

Predictive modeling of inbound demand at major European airports with Poisson and Pre-Scheduled Random Arrivals

Carlo Lancia^{a,*}, Guglielmo Lulli^b

^a*Leiden University Mathematical Institute, Niels Bohrweg 1, 2333 CA, Leiden, NL*

^b*Department of Management Science & Center for Transportation and Logistics, Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK*

Abstract

This paper presents an exhaustive study of the arrivals process at eight major European airports. Using inbound traffic data, we define, compare, and contrast a data-driven in-homogeneous Poisson and Pre-Scheduled Random Arrivals (PSRA) point process with respect to their ability to predict future demand. As part of this analysis, we show the weaknesses and difficulties of using a non-homogeneous Poisson process to model the arrivals stream. On the other hand, our novel and simple data-driven PSRA model captures and predicts the main properties of the typical arrivals stream with good accuracy. These results have important implication for the modeling and simulation-based analyses of inbound traffic and can improve the use of available capacity, thus reducing air traffic delays. In a nutshell, the results lead to the conclusion that, in the European context, the PSRA model provides more accurate predictions.

Keywords: Transportation, Air traffic, Demand prediction, Data-driven modeling

2010 MSC: 90B06, 62P30

*Corresponding author.

Email addresses: c.lancia@math.leidenuniv.nl (Carlo Lancia),
g.lulli@lancaster.ac.uk (Guglielmo Lulli)

URL: <https://sites.google.com/view/clancia> (Carlo Lancia)

1. Introduction

Air congestion is a regular and persistent phenomenon in the air traffic system in both the US and Europe. Over the years, air traffic demand has increased at a much faster pace compared to the increment of air traffic system capacity. In the last decade, we have witnessed a mitigation of congestion phenomena, with air traffic demand just recovering from the 2008 economic crisis (EUROCONTROL, 2018, §1.2). Yet, the latest air traffic statistics published by EUROCONTROL show a significant deterioration of on-time performance in the European Civil Aviation Conference area: the average delay per flight is at its highest in the last 10 years (EUROCONTROL, 2016a). As a consequence of these level of congestion, 7,167 flights were canceled and 107,426 delayed in Europe between November 28 and December 27, 2016. The situation was even worse in the US, as the numbers of canceled and delayed flights were twice as great as the figures recorded in Europe (FlightStats). However, for the sake of completeness, the number of controlled flights in the US is much larger: 15.3 million in the US versus 9.9 million flights in Europe in 2015 (EUROCONTROL & FAA, 2015).

Airports are the most critical bottlenecks of the air traffic system. The Arrival Sequencing and Metering Area (ASMA) additional time –which is a proxy for the average arrival runway queuing time of the inbound traffic flow– during times when the airport is experiencing high demand, is an indicator of airport congestion (Cappelleras). In 2015, the average ASMA additional time at the top 30 European airports amounted to 2.27 minutes per arrival, increasing by about 18% with respect to the previous year. The ASMA performance deterioration in 2015 was largely driven by an increase in average additional ASMA time at London Gatwick, Stockholm Arlanda, Dublin, and Brussels. London Heathrow has by far the highest level of average additional ASMA time in Europe, which is almost 9 minutes per arrival, followed by London Gatwick with more than 4 minutes per flight (EUROCONTROL, 2016b). Similar situations occur in the US, although with less contrast in additional time reported across airports (EU-

ROCONTROL & FAA, 2015). This situation occurs despite the fact that the principal airports in Western and Central Europe are treated as *fully coordinated*, meaning essentially that the number of flights that can be scheduled per hour (or other unit of time) is not allowed to exceed airport *declared capacity* (de Neufville & Odoni, 2003). In the U.S., scheduling limits are applied only to airports of the New York region, Washington Reagan, and Chicago O’Hare airport, under the High Density Rule.

Starting with the pioneering work of Blumstein (1959), airport operations have attracted the interest of the scientific community in the attempt to alleviate congestion. Many quantitative methods have been developed to understand the various causes of congestion. These methods aim to ameliorate the level of congestion by detecting possible actions for improving the use of capacity and reducing delays. In particular, a great amount of work has been devoted to study the arrivals process at airports and the corresponding queues. Given the stochastic nature of the phenomenon, most of these studies rely on either queuing theory or simulation models. To estimate congestion with reasonable accuracy, models should include both *i*) fluctuations in the inbound demand rate over time due to hub-and-spoke operations carried out at major airports and *ii*) randomness affecting the arrivals. Koopman assumed that the statistics of arrivals follow a Poisson law, but with an arrival rate that is a strongly-varying function of time according to quantities actually observed at airports. According to Hengsbach & Odoni (1975), the assumption of Poisson arrivals for airport demand has two very appealing properties: *i*) it is mathematically tractable and is consistent with observations at major airports, and *ii*) it has been extensively used in the transportation literature.

Poisson arrivals have been assumed to study the arrival streams at several airports: J.F. Kennedy and La Guardia (Koopman, 1972), Toronto Pearson (Bookbinder, 1986), and Boston Logan (Hengsbach & Odoni, 1975) among others. Yet, this assumption has been corroborated only in more recent times (Willemain et al., 2004). In that paper, the authors examined data on arrivals to nine major US airports during December 2003 for evidence of exponentiality in the

distribution of the interval between two estimated arrival times, i.e. the arrival times computed by the Enhanced Transportation Management System algorithm in use by the Federal Aviation Administration when the aircraft were 100 miles from their destinations. Willemain et al. (2004) performed the analysis of these intervals under the assumption that they “are independent samples from a Weibull distribution with a fixed shape parameter (equal to 1 for an exponential distribution) and a slowly varying scale parameter.” The results confirmed the near-exponentiality of the inter-arrivals, therefore supporting the idea of describing arrivals through a non-homogeneous Poisson process –it is well known that a process is Poisson if and only if it has independent and exponentially-distributed inter-arrival times. However, there are some inherent issues with (in-)homogeneous Poisson arrivals. First, any Poisson process –homogeneous or not– is by definition not capable of modeling any correlation between arrivals in consecutive time periods. This leads to an overestimation of the queue length presumably because the uncaptured correlation is negative (Caccavale et al., 2014). The overestimation of the queue length has a strong impact on the determination of control actions (decisions) to make efficient use of the available capacity: models adopting homogeneous Poisson processes may overestimate congestion and yield too conservative decisions. Second, if we model the arrival stream as a non-homogeneous Poisson process, a possibly large number of parameters has to be estimated unless the intensity of the process changes slowly over time. Third, the arrival rates at several major European airports tend in fact to change rather fast, as highlighted by the average-demand curves of Figure 4 in Section 3.2; this situation implies that the methods and results proposed by Willemain et al. (2004) not applicable in a European context.

To overcome these issues, Guadagni et al. (2011) have recently proposed to model the arrival stream at airports with a Pre-Scheduled Random Arrivals (PSRA) process, which is obtained from a deterministic schedule by superimposing Independent Identically Distributed (IID) random delays. The list of actual arrival times is the result of mixing-up the fixed schedule by the addition of random perturbations. The resulting process –which has been known since the

60's (Kendall, 1964)– was able to provide a very good fit for the simulated congestion levels for arrivals at London Heathrow airport (Caccavale et al., 2014). Further, the PSRA process is easy to study numerically, and some significant analytical results have recently been achieved by Lancia et al. (2018). Nikoleris & Hansen (2012) used PSRA to develop a single-server queuing model for trajectory-based aircraft operations which accounts explicitly for varying levels of imprecision in meeting prescribed times of arrival at either a point in the airspace or a runway's threshold. With the purpose of gaining insight into the generation of the observed delays and balancing congestion delays more efficiently between ground and en-route, Gwiggner & Nagaoka (2014) compared two single-server queuing models ($\cdot/D/1$ and $\cdot/G/1$ in Kendall's notation) using both Poisson and PSRA as arrival processes. Using analysis of the east-bound arrivals at Tokyo International Airport, they concluded that PSRA and a Poisson stream behave equivalently during moderate congestion but differ substantially during very high congestion. However, this comparison is based on the output of queuing models. Analysis of radar data gave arguments both in favor and against the hypothesis of Poisson arrivals.

In contrast to the work of (Caccavale et al., 2014) and Gwiggner & Nagaoka (2014), here we focus on the *direct* comparison of in-homogeneous Poisson arrivals and PSRA, rather than studying the output of a queue model. This represents a fundamental difference and a strong element of novelty with respect to the existing literature. Further, Gwiggner & Nagaoka (2014) tested radar data against the null hypothesis of exponential inter-arrivals without checking whether arrivals were correlated or not. A thorough analysis of serial correlations in the arrival stream is another important contribution of this paper.

In this paper, we present data-driven models for both PSRA and non-homogeneous Poisson and compare their performances in predicting future demand. Shifting the focus to demand-prediction accuracy offers more appropriate metrics for comparing those models in a stochastic optimization framework. An important element of novelty introduced by this work is the use of a regression model for PSRA delays ξ_i (see (3) below) instead of a parametric distribu-

tion (Ball et al., 2001; Guadagni et al., 2011; Nikoleris & Hansen, 2012). The use of a regression model allows the modeling of flight delays and enables precise prediction of the demand. The paper introduces elements of novelty also in the derivation of the Poisson process, which is learned from the data using an original combination of online change-point detection and clustering.

We study the arrival process in the period from June 15 to September 15, 2016, at some of the busiest and most congested airports in Europe: London Heathrow (International Air Transport Association (IATA) code: [LHR](#)), London Gatwick ([LGW](#)), Frankfurt am Main ([FRA](#)), Amsterdam Schiphol ([AMS](#)), and Paris C. De Gaulle ([CGD](#)). As we are also interested in the modeling of medium-intensity traffic, we include in the dataset arrivals at three other important airports: Madrid Barajas ([MAD](#)), Rome Fiumicino ([FCO](#)), and Athens International ([ATH](#)).

Inter-arrival data seemingly suggest that the underlying arrival stream is homogeneous Poisson over three time intervals, namely, 08:00–09:30, 12:00–13:30, and 18:00–19:30, local time; these intervals were chosen to capture different operational phases of the airports, especially those hosting hub-and-spoke operations. Nevertheless, using such a process to model the arrival stream with good approximation presents some weaknesses and difficulties, which we describe in detail in the following sections. On the other hand, PSRA combine a simple formulation with good predictive qualities of the inbound arrival stream. The results presented below are relevant to analyses and simulation-based studies of the air traffic system. Indeed, (fast-time) simulation is one of the most common tools used by practitioners and experts of Air Navigation Service Providers and Network Manager to determine fine-tuned control actions to improve the performances of the air traffic system and to alleviate congestion especially at airports. The approach described herein will allow more accurate analysis, and therefore enables better decision-making. As a consequence, it can also contribute to the improvement of the ATM system efficiency: even a small reduction in terms of average ASMA time can have a huge impact in terms of fuel costs, greenhouse emissions, air traffic controllers' workload and safety.

In summary, the contribution of this paper is three-fold. First, we verify that inter-arrivals appear to follow an exponential distribution, translating (part of) the results of Willemain et al. (2004) and (Gwiggner & Nagaoka, 2014) to a European context; next, we explore serial correlations in the arrival stream and show how our findings do not directly support the assumption of (in-homogeneous) Poisson arrivals. Second, we propose a novel data-driven approach to the modeling of the inbound stream, showing all procedural details to define a non-homogeneous Poisson process and PSRA. Third, we compare the processes obtained in this way with respect to the prediction of future demand.

The remainder of this paper is organized as follows. In Section 2, we describe the dataset and the data analysis methodology used for this study. Section 3 presents the main finding of the paper: in Sections 3.1 and 3.2 we study the exponentiality of the inter-arrival times and its modeling consequences; in Section 3.3, we show how to construct a non-homogeneous Poisson and a PSRA process in a data-driven manner; in Section 3.4 we compare these processes with respect to the prediction of future demand. Finally, in Sections 4 and 5 we discuss the results and provide closing comments.

2. Data and Methods

2.1. Data

Inbound flight data were extracted from EUROCONTROL’s Demand Data Repository (DDR) between June 15 and September 15, 2016. In the summer period, cancellations, diversions, rerouting, and temporary closures of runways and airports tend to occur less frequently than in other periods of the year. Thus, choosing this period allows us to compare the proposed models in a situation that ideally corresponds to a baseline scenario from the point of view of the weather-related disruptions. On the other hand, the summer period is arguably more challenging for the modelling of the demand, since demand levels are notoriously higher than in other periods of the year. In this respect, the performance results presented in Section 3.4 should not be considered as

Table 1: Size of inbound sample for each airport considered. Observation period: from June 15 to September 15, 2016.

airport name	IATA code	sample size
Frankfurt am Main International Airport	FRA	58167
London Gatwick Airport	LGW	39746
London Heathrow Airport	LHR	56716
Amsterdam Airport Schiphol	AMS	63279
Madrid Barajas International Airport	MAD	48162
Charles de Gaulle International Airport	CGD	60122
Athens International Airport	ATH	29503
Rome Fiumicino International Airport	FCO	43333

a worst-case benchmark, but rather as results obtained in a favourable setting weather-wise. Nevertheless, since models are built on the same data, comparing their predictive performance is nothing short of fair. Table 1 displays the total count of inbound flights in the study sample for each airport.

We queried the DDR to extract the so-called regulated and actual flight plans. The former is the last flight plan agreed with the NETWORK MANAGER (EUROCONTROL); it can be negotiated until 20 minutes before departure. The latter is the flight trajectory actually flown; it reflects adjustments of Air Traffic Control to the regulated plan. We denote by t^r and t^a the time at which, according to the regulated and the actual flight plan (respectively) the aircraft enters a cylinder of 40 NM (Nautical Miles) around the destination airport. This procedure is in agreement with the computation of the ASMA times (Cappelleras). Indeed, the passage time at 40 NM is a proxy for the time when the flight starts the approach phase and is handed over to the Terminal Control. This time could have been measured with more accuracy by considering the instantaneous latitude and longitude of each aircraft as reported by the DDR, yet the general data analysis methodology and results illustrated hereafter remain valid. Unless

explicitly stated, times are local.

Time-stamps of the passage at 40 NM form a Time Series (TS) for each airport. Once the TS has been created, *inter-arrival times* are defined as the time lapse in seconds between two successive events. As the arrival rate is not constant, this TS has no fixed frequency. Since time-stamps are measured as precisely as the nearest second, the inter-arrival TS generally contains ties, i.e. a set of two (or more) equal values.

2.2. Exponentiality of the inter-arrival times

We investigate evidence of exponentiality in the inter-arrivals through a QQ-plot using theoretical quantiles from the Weibull distribution

$$f_W(x; \lambda, \beta) = \frac{\beta}{\lambda} \left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^\beta}. \quad (1)$$

The use of a Weibull is in line with Willemain et al. (2004) and is preferable over an exponential law because the *shape parameter* β can appreciably modify the probability of observing small inter-arrivals, i.e. a large number of arrivals in a fixed interval, while the chance of observing large inter-arrivals still decays exponentially fast. The presence of ties in the sample is overcome by using the discrete version of (1) (Nakagawa & Osaki, 1975; Barbiero, 2013)

$$P_W(X = x; q, \beta) = q^{x^\beta} - q^{(x+1)^\beta}, \quad (2)$$

where q plays now the role of the *scale parameter* λ in (1). When $\beta = 1$, (1) and (2) become respectively an exponential distribution and a geometric probability mass function. QQ-plots are drawn for three different time intervals, namely, 08:00–09:30, 12:00–13:30, and 18:00–19:30, local time; these are meant to capture different operational phases of the airports considered in this study. We use the Kolmogorov-Smirnov test (Arnold & Emerson, 2011) to evaluate goodness of fit.

2.3. Arrival process: average demand and serial correlations

Typical characteristics of the inbound stream are assessed with an exploratory analysis of both average properties and serial correlations of the demand; the

latter is a key point that can motivate the use of the PSRA process. These characteristics of the demand are investigated by aggregating the arrivals TS by intervals of ten minutes. The reason for choosing an interval of ten minutes is two-fold. On the one hand, it is sufficiently small to capture changes in the regime of the underlying (stochastic) arrival process. On the other, it is sufficiently large to overlook noisy variations of the demand that would challenge the interpretability of the results.

Average demand is estimated per interval of ten minutes, yielding a daily average profile of the demand; typical fluctuations in the daily average are described by 95% point-wise confidence intervals. We look for evidence of serial correlations in the arrival stream by computing the Autocorrelation function (ACF) on the premise that the capacities of both en-route sectors (airspace) and airports impose constraints (dependencies) on the number of arrivals in consecutive time intervals. Stationarity of the arrivals TS is verified by taking first-order differences and then performing the augmented Dickey-Fuller test (Fuller, 2009; Seabold & Perktold, 2010) in a 24-hour window. Further, we explore periodicity of the demand using a continuous wavelet transform based on the Ricker wavelet (Ryan, 1994). We conclude the descriptive analysis of the arrival stream by computing demand correlations over consecutive time intervals.

2.4. Data-driven modeling of the arrival processes

Two models for the inbound stream at airports are presented: a non-homogeneous Poisson process and PSRA. Instead of doing inference on the Poisson intensity $\lambda(t)$ or the distribution of the PSRA delays, we adopt data-driven modeling procedures described in the following subsections. The two models will then be compared and contrasted with respect to their prediction capabilities in Section 3.4.

2.4.1. Construction of the non-homogeneous Poisson process

We approximate the intensity of the in-homogeneous Poisson process with a step-function. The intensities and the corresponding time intervals are com-

puted using first the PELT algorithm (Killick et al., 2012) to detect change-points¹ in the arrival stream, and then DBSCAN (Ester et al., 1996; Pedregosa et al., 2011) to cluster change-points and estimated intensities in the (t, λ) plane; the intensity of the learned Poisson process is obtained by the centroid of these clusters. This results in the formulation of a non-homogeneous Poisson process whose intensity function is daily periodic. The assumption of a 24-hour periodic process is partly supported by the results in Figure 5, but it is also a simplifying one. Airline schedules often change between weekdays and weekends. However, a relatively simple formulation is an appealing feature if one intends to use and/or derive analytical results (e.g. by using this arrival process in a queue model).

The use of PELT is appealing for its low computational cost compared to other change-point detection methods² (Killick et al., 2012). DBSCAN is appropriate because it works with areas of high/low density points, which is exactly the situation depicted in Figure A.2. Both the behavior and performance of PELT strongly depend on the settings of the penalty function. Thus, Appendix A in the supplementary material offers a sensitivity analysis using penalties based on Akaike Information Criterion (AIC) (Akaike, 1998) (our choice) and Modified Bayesian Information Criterion (MBIC) (Zhang & Siegmund, 2007) (the R package default). The details of the procedure are described by Algorithm 1 below.

2.4.2. Construction of the PSRA process

For PSRA, we simulate the process

$$t_i = t_i^r + \xi_i, \quad (3)$$

¹Time points where either the mean or the variance of the arrival TS undergoes a *structural* change.

²As a matter of fact, Functional Pruning Optimal Partitioning (Maidstone et al., 2017) might perform even better since the intensity of the process is the only parameter that is going to change under the null. However, PELT usage is established and its code well-documented.

Algorithm 1 Configuration of data-driven non-homogeneous Poisson process

- 1: $B \leftarrow$ time series of binned arrivals
 - 2: $H_0 \leftarrow$ ‘Poisson’
 - 3: $T, L \leftarrow \text{PELT}(B, H_0)$
 - 4: $n, C \leftarrow \text{DBSCAN}(T, L)$
 - 5: $\{C$ is a list of length n that contains 2-dimensional lists of clustered t, λ -points}
 - 6: $\hat{T} \leftarrow []$
 - 7: $\hat{L} \leftarrow []$
 - 8: **for** $k = 1$ to n **do**
 - 9: $\hat{t}_k, \hat{\lambda}_k \leftarrow \text{centroid}(C[i])$
 - 10: Append \hat{t}_k to \hat{T}
 - 11: Append $\hat{\lambda}_k$ to \hat{L}
 - 12: **end for**
 - 13: $\lambda^*(t) \leftarrow \text{step}(\hat{T}, \hat{L})$ {A step function taking on value $\hat{\lambda}_i$ for $\hat{t}_i \leq t < \hat{t}_{i+1}$ }
 - 14: **return** $\lambda^*(t)$
-

where t_i^r is the regulated arrival time at 40 NM, ξ_i is a delay that is generated from a regression model for $\{t_i^a - t_i^r\}_i$, and t_i^a is the observed arrival time at 40 NM. Since the regulated flight plan can change up to 20 minutes before take-off, model (3) is tailored for tactical predictions. However, the regulated flight plan undergoes a number of regulation rounds. This means that model (3) could be used any time before take-off with the most recent regulations available – with an expectable predictive-performance reduction compared to the observed arrival time.

The regression model for the delays ξ_i 's is obtained by training an ε -support vector machine (Cristianini & Shawe-Taylor, 2000) on the following features: flight origin (national, continental, or intercontinental), arrival time according to the regulated flight plan, and day of the week. The last two features are encoded as two-dimensional cyclic variables using sine/cosine transformations.

Cross-validation is used both for tuning the hyper-parameters of the support vector machine and for assessing the model performance. The recommendation in this case is to use two rounds of cross-validation (Cawley & Talbot, 2010), respectively called inner and outer, on different train-test splits of the dataset. Since we are dealing with a TS, we split the dataset into k parts, labeled $j = 1, \dots, k$; then, for $j = 2, \dots, k$, parts $1, \dots, j - 1$ are used as a training set to predict part j . An inner cross-validation with $k = 6$ each is used to search for the optimal values of the parameters C and ε on a logarithmic grid: ε defines a margin of tolerance within which no penalty (C) is associated in the hinge loss function (Rosasco et al., 2004). The limited number of features used and the resulting impact on this research are discussed in Section 4. The simulation procedure is outlined in Algorithm 2.

2.5. Prediction capabilities of data-driven Poisson and PSRA

We compare the PSRA and the Poisson process defined in Sections 2.4.1 and 2.4.2 with respect to their capabilities for predicting future demand. Due to the formulation of the two processes, predictions can be cast for the Poisson process at any point in time, but predictions for the PSRA are possible after

Algorithm 2 Simulation of data-driven PSRA process

```
1:  $T^r \leftarrow$  sequence of regulated arrival times at 40 NM
2:  $T \leftarrow$  empty list
3:  $X \leftarrow$  matrix of features
4:  $\Xi \leftarrow$  epsilon-support vector machine
5:  $C, \varepsilon \leftarrow$  optimal hyper-parameters via inner cross validation
6: for each  $t$  in  $T^r$  and each row  $x$  in  $X$  do
7:    $\xi \leftarrow \Xi(x)$ 
8:   Append  $t + \xi$  to  $T$ 
9: end for
10: sort  $T$ 
11: return  $T$ 
```

the regulated flight plan becomes known (as discussed in Section 2.4.2). The comparison is made on two predictions tasks, namely:

1. predicting the aggregated demand for each 10-minute interval on the last day of the dataset (September 14);
2. predicting the mean aggregated demand for each 10-minute interval averaged over the last week of the dataset (September 5-11).

Both tasks yield a vector of 144 predictions (one for each 10-minute interval), upon which we compute the following scores: Mean Absolute Error (MAE), Mean Squared Error (MSE), and r^2 . An outer cross-validation with $k = 12$ is used to compare true demand with that predicted by model (3). Cross-validation of the data-driven Poisson model is not meaningful, because the method is based on clustering and not on the prediction of a known target variable (*ground truth*). Evidence supporting this assumption is given in Section 3.2. Further, this assumption has the desirable property of simplifying the model formulation.

All figures and statistical analyses were produced using `Python v.3.6` and `R v.3.3` (via `rpy2`). The code used for generating the analyses is freely available at the address <http://tiny.cc/yznu8y>, a repository maintained by one of the authors (C. Lancia).

3. Results

In this section, we present the results of our analysis using the dataset and the methods described in Section 2.

3.1. Exponentiality of the inter-arrival times

Figures 1-3 show, for each of the eight airports and three time intervals considered, the QQ-plot of the inter-arrivals against the corresponding fitted Weibull (2). Regardless of the time interval, there is quite a good accordance between empirical and theoretical quantiles in the bulk of the distribution. This can be observed as a general flat adherence of the QQ-plot onto the 45-degree dotted line and should be interpreted as the capability of Weibull inter-arrivals to describe small-to-moderate inter-arrival times, i.e. situations of high demand. However, the goodness of the fit deteriorates on the tails and it typically shows over-dispersion, which can be severe at [FRA](#), [AMS](#), [MAD](#), and [CGD](#). A remarkable exception is [LHR](#), for which the demand fluctuates around the value of 40 aircraft/hour (corresponding to an average inter-arrival of 90 seconds). Accordingly, [LHR](#) exhibits the smallest degree of over-dispersion on the tails among the airports considered in this paper.

Table 2 reports the parameters q and β , the mean of the fitted distribution, the Kolmogorov-Smirnov D -statistic, and the p -value of the corresponding goodness-of-fit test for each time interval and airport considered. The fitted shape-parameter β is always fairly close to 1, meaning that the fitted Weibull looks like an exponential/geometrical. Very often we can reject the null hypothesis of Weibull inter-arrivals at the 1% significance level. This is due to the large size of the sample considered, which makes the test very powerful even against small deviations from the expected behavior.

3.2. Arrival process: average demand and serial correlations

Figure 4 shows the daily mean profile of the demand along its 95% point-wise confidence band. Most of the airports analyzed in this paper show the

Table 2: Parameters and goodness of fit of the fitted discrete Weibull (2). The parameter β is measured in seconds. The interpretation of the p -values of the goodness-of-fit test is not straightforward due to the large sample size.

		q	β	mean	D -stat.	p -value
time	airport					
08:00–09:30	FRA	0.992	1.046	93.131	0.020	0.03
	LGW	0.998	1.127	214.506	0.032	0.02
	LHR	0.995	1.135	103.273	0.024	<0.01
	AMS	0.993	1.118	82.209	0.033	<0.01
	MAD	0.996	1.064	166.540	0.019	0.21
	CGD	0.993	1.143	71.336	0.033	<0.01
	ATH	0.996	1.000	261.533	0.021	0.40
	FCO	0.994	1.060	129.793	0.023	0.03
12:00–13:30	FRA	0.994	1.064	122.668	0.018	0.13
	LGW	0.997	1.166	137.588	0.033	<0.01
	LHR	0.995	1.158	90.262	0.026	<0.01
	AMS	0.994	1.157	74.990	0.035	<0.01
	MAD	0.994	1.041	133.073	0.016	0.29
	CGD	0.996	1.119	127.154	0.027	<0.01
	ATH	0.994	1.000	175.291	0.013	0.74
	FCO	0.992	1.067	88.120	0.021	0.01
18:00–19:30	FRA	0.990	1.040	82.466	0.015	0.12
	LGW	0.997	1.167	135.505	0.034	<0.01
	LHR	0.996	1.199	85.679	0.031	<0.01
	AMS	0.993	1.195	61.895	0.033	<0.01
	MAD	0.995	1.095	116.381	0.013	0.43
	CGD	0.995	1.130	93.981	0.027	<0.01
	ATH	0.996	1.064	189.090	0.013	0.77
	FCO	0.994	1.072	117.987	0.016	0.21

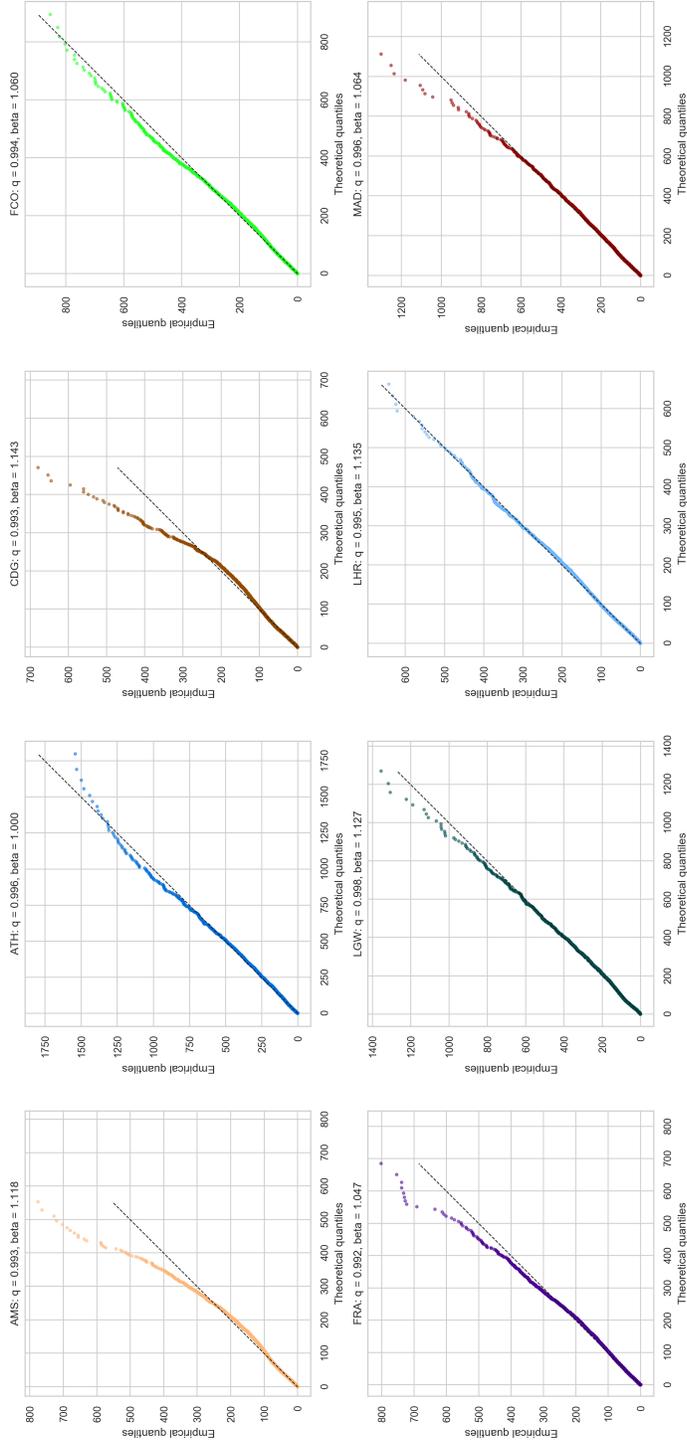


Figure 1: QQ-plot of inter-arrivals at 40 NM in the period 08:00 – 09:30. Theoretical quantiles obtained from (2); parameters listed in Table 2; times are local.

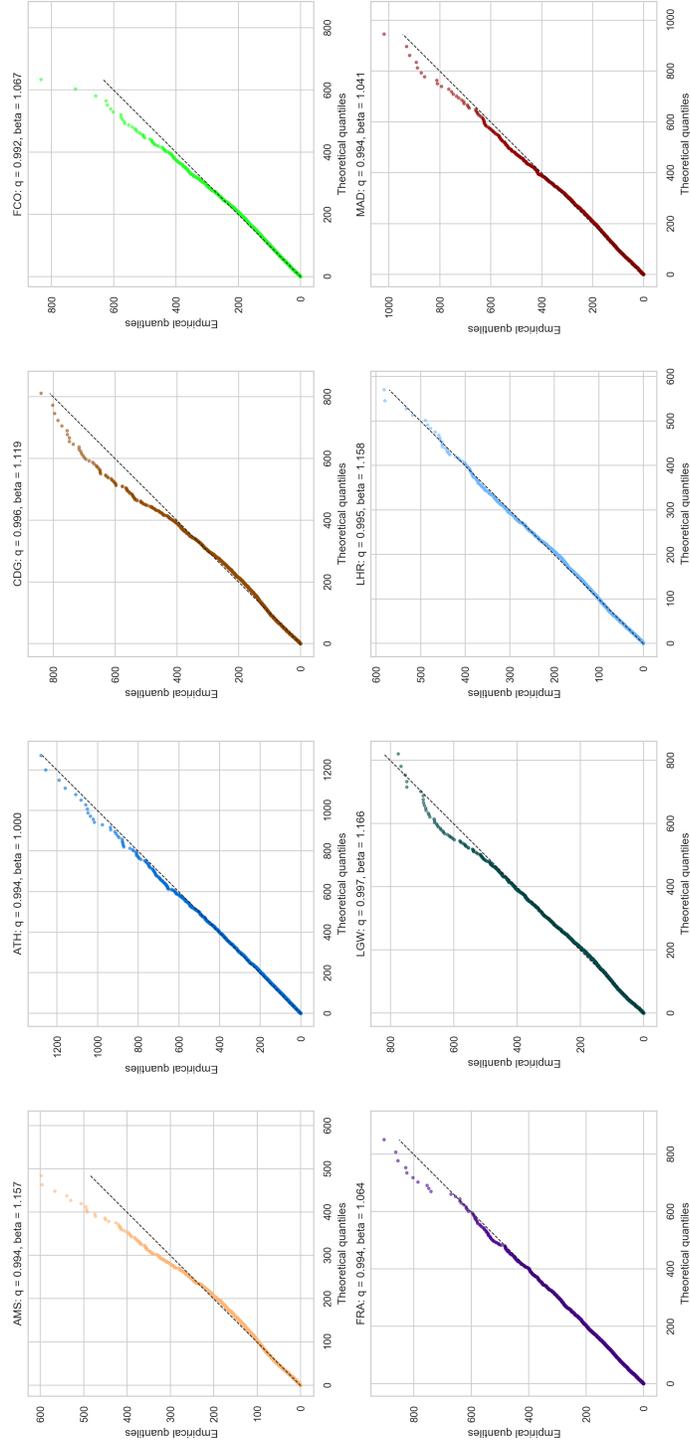


Figure 2: QQ-plot of inter-arrivals at 40 NM in the period 12:00 – 13:30 local time. Theoretical quantiles obtained from (2); parameters listed in Table 2; times are local.

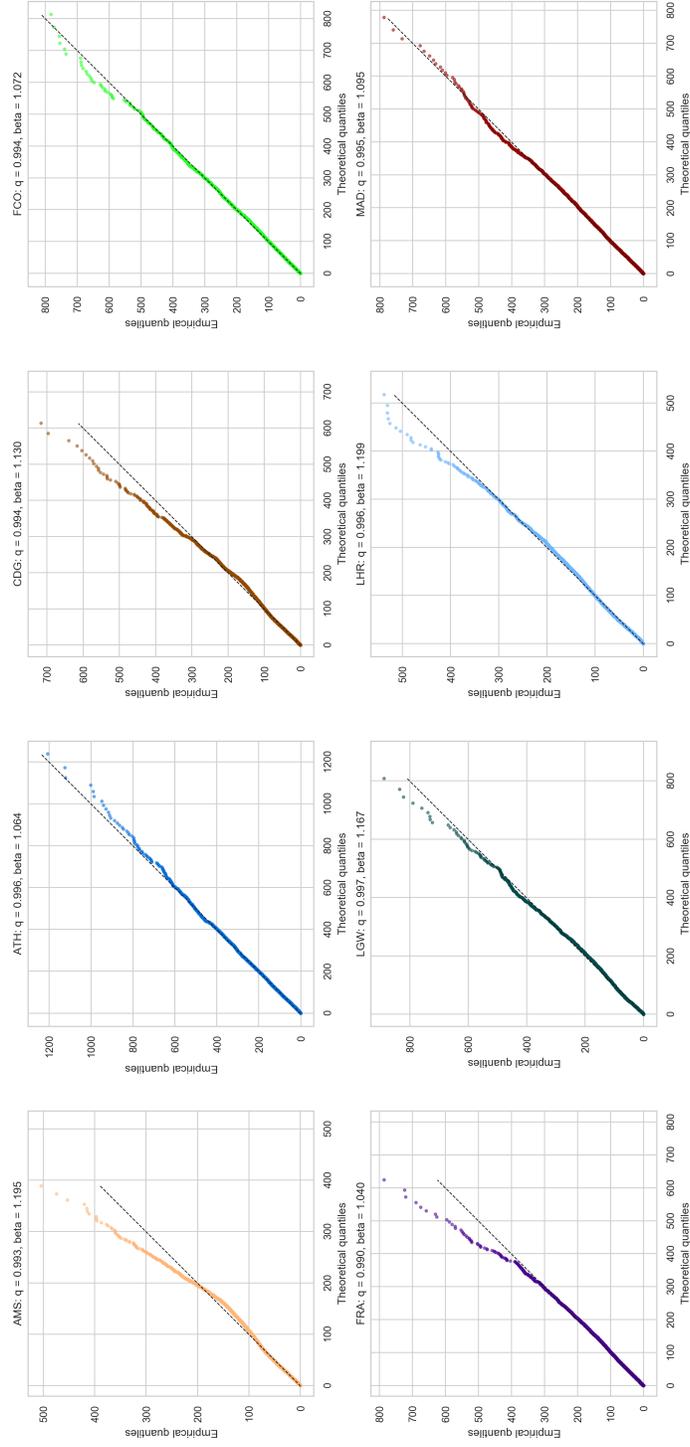


Figure 3: QQ-plot of inter-arrivals at 40 NM in the period 18:00 – 19:30 local time. Theoretical quantiles obtained from (2); parameters listed in Table 2; times are local.

characteristic *wavy pattern* for arrivals, typical of airports hosting hub-and-spoke operations. A notable exception is [LHR](#), which has a fairly constant mean arrival rate and a corresponding arrival process that is characterized by a single regime: the demand oscillates around 7 arrivals every 10 minutes –the declared arrival capacity is 45 aircraft/hour. This also explains why Heathrow shows the best fit with a nearly-exponential distribution in Figures 1–3.

The correlograms in Figure 5 highlight the presence of two kinds of serial correlation in the demand. First, all airports present a statistically significant strong negative lag-1 correlation. These correlations cannot be appreciated in full from Figure 5 because the y -axis is clipped at ± 0.2 for making other correlations more readable. For this reason, their exact numerical values are reported in Table 3. The presence of these correlations suggests that the net variation of the demand³ over an interval of 10 minutes is negatively correlated with the demand variation in the following 10 minutes. In other words, intervals where the demand increases (resp. decreases) are more likely to be followed by intervals where the demand decreases (resp. increases). This property, which can also be guessed from Figure 4, might have an interesting connection with capacity constraints. In fact, if an interval of increased demand were likely followed by another interval of increased demand, then the capacity of the airport could be temporarily exceeded.

Second, many airports show the presence of statistically significant correlations at lags of 1, 2, and 3 days. These correlations are not strong in absolute terms, yet they are the strongest shown by the correlograms, and their relatively low magnitude can be explained by the presence of natural daily variation of the demand evolution in a very large sample. Appendix B in the supplementary material offers a more-in-depth analysis of these serial correlations through a continuous wavelet transform of the demand. This analysis shows that correlations at lag of one or more days are of practical significance. Thus, we have the following:

³The demand TS is made stationary by taking first-order differences, see Section 2.3.

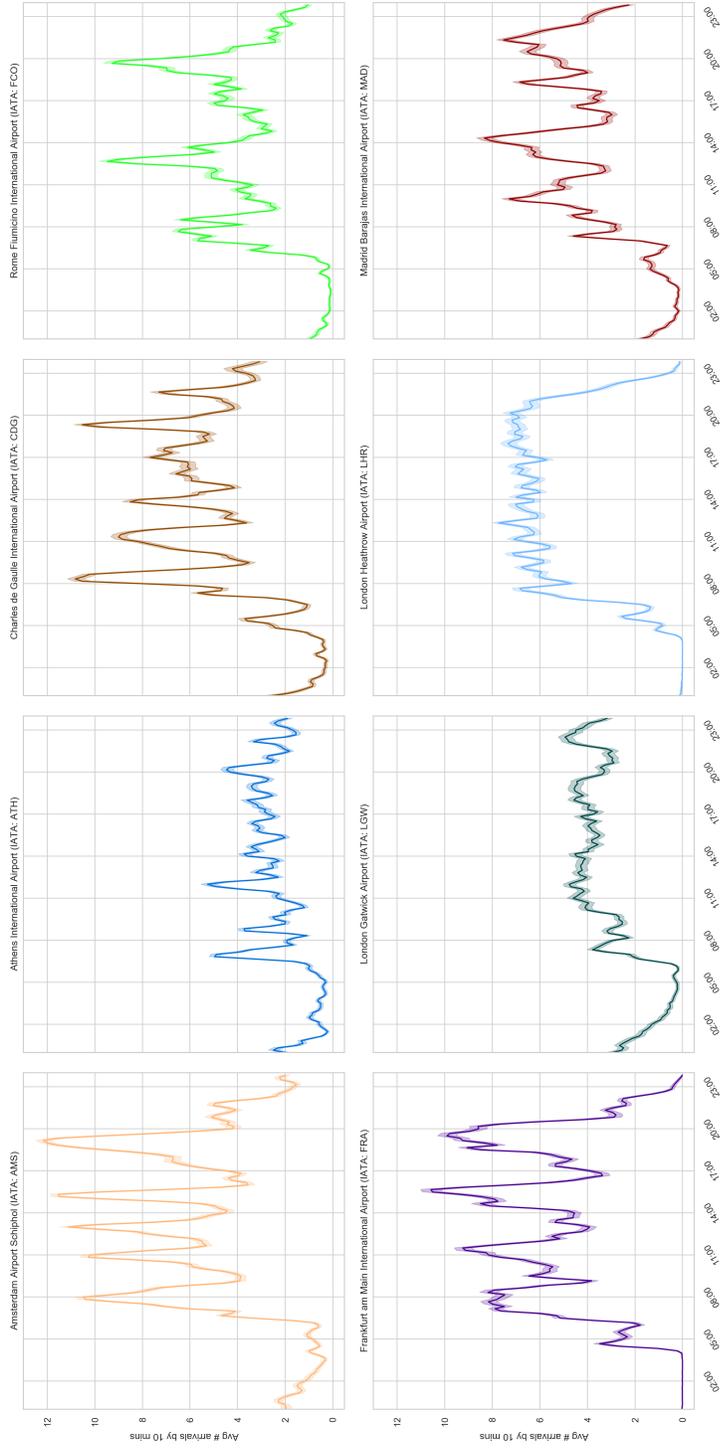


Figure 4: Average demand aggregated by 10 min. Shaded areas display 95% point-wise confidence bands.

Table 3: Values of lag-1 autocorrelations.

IATA code	lag-1 autocorr.
FRA	-0.447
LGW	-0.526
LHR	-0.440
AMS	-0.359
MAD	-0.466
CGD	-0.415
ATH	-0.479
FCO	-0.535

Key Point 1. Significant correlations at lags multiple of one day motivate the idea of learning a daily-periodic non-homogeneous Poisson process in the next section.

Figure 6 shows the Pearson’s correlation coefficients $\rho_{t_i, t_{i+1}}$ between the *demand variation* in the intervals $[t_i, t_{i+1})$ and $[t_{i+1}, t_{i+2})$. The demand variation is computed as the difference between the number of arrivals observed from t_i^a data and the number of arrivals that were expected according to t_i^r data. The values of $\rho_{t_i, t_{i+1}}$ are mostly negative, especially during normal operations hours. The majority of these correlations are different or borderline-different from zero at a 5% significance level. This finding is expected in view of the lag-1 autocorrelations reported by Table 3.

Key Point 2. The signs of the correlations are in line with the general result that PSRA generate negatively autocorrelated streams (Guadagni et al., 2011). Please note that Guadagni et al. (2011) compute correlations from the observed inbound stream, but have an equally-spaced-in-time arrival schedule. In our formulation, pre-scheduled arrivals are not equally spaced, but come from the regulated flight plan. Accordingly, the correct quantity to consider is the demand variation.

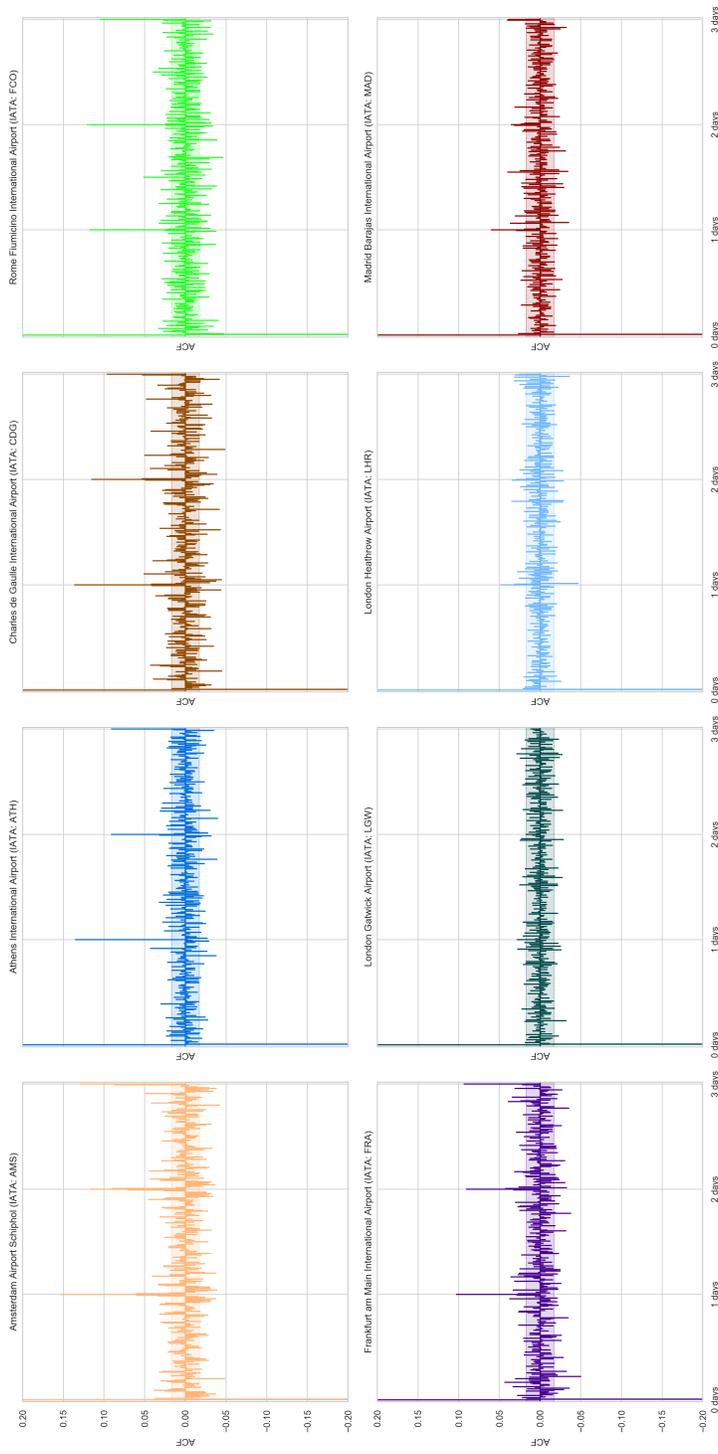


Figure 5: Autocorrelation of arrivals TS aggregated by intervals of 10 minutes. Shared areas mark the limits of statistical significance at the 99% level by Bartlett's Formula. For the sake of readability, the y -axis of the autocorrelation (1st column) is capped at ± 0.2 . The values of the lag-0 and lag-1 correlations are off scale, and the latter are reported by Table 3.

The negative correlations are a very appealing feature of the model as they reflect bounds on the available capacity of both the terminal airspace and the arrival airport. If demand variations were mostly uncorrelated or even positively correlated, random fluctuations in the inbound flow might cause the capacity of airports and/or terminal sectors to be temporarily exceeded. On the contrary, the regulated schedule imposes a structure in the sequence of arrivals (3). For the sake of completeness, correlations computed on simulations of 3 are shown in Appendix C of the supplementary material.

3.3. Data-driven modeling of the arrival processes

Using the data-driven methodology described in Section 2.4, we now detail the parameters characterizing the inbound-stream models, i.e. non-homogeneous Poisson and PSRA, at the considered airport. We begin with the construction of the Poisson process, which is daily-periodic in our formulation. This modeling choice is motivated by Key Point 1 above.

3.3.1. Construction of the non-homogeneous Poisson process

Figure 7 shows the DBSCAN clustering of change-points identified by PELT and the daily average rate of arrivals at 40 NM per 10-minute interval. For each cluster, black diamonds mark the average Poisson intensity $\hat{\lambda}_i$ and the corresponding average time of the day \hat{t}_i , where i is the index of the cluster. In view of the 24-hour periodicity of the demand highlighted in Section 3.2, we define our data-driven in-homogeneous Poisson model by a periodic step-function, which takes on value $\hat{\lambda}_i$ for $\hat{t}_i \leq t < \hat{t}_{i+1}$. The values of \hat{t}_i and $\hat{\lambda}_i$ are reported by Table D.1 in Appendix D of the supplementary material.

Remark 1. Note that the values of $\hat{\lambda}$ are substantially in line with the fitted values from Table 2, since $\lambda \approx 60 \times \text{mean}^{-1}$ in the approximation of exponential inter-arrivals.

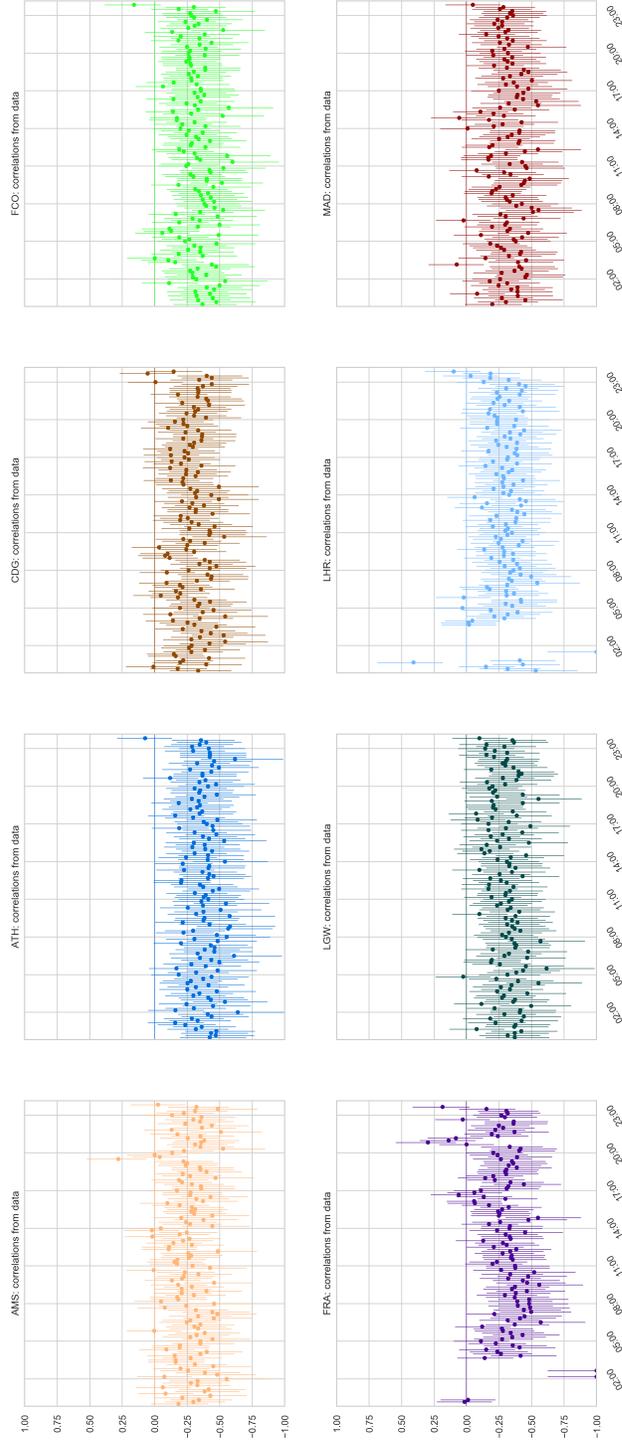


Figure 6: Correlations from t_i^a data. Error bars show 95% confidence intervals for Pearson's ρ .

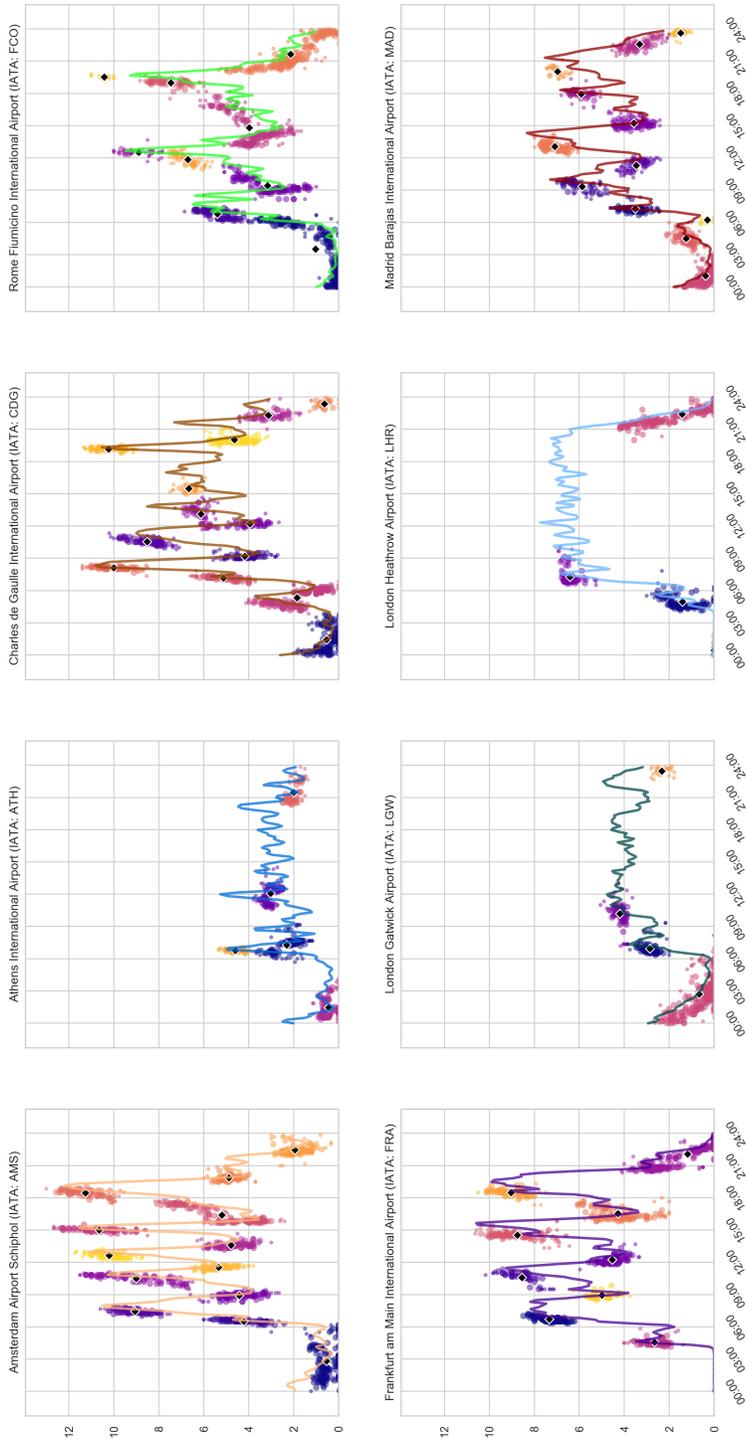


Figure 7: Result of DBSCAN clustering after PELT changepoint analysis. Points belonging to the same clusters are coloured with the same colour. The clustering result is shown superimposed on the daily average from Figure 4.

Table 4: Inner cross-validation of PSRA model (3). The table shows the best values for the parameters of the support vector machine regression model using a sequential two-weekly split, and the corresponding average values of Mean Absolute Error (MAE), Mean Squared Error (MSE), and r^2 using 11-fold consecutive weekly splits.

airport	C	ε	MAE	MSE	r^2
FRA	100.0	100.000	0.055680	0.330237	0.034639
LGW	1000.0	0.001	0.024724	0.088728	0.027090
LHR	100.0	0.100	0.042192	0.257391	0.026009
AMS	10.0	100.000	0.037729	0.296439	0.023622
MAD	100.0	100.000	0.042015	0.205384	0.024280
CGD	100.0	0.010	0.025096	0.147226	0.018576
ATH	1000.0	1.000	0.034228	0.119433	0.042989
FCO	1000.0	1.000	0.024727	0.146581	0.026880

3.3.2. Construction of the PSRA process

Tables 4 and 5 show the results of a nested cross validation of PSRA process (3), where the delays ξ_i are obtained via a regression model. The model is validated on its capability of predicting the aggregated demand and not the delays ξ_i . Based on 5 consecutive two-weekly splits, the inner cross validation estimates the best parameters C and ε of the support vector machine on a logarithmic grid $C = 10^{-2}, \dots, 10^5; \varepsilon = 10^{-5}, \dots, 10^2$. The outer cross validation evaluates demand prediction on an 11-fold cross validation based on consecutive weekly splits. The r^2 metric is often around 0.5, meaning that the model is capturing a significant proportion of the variance of the demand, while the mean absolute error is about 1.5 aircraft/10 minutes.

Key Point 3. It is clear that even a simple regression model like the one used for the delays is performing quite well.

Table 5: Outer cross-validation of PSRA model (3). The table shows average value (μ) and standard deviation (σ) of Mean Absolute Error (MAE), Mean Squared Error (MSE), and r^2 using 11-fold consecutive weekly splits.

airport	MAE (μ)	MAE (σ)	MSE (μ)	MSE (σ)	r^2 (μ)	r^2 (σ)
FRA	1.591	0.071	5.531	0.366	0.594	0.036
LGW	1.436	0.043	3.911	0.204	0.126	0.069
LHR	1.582	0.061	5.414	0.370	0.509	0.033
AMS	1.831	0.078	6.490	0.530	0.530	0.041
MAD	1.563	0.044	4.765	0.281	0.378	0.035
CGD	1.700	0.059	5.477	0.323	0.500	0.033
ATH	1.199	0.040	2.752	0.156	0.141	0.052
FCO	1.356	0.037	3.844	0.143	0.514	0.026

3.4. Prediction capabilities of data-driven Poisson and PSRA

The prediction capabilities of both in-homogeneous Poisson and PSRA are compared by Figures 8 and 9. The first figure shows the average demand prediction obtained for the last week included in the dataset (September 5–11, 2016), while the second figure shows the prediction of the demand for the last day (September 14, 2016). Results are plotted as differences between true and predicted demand (solid: PSRA, dotted: Poisson). Figure 8 clearly shows that PSRA are much more accurate than Poisson in predicting the average future demand. Table 7 compares the models by presenting MAE, MSE, and r^2 scores for the prediction task at each airport. A close look at Figure 9 shows that, for the prediction on a single day, the demand predicted by the Poisson model fluctuates more than PSRA around the true value. Table 7 shows that PSRA still achieve greater predictive accuracy in this task, due to a smaller error and a higher r^2 score.

Difference between true and predicted demand on 2016-09-05/11 aggregated by 10 mins

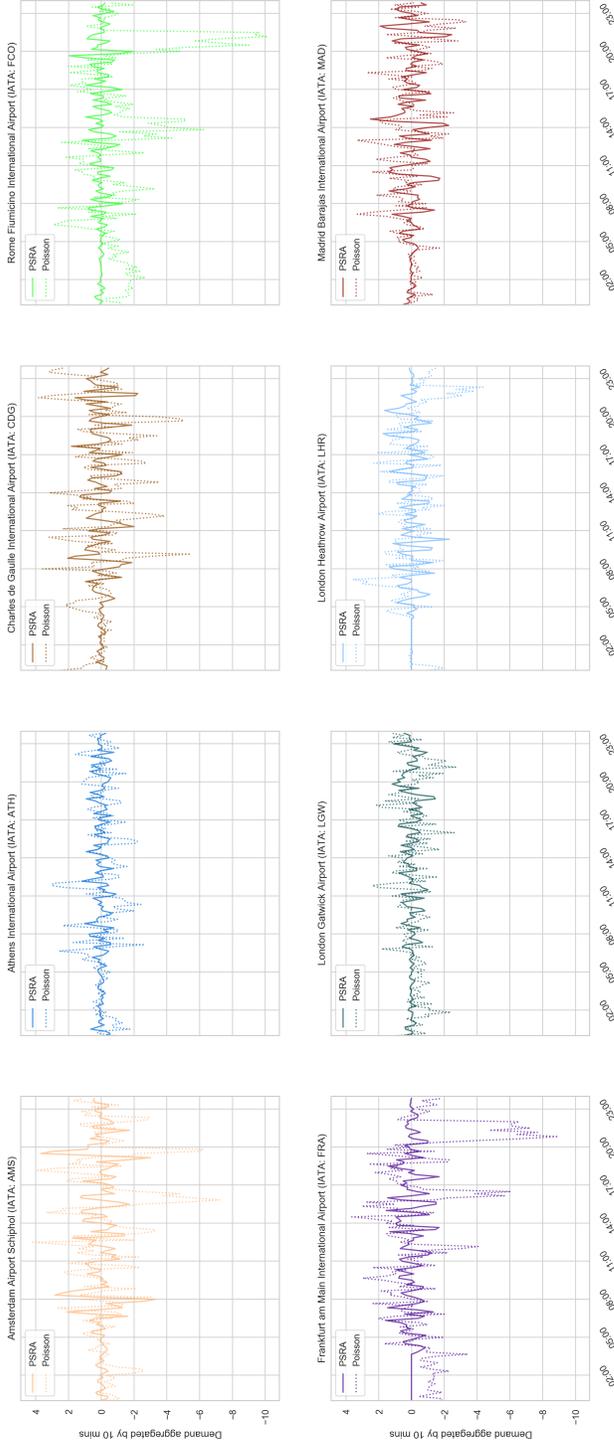


Figure 8: Differences between true and predicted average demand between September 5 and 11, 2016. A solid line shows the prediction of the PSRA while the dotted one shows that of the non-homogeneous Poisson process defined by Table D.1.

Table 6: Scores of prediction tasks for PSRA and the non-homogeneous Poisson model. True and predicted aggregated demand, averaged in the week September 5-11, 2016, are compared using the following scores: Mean Absolute Error (MAE), Mean Squared Error (MSE), and r^2 .

airport	prediction	model	MAE	MSE	r^2
FRA	Sep 5-11	PSRA	0.451	0.398	0.964
		Poisson	1.624	5.199	0.531
LGW	Sep 5-11	PSRA	0.320	0.187	0.920
		Poisson	0.748	0.928	0.602
LHR	Sep 5-11	PSRA	0.395	0.355	0.957
		Poisson	0.959	1.661	0.798
AMS	Sep 5-11	PSRA	0.629	0.875	0.921
		Poisson	1.262	3.635	0.672
MAD	Sep 5-11	PSRA	0.555	0.593	0.894
		Poisson	0.876	1.491	0.733
CGD	Sep 5-11	PSRA	0.520	0.516	0.932
		Poisson	1.199	2.764	0.635
ATH	Sep 5-11	PSRA	0.289	0.143	0.906
		Poisson	0.704	0.869	0.427
FCO	Sep 5-11	PSRA	0.354	0.252	0.956
		Poisson	1.668	6.104	-0.064

Difference between true and predicted demand on 2016-09-14 aggregated by 10 mins

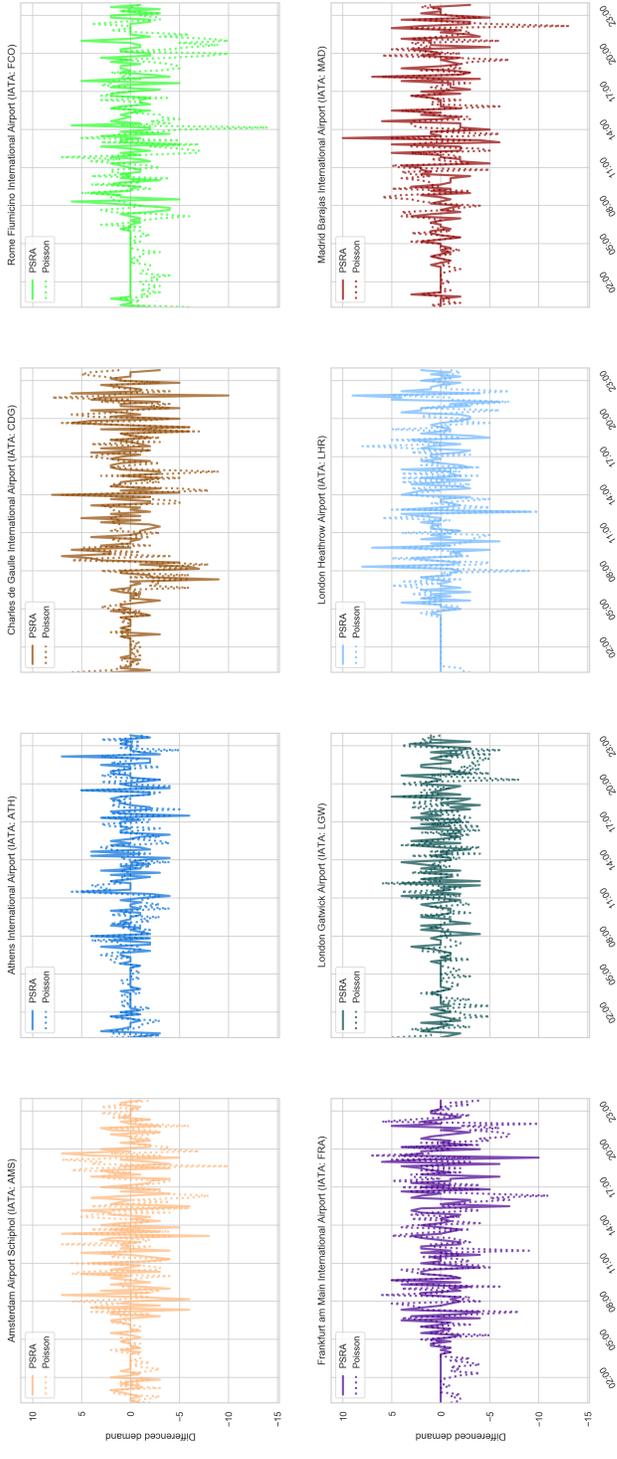


Figure 9: Differences between true and predicted demand on September 14, 2016. A solid line shows the prediction of the PSRA model while the dotted one that of the non-homogeneous Poisson process defined by Table D.1.

Table 7: Scores of prediction tasks for PSRA and the non-homogeneous Poisson model. True and predicted aggregated demand on September 14, 2016 are compared using the following scores: Mean Absolute Error (MAE), Mean Squared Error (MSE), and r^2 .

airport	prediction	model	MAE	MSE	r^2
FRA	Sep 14	PSRA	1.611	5.472	0.597
		Poisson	2.597	11.972	0.119
LGW	Sep 14	PSRA	1.292	3.069	0.286
		Poisson	1.660	5.174	-0.203
LHR	Sep 14	PSRA	1.646	5.979	0.427
		Poisson	1.944	8.236	0.211
AMS	Sep 14	PSRA	1.826	6.493	0.541
		Poisson	2.375	10.347	0.268
MAD	Sep 14	PSRA	1.618	5.688	0.332
		Poisson	2.042	8.611	-0.012
CGD	Sep 14	PSRA	1.958	7.778	0.170
		Poisson	2.785	13.813	-0.475
ATH	Sep 14	PSRA	1.243	3.396	-0.072
		Poisson	1.549	4.771	-0.506
FCO	Sep 14	PSRA	1.222	3.444	0.561
		Poisson	2.222	9.097	-0.159

4. Analysis

Analysis of the inter-arrival times at 40 NM shows a fair accordance between the data and a Weibull distribution; this has already been observed in the literature for *projected* inter-arrivals at some major US airports (Willemain et al., 2004). The shape parameter of the fitted Weibull is sufficiently close to unity to seemingly justify the *classical* assumption of Poisson arrivals⁴. The null hypothesis of Weibull-distributed IID inter-arrivals can often be rejected based on the Kolmogorov-Smirnov goodness-of-fit test, yet this finding requires caution: due to the large sample size, the test is extremely powerful (especially at the most congested airports) and likely to flag even small deviations of the empirical distribution from the theoretical one. The average demand plotted in Figure 4 suggests that the goodness-of-fit might be related to the wavy variation of the demand. On the one hand, QQ-plots in Figures 1-3 show that the best fit is achieved by airports where the demand is stable over time, e.g. [LGW](#) or [LHR](#). On the other hand, the selected time windows often span moments of lower and higher demand, so that a better fit might be obtained by letting β , the shape parameter in (2), vary over time. However, the analysis of the demand correlations in Figure 6 indicates that a non-homogeneous Poisson process is not realistic anyway due to the independence of its increments. In fact, the demand over two consecutive intervals is characterized by non-zero correlations (point estimates mostly negative and often significant at the 5% level), which is not consistent with the assumption of independent increments.

Homogeneous or not, Poisson processes remain a popular choice for modeling inbound air traffic (Gwiggner & Nagaoka, 2014). Thus, we have proposed a novel data-driven approach to the modeling of a non-homogeneous Poisson process by use of PELT (a change-point detection algorithm) and DBSCAN (clustering). We compare the resulting Poisson processes with a PSRA point process, in which the

⁴A necessary and sufficient condition for a random process to be Poisson is that inter-arrival times are *independent* and exponentially distributed.

observed arrival time is given by the sum of the last-agreed arrival time and a random fluctuation. A parametric distribution is the typical modeling approach for the delays ξ_i in (3), but we introduced an element of novelty by adopting also for PSRA a data-driven approach. We used a support vector machine, but other choices, such as ordinary/generalized least squares or other penalized regressions, are viable options worthy of exploration.

The PSRA model presented here was based on only a small number of real-world features. While it is true that a larger number of covariates could be obtained from the DDR, recent constraints imposed to the use of such repository limited the amount of data that we could use. Further, the use of meteorological conditions as a predictor of delay would arguably increase the training performance of the model, but worsen its testing metrics due to the high uncertainty associated with weather forecasts. Tables 6 and 7 show that PSRA often achieves an r^2 score much larger than 0 even with a small number of features, indicating that a substantial part of the demand variance is already captured by the model. *On the premise that the predictive power of our PSRA formulation could only increase if a larger number of features were used, this work validates the idea that PSRA are preferable over a Poisson model for modeling inbound air traffic.*

The predictive accuracy of PSRA over the Poisson process is demonstrated by Figures 8 and 9, in which Poisson arrivals show a greater fluctuation of the predicted demand around the true value. As a consequence, using such an arrival process in a queue model would lead to overestimating the congestion. As reported by Caccavale et al. (2014), such overestimation can be substantial.

One could argue that the accuracy of the Poisson process could be increased by modeling the intensity of the process on a finer time scale, i.e. by estimating it in small prescribed intervals. It is well-known that the Maximum Likelihood Estimator (MLE) of the parameter λ is the sample mean. Thus, a Poisson process with $\lambda(t)$ forced to vary every 10 minutes would exactly reproduce the daily average aggregated demand. However, it is not obvious that such a highly-parametric model would score much better than the one presented here.

Additionally, it would fail to capture the correlation structure of the arrival data because a non-homogeneous Poisson process has independent increments regardless of the functional form of the intensity $\lambda(t)$. Conversely, our PSRA model preserves the correct correlation structure of the demand by inheriting it from the regulated flight plan; see Appendix C in the supplementary material. In addition, PSRA have the general property of converging to a Poisson stream in the regime of delays with very large standard deviation (Guadagni et al., 2011).

5. Conclusions

In this paper, we have provided a thorough analysis of inbound air traffic at eight European airports. We developed two data-driven models for the prediction of the demand, one in the family of Poisson arrivals and the other in the family of PSRA. In all of the airports considered, the PSRA process provides better predictions for the arrival stream. The superior predictive power of the proposed PSRA model over the proposed non-homogeneous Poisson process is not a reason *per se* to recommend one model or the other. Yet, given the correlation of the demand over consecutive intervals, it is clear that a Poisson process is not a good candidate for the description of the inbound traffic flow in Europe. Their use for the estimation of inbound congestion might result in pronounced overestimation, as discussed in the introduction; conversely, PSRA are correlated, provide more accurate predictions, and can be viewed as a generalization of Poisson arrivals from a mathematical point of view. For all these reasons, we recommend the consideration of PSRA as a modelling option because they are preferable for many reasons in a European context.

The proposed approach should be envisioned in simulation-based analysis of air traffic management initiatives at either the strategic or the planning phase. More specifically, it can be used as an engine to support and assess flight schedule development and strategic slot allocation schemes. The analysis can target either a single airport and the implications for its operations or the ATM net-

work and its overall performance. In the latter case, the proposed model will be one of the atomic components of a larger embedded simulation model. Over a longer time horizon, the proposed model can be used to evaluate long-term growth initiatives such as airport expansions. Several airports around the world, like London Heathrow and Rome Fiumicino in Europe, are planning to increase their capacity. Thus, it would be advisable to conduct accurate studies on the possible gains in terms of available capacity and performance of the system.

The proposed model for inbound air traffic demand also offers the potential to fine-tune Traffic Management Initiatives on a shorter time scale such as Ground Delay Programs. However, in this case it might be advisable to extend the predictive models to include weather features, as in Liu et al. (2017) and Gopalakrishnan & Balakrishnan (2017).

References

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer.
- Arnold, T. B., & Emerson, J. W. (2011). Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *The R Journal*, *3*, 34–39. URL: http://journal.r-project.org/archive/2011-2/RJournal_2011-2_Arnold+Emerson.pdf.
- Ball, M., Vossen, T., & Hoffman, R. (2001). Analysis of demand uncertainty effects in ground delay programs. In *4th USA/Europe Air Traffic Management R&D Seminar* (pp. 51–60).
- Barbiero, A. (2013). *DiscreteWeibull: Discrete Weibull distribution*. URL: <http://CRAN.R-project.org/package=DiscreteWeibull> r package version 1.0.
- Blumstein, A. (1959). The landing capacity of a runway. *Operations Research*, *7*, 752–763.
- Bookbinder, J. H. (1986). Multiple Queues of Aircraft Under Time-Dependent Conditions. *INFOR*, *24*, 280–288.
- Caccavale, M. V., Iovanella, A., Lancia, C., Lulli, G., & Scoppola, B. (2014). A model of inbound air traffic: The application to Heathrow airport. *Journal of Air Transport Management*, *34*, 116–122.
- Cappelleras, L. (). Performance Indicator – Additional ASMA Time. Available at http://ansperformance.eu/references/methodology/additional_asma_time_pi.html.
- Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, *11*, 2079–2107.

- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X. et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd* (pp. 226–231). volume 96.
- EUROCONTROL (2016a). *CODA DIGEST Q2 2016*. Technical Report Eurocontrol. Available at <http://www.eurocontrol.int/sites/default/files/content/documents/official-documents/facts-and-figures/coda-reports/digest-q2-2016-final.pdf>.
- EUROCONTROL (2016b). *Performance Review Report on European Air Traffic Management Performance in 2015*. Technical Report Eurocontrol. Available at www.eurocontrol.int/documents/performance-review-report-european-air-traffic-management-performance-2015.
- EUROCONTROL (2018). Performance review report 2017. Available at <https://www.eurocontrol.int/sites/default/files/publication/files/prr-2017.pdf>.
- EUROCONTROL, & FAA (2015). *Comparison of Air Traffic Management-Related Operational Performance: U.S./Europe*. Technical Report Federal Aviation Administration Available at https://www.faa.gov/air_traffic/publications/media/us_eu_comparison_2015.pdf.
- FlightStats (2017). www.flightstats.com [Accessed: June, 2017].
- Fuller, W. A. (2009). *Introduction to statistical time series* volume 428. John Wiley & Sons.
- Gopalakrishnan, K., & Balakrishnan, H. (2017). A comparative analysis of models for predicting delays in air traffic networks. In *12th USA/Europe Air Traffic Management R&D Seminar* (pp. 1–10).

- Guadagni, G., Ndreca, S., & Scoppola, B. (2011). Queueing systems with pre-scheduled random arrivals. *Mathematical Methods of Operations Research*, *73*, 1–18.
- Gwiggner, C., & Nagaoka, S. (2014). Data and queueing analysis of a japanese air-traffic flow. *European Journal of Operational Research*, *235*, 265–275.
- Hengsbach, G., & Odoni, A. R. (1975). *Time dependent estimates of delays and delay costs at major airports*. Technical Report Cambridge, Mass.: MIT, Dept. of Aeronautics & Astronautics, Flight Transportation Laboratory, 1975.
- Kendall, D. G. (1964). Some recent work and further problems in the theory of queues. *Theory of Probability & Its Applications*, *9*, 1–13.
- Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of change-points with a linear computational cost. *Journal of the American Statistical Association*, *107*, 1590–1598.
- Koopman, B. O. (1972). Air-Terminal Queues under Time-Dependent Conditions. *Operations Research*, (pp. 1089–1114).
- Lancia, C., Guadagni, G., Ndreca, S., & Scoppola, B. (2018). Asymptotics for the late arrivals problem. *Mathematical Methods of Operations Research*, . URL: <https://doi.org/10.1007/s00186-018-0643-3>. doi:10.1007/s00186-018-0643-3.
- Liu, Y., M. Hansen, D. Z., Liu, Y., & Pozdnukhov, A. (2017). Modeling ground delay program incidence using convective and local weather information. In *12th USA/Europe Air Traffic Management R&D Seminar* (pp. 1–10).
- Maidstone, R., Hocking, T., Rigail, G., & Fearnhead, P. (2017). On optimal multiple changepoint algorithms for large data. *Statistics and Computing*, *27*, 519–533. URL: <https://doi.org/10.1007/s11222-016-9636-3>. doi:10.1007/s11222-016-9636-3.

- Nakagawa, T., & Osaki, S. (1975). The discrete Weibull distribution. *IEEE Transactions on Reliability*, *24*, 300–301.
- de Neufville, R., & Odoni, A. (2003). *Airport Systems: Planning, Design and Management*. McGraw-Hill.
- Nikoleris, T., & Hansen, M. (2012). Queueing models for trajectory-based aircraft operations. *Transportation Science*, *46*, 501–511.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.
- Rosasco, L., Vito, E. D., Caponnetto, A., Piana, M., & Verri, A. (2004). Are loss functions all the same? *Neural Computation*, *16*, 1063–1076.
- Ryan, H. (1994). Ricker, Ormsby, Klander, Butterworth – A Choice of Wavelets. Available at <http://74.3.176.63/publications/recorder/1994/09sep/sep94-choice-of-wavelets.pdf>.
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with Python. In *Proceedings of the 9th Python in Science Conference* (p. 61). volume 57.
- Willemain, T. R., Fan, H., & Ma, H. (2004). Statistical analysis of intervals between projected airport arrivals. *Rensselaer Polytechnic Inst., DSES Tech. Rept*, (pp. 38–04).
- Zhang, N. R., & Siegmund, D. O. (2007). A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data. *Biometrics*, *63*, 22–32.