security areas, this file also contains probabilities of these passengers being late for their connecting flights. Here, a passenger is considered to be late if they arrive at immigration and security later than 30 minutes before the scheduled departure time of the connecting flight. The 30 minutes threshold is often used by the airlines to rebook late passengers. Heathrow and the airlines can easily identify which passengers are at risk of missing their onward flight based on the forecast probabilities. Given this information, they would be able to expedite late passengers and facilitate early re-booking.

Figure 6 Output from the Application: Individual Connection Times

	A	B	C	D	E	F	G	H	I	J
1	passenger_id	on_chock_time	q0.05	q0.25	median	q0.75	q0.95	ib_flight_no	ob_flight_no	P(missing connecting flight)
2	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	BA847	BA293	0
3	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	BA479	BA069	0
4	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	BA309	BA115	0
5	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	BA763	BA279	0.11
6	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	BA565	BA287	0.01
7	322652	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	MS777	BA269	0
8	323561	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	BA757	BA1394	0.80
9	323767	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	BA343	BA049	0
10	323759	01/07/2016 12:28	01/07/2016 12:39	01/07/2016 12:43	01/07/2016 12:47	01/07/2016 12:52	01/07/2016 13:03	BA343	BA227	0
11	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	BA805	BA1484	0.96
12	324030	01/07/2016 11:42	01/07/2016 11:59	01/07/2016 12:07	01/07/2016 12:14	01/07/2016 12:22	01/07/2016 12:36	BA309	BA113	0
13	323657	01/07/2016 12:11	01/07/2016 12:28	01/07/2016 12:32	01/07/2016 12:38	01/07/2016 12:44	01/07/2016 12:57	BA573	BA197	0.22
14	323967	01/07/2016 11:34	01/07/2016 11:45	01/07/2016 11:49	01/07/2016 11:52	01/07/2016 11:57	01/07/2016 12:07	BA431	BA227	0
15	324058	01/07/2016 12:10	01/07/2016 12:34	01/07/2016 12:39	01/07/2016 12:45	01/07/2016 12:51	01/07/2016 13:04	EI712	BA279	0
16	321912	01/07/2016 12:28	01/07/2016 12:51	01/07/2016 12:58	01/07/2016 13:05	01/07/2016 13:12	01/07/2016 13:25	QR003	BA203	0

ID of the passenger

Inbound flight estimated on-chock time

0.05, 0.25, 0.50, 0.75 and 0.95 quantiles of a passenger's arrival time at immigration and security areas

Inbound and outbound flight no.

probability of a passenger being late for her connection flight

In the output shown in Figure 7, we grouped passengers by their outbound flights, and calculated how many of them are expected to be late based on the simulation results. We also calculated the number of passengers that would be still at risk if the airline delayed the departure time by 5, 10, 20, and 30 minutes. These forecasts can help improve the predictability and stability of outbound flights' scheduled departure time. During the aircraft turn around, its scheduled departure time can be adjusted, but only a few times. Many airlines consider passenger delays when they amend flights' scheduled departure times. In this case, predictions of passenger delays ahead of time can help airlines make accurate adjustments. Moreover, if our predictions are accurate, no further changes are to be expected because of late transfer passengers; and therefore, the stability of the scheduled departure time can be improved.

Figure 7 Output from the Application: Expected Number of Late Passengers for Each Outbound Flight

	A	B	C	D	E	F
1	ob_flight_no	current SDT	+5 min	+10 min	+20 min	+30 min
2	BA005	3	2	2	1	0
3	BA007	22	20	17	11	6
4	BA009	1	1	0	0	0
5	BA011	0	0	0	0	0
6	BA017	2	2	1	1	1
7	BA031	0	0	0	0	0
8	BA033	0	0	0	0	0

Outbound flight no.

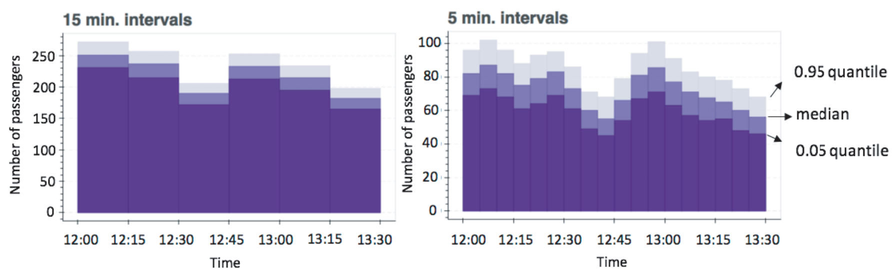
Expected number of late passengers with current scheduled time for departure, and current scheduled time for departure + 5 min(C), +10 min(D), +20 min(E), and +30 min(F)

Figure 8 Output from the Application: Arrivals at Immigration and Security

	A	B	C	D	E	F	G
1	inter_begin	inter_end	q0.05	q0.25	median	q0.75	q0.95
2	01/07/2016 12:00	01/07/2016 12:15	233	244	253	260	271
3	01/07/2016 12:15	01/07/2016 12:30	217	227	235	244	257
4	01/07/2016 12:30	01/07/2016 12:45	172	183	190	197	207
5	01/07/2016 12:45	01/07/2016 13:00	214	225	233	240	252
6	01/07/2016 13:00	01/07/2016 13:15	198	207	216	224	234
7	01/07/2016 13:15	01/07/2016 13:30	163	173	179	189	199

Beginning and ending time of the 15 minutes time interval

0.05, 0.25, 0.50, 0.75 and 0.95 quantiles of the number of passengers arriving at immigration and security areas in windows 15 minutes intervals



In the output file of passenger flow forecasts (Figure 8), each row represents the forecast for a 15-minute time interval. The 90% prediction intervals and the median of the 5-minute and 15-minute passenger flows are also visualized in the output figures. These real-time predictions allow the dynamic planning of immigration and transfer security resourcing, by applying the real-time demand forecasts to Heathrow's existing lane planning tool.

5.3. Implementation and Impact

Our predictive system has been implemented at Heathrow since 2017, integrated into the APOC's Dynamic Model of Operations (DMO) system. The probability of a passenger missing her connection appears in a "Connection at Risk" table, and the passenger flow at the Conformance desks is shown in a "Transfer Security Flow" table. The forecasts of the number of late passengers for each departing flight are provided to APOC's connections team who liaise with the airline.

The predictive system is based on the Azure Machine Learning platform. The regression tree is retrained daily with rolling five years of historical data. Unlike our initial approach that fits only Gamma distribution to the terminal nodes, the predictive system in practice fits multiple distributions, including the normal, exponential, Gumbel, Weibull, and Gamma distribution. The system then selects the best distribution for each leaf, according to the in-sample fitting performance. The results show that the Gamma distribution appears to be the best distribution in more than 80% of the cases. On the day of operation, the real-time data is read from Heathrow's SQL server into Azure ML in every 15 minutes. The forecasts for the next 90 minutes are then saved in CSV format to an Azure Blob storage, which is designed to store excessively large quantities of data files.

The Flow Managers in the APOC are currently using the predictive system to optimize the operation for a smoother and more predictable service. Passenger experience has been improved through reduced queuing, as capacity and resourcing along the journey more closely match with the dynamic demand. In addition, the managers are able to identify passengers who are at risk of missing their connection, and work with Heathrow and airline teams to assist and expedite their journeys.

An accuracy test of the expected passenger flows has been conducted over July and August 2017. The MAE of the forecasts generated from the new predictive system is 38.7, which is 21% lower in comparison to Heathrow's previous system that only generates static predictions the day before operation.

Encouraged by the good results, Heathrow's APOC has expanded the usage of the predictive system to enhance other airport services. Demand forecasts for bags and passengers with restricted mobility have now all been developed using predictive techniques. In addition, real-time prediction of the direct departure flows has been developed to support the optimization of direct security operation. The team is now looking at how these techniques could be applied to other parts of the passenger journey such as surface access and trolley operations.

6. Conclusion and Future Work

In this paper, we offer a first study of passengers' transfer journeys using data provided by Heathrow airport. We develop a system to provide real-time information about transfer passengers' journeys through the airport. This information is vital for the airport in order to best serve the passengers, the airlines, and their employees.

Our study makes advanced machine learning accessible to managers with no background working in data science at the APOC. It also enables the APOC to move from fragmented data and Excel based reporting to a more sophisticated data science environment. The capabilities demonstrated by this study inspired institutional investment in improving data science skills. Moreover, the impact of this work extends beyond Heathrow's APOC, with interest expressed by Aéroports de Paris.

The work here demonstrates the usefulness of the regression tree method. Although simple in nature, it provides accurate predictions and easy interpretations. While other models served only as benchmarks here, we encourage continuous exploration of alternative models for improvement. For instance, we can imagine that ensembles, or a combination of several models, could result in improved performance. In addition, since only few existing machine learning methods are designed to generate the entire quantile function, we encourage future work to focus on developing superior models in forecasting whole probability distributions.

Appendix

Table A.1 Descriptions of the Variables in the Data Set

Data set	Variable name	Description
BOSS[a]	on chocks time	The time when an aircraft is parked at gate.
	aircraft body	A flight's aircraft body type: W (wide) or N (narrow).
	aircraft type	A flight's aircraft type. There are 23 types in total.
	pax capacity	The capacity of the flight.
	pax total	Total number of passengers on the flight.
	pax transfer	Number of transfer passengers on the flight.
	runway no.	Runway number of the flight. There are four runway numbers in the dataset: 27L, 27R, 09L, and 09R.
	scheduled time	Scheduled arrival/departure time of the flight.
	stand no.	Stand number of the flight. There are 213 stand numbers in total.
	inbound date [c]	The date of a flight arrives at the airport.
	flight no.[c]	There are 692 and 399 unique flight numbers for arriving and connecting flights, respectively.
BDD	**pax travel class**	Passengers' travel class on the arriving flight.There are five classes in the data set. We grouped them into two categories: economics (EC) or business and first class (NEC).
	ib terminal	A passenger's arriving terminal.
	ib stand type	Stand type of the arriving flight: P (Pier served stand) or R (Remote stand).
	ob stand type	Stand type of the connecting flight.
	pax seat no.[b]	A passenger's seat number on the arriving flight.
Conformance	local conform time	Timestamp of when a passenger arrives at Conformance desk.
	conform location code	Code of the conformance desk.
	conform location descrp	Terminal number, conformance desk number, and international or domestic connecting flight.
Created variables	**ib region**	The region of the departure airport for the arriving flight. There are four regions: UK, Europe, North America, and the rest of world.
	ob region	The region of the connecting airport for the arriving flight.
	ib punct	Punctuality of the arriving flight.
	ib hour	Hour of the day when the arriving flight lands at the airport.
	perceived delta	Time difference between the inbound flight's on-chock time and the outbound flight's scheduled departure time.
	ib load	Load factor of the arriving flight. Defined as the ratio of the actual number of passengers to the capacity of the flight for inbound flight and outbound flight.
	ob load	Load factor of the connecting flight.

Variables in bold represent the 19 predictors used to train the model.

[a] The BOSS data set stores information for every flight landing or departing from Heathrow. The variables listed here under the BOSS data set are for both arriving and connecting flights. Therefore, we have obtained 22 variables from the BOSS data set.

[b] This variable also appears in the Conformance data set.

[c] This variable also appears in the BDD and the Conformance data set.

Table A.2 Summary Statistics of the Numerical Predictors

	mean	median	standard deviation
inbound flight pax capacity	221	189	84
outbound flight pax capacity	228	205	84
inbound pax total	176	151	86
outbound pax total	187	161	90
inbound pax transfer	62	48	50
outbound pax transfer	91	76	61
ib punc	42	49	55
perceived delta	71	56	87
ib load	0.79	0.83	0.18
ob load	0.80	0.85	0.18

Table A.3 Summary Statistics of the Categorical Predictors

	Summary
inbound aircraft body	55% of the flights' body type was "Narrow", the others were "Wide"
outbound aircraft body	51% of the flights' body type was "Narrow", the others were "Wide"
ib hour	The busiest hours in 2015 were 6:00 - 7:00 and 12:00 - 13:00, both with an average of 2 8 international flights landing at the airport.
pax travel class	73% of the passengers traveled in economy class
ib terminal	65% of the passengers arrived at T5, the others arrived at other terminals
ib stand type	91% were pier serviced, the others were remote stand
ob stand type	90% were pier serviced, the others were remote stand
ib region	39%, 36%, 6%, and 19% of the passengers were from EU countries, North American, Asia, and rest of the world
ob region	49%, 30%, 5%, and 16% of the passengers traveled to EU countries, North American, Asia, and rest of the world

Table A.4 Descriptions of the 16 Passenger Segments

node number[a]	Average conecting time	Std. of the connecting times	Arriving terminal	Departing region of the arriving flight	Travel class of the arriving flight	Perceived connection time	Stand type of the arriving flight	Hour of the day the arriving flight lands at the airport	Total numbber of passegers on the arriving flight
8	13.7	6.9	Terminal 5	EU countries	Business/First		Pier		
4	18.8	8.0	Terminal 5	EU countries	Economy	<= 102 minutes			
7	19.4	7.3	Terminal 5	EU countries	Business/First		Remote		
5	23.5	11.9	Terminal 5	EU countries	Economy	> 102 minutes			
12	24.6	8.8	Terminal 5	Non-EU countries	Business/First	<= 152 minutes			
15	27.0	11.3	Terminal 5	Non-EU countries	Business/First	> 152 minutes			
11	30.1	11.7	Terminal 5	Non-EU countries	Economy	<= 152 minutes			
27	33.3	8.1	Terminal 2/3/4	EU countries	Business/First			After 14:00	
26	37.3	10.3	Terminal 2/3/4	EU countries	Business/First			Before 14:00	
20	37.9	9.5	Terminal 2/3/4		Economy	<=147		After 14:00	
29	38.6	10.3	Terminal 2/3/4	Non-EU countries	Business/First	<= 123			
14	39.5	17.8	Terminal 5	Non-EU countries	Economy	> 152 minutes			
30	42.1	13.4	Terminal 2/3/4	Non-EU countries	Business/First	>123			
19	44.8	12.3	Terminal 2/3/4		Economy	<=147		Before 14:00	
22	47.6	15.4	Terminal 2/3/4		Economy	>147			<= 190
23	53.0	17.0	Terminal 2/3/4		Economy	>147			>190

[a] The numbered nodes are the leaf nodes shown in Figure 3.

References

Atkinson SE, Ramdas K, Williams JW. 2016. Robust scheduling practices in the us airline industry: Costs, returns, and inefficiencies. *Management Science* **62**(11) 3372–3391.

Barnhart C, Cohn A. 2004. Airline schedule planning: Accomplishments and opportunities. *Manufacturing & service operations management* **6**(1) 3–22.

Barnhart C, Fearing D, Vaze V. 2014. Modeling passenger travel and delays in the national air transportation system. *Operations Research* **62**(3) 580–601.

Bassett Jr G, Koenker R. 1982. An empirical quantile function for linear models with iid errors. *Journal of the American Statistical Association* **77**(378) 407–415.

Bertsimas D, Patterson SS. 2000. The traffic flow management rerouting problem in air traffic control: A dynamic network flow approach. *Transportation Science* **34**(3) 239–255.

Breiman L, Stone CJ, Friedman J, Olshen RA. 1984. *Classification and regression trees*. CRC press.

Chernozhukov V, Fernández-Val I, Galichon A. 2010. Quantile and probability curves without crossing. *Econometrica* **78**(3) 1093–1125.

De Reyck B, Guo X, Grushka-Cockayne Y, Lichtendahl Jr. KC, Karasev A, Garside T, Coss N, Tasker F. 2016. Apoc business process reengineering big data study. Tech. rep.

Domingos P. 2012. A few useful things to know about machine learning. *Communications of the ACM* **55**(10) 78–87.

Eliashberg J, Hui SK, Zhang ZJ. 2007. From story line to box office: A new approach for green-lighting movie scripts. *Management Science* **53**(6) 881–893.

Eurocontrol. 2010. The potential role of the Airport Operations Centre (APOC) in the SESAR airport concept. Tech. rep.

Ferreira KJ, Lee BHA, Simchi-Levi D. 2015. Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management* **18**(1) 69–88.

Ferrovial. 2016. @techreport recognized for excellent passenger service by skytrax. `https://www.ferrovial.com/en/press-room/news/heathrow-recognized-for-excellent-passenger-service-by-skytrax/`.

Friedman J, Hastie T, Tibshirani R. 2001. *The elements of statistical learning*. Springer Series in Statistics.

Grushka-Cockayne Y, Jose VRR, Lichtendahl Jr. KC, Winkler RL. 2017. Quantile evaluation, sensitivity to bracketing, and sharing business payoffs. *Operations Research* **65**(3) 712–728.

Heathrow Airport. 2015. Flight performance annual report. `https://www.heathrow.com/file_source/HeathrowNoise/Static/Heathrow_flight_performance_annual_report_2015.pdf`.

Heathrow Airport. 2016. Annual report and financial statements for the year ended 31 december 2015. `https://www.heathrow.com/company/investor-centre/document-centre/annual-accounts`.

Hong T, Fan S, Pinson P, Zareipour H, Troccoli A, Hyndman RJ. 2016. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *International Journal of Forecasting* 357–363.

Jose VRR, Winkler RL. 2009. Evaluating quantile assessments. *Operations Research* **57** 1287–1297.

Lan S, Clarke JP, Barnhart C. 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation science* **40**(1) 15–28.

Lohatepanont M, Barnhart C. 2004. Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment. *Transportation Science* **38**(1) 19–32.

MacDonald M. 2016. Annual analyses of the EU air transport market. Tech. rep.

Meinshausen N. 2006. Quantile regression forests. *Journal of Machine Learning Research* **7**(Jun) 983–999.

Mukherjee A, Hansen M. 2009. A dynamic rerouting model for air traffic flow management. *Transportation Research Part B: Methodological* **43**(1) 159–171.

Nelsen RB. 2007. *An introduction to copulas*. Springer Science & Business Media.

Shang Y, Dunson D, Song JS. 2017. Exploiting big data in logistics risk assessment via bayesian nonparametrics. *Operations Research* **65**(6) 1574–1588.

Wei W, Hansen M. 2006. An aggregate demand model for air passenger traffic in the hub-and-spoke network. *Transportation Research Part A: Policy and Practice* **40**(10) 841–851.

Xue Z, Wang Z, Ettl M. 2015. Pricing personalized bundles: A new approach and an empirical study. *Manufacturing & Service Operations Management* **18**(1) 51–68.