XXI SITRAER
2024 | Air Transportation Symposium

OPTIMIZING AIRPORT DISEMBARKATION PROCESSES: A NOVEL APPROACH USING APPROXIMATE BAYESIAN COMPUTATION

João Paulo Silva Lima[a], Michelle Carvalho Galvão da Silva Pinto Bandeira[a], Viviane Adriano Falcão[b]
[a] Federal University of Goiás (UFG), Faculty of Science and Technology (FCT), Laboratory of Intelligence and Innovation in Transportation (LI2T)
[b] Federal University of Pernambuco (UFPE), Department of Civil and Environmental Engineering

*** Corresponding author e-mail address**: joaoplima@ufg.br

PAPER ID: SIT909488

ABSTRACT

This study develops a macrosimulation model for the disembarkation processes at an international airport, utilizing Approximate Bayesian Computation (ABC) to estimate queue model parameters with intractable likelihood functions. The proposed approach employs ABC with weighted local-linear regression, enabling the inference of marginal posterior distributions for unknown parameters such as the mean pedestrian walking speed and the service rate of immigration queuing systems for both local and foreign passengers. The results demonstrate that ABC effectively transformed prior distributions, even with limited data, into reliable posterior distributions with a 95% confidence interval. This methodological advancement is particularly relevant for the operational management of airport terminals, providing valuable support for future decision-making regarding the capacity and efficiency of queuing systems. The originality of this work lies in the application of ABC within an airport queuing context, where traditional methods fail due to system complexity. By focusing on flight characteristics and immigration passing rates, the study addresses a significant gap in the literature, which has thus far lacked robust approaches for modeling complex queuing systems with intractable likelihood functions. Furthermore, this work paves the way for future applications of ABC in other areas of the airport environment, such as boarding and transfer processes, and suggests integrating other methodologies to further enhance the understanding of these systems.

Keywords: Queue problem, Pedestrian Traffic, Approximate Bayesian Computation, Decision Support.

1. INTRODUCTION

Passengers use different handling processes inside an airport terminal building to board or land the aircraft. The quality of these processes, crucially, is not only measured by the time these procedures require but also by the level of comfort experienced by passengers (Rodríguez-Sanz *et al*., 2021). From an operational perspective, these processes integrate diverse elements of the airport's complex system under different operational scenarios associated with the airport's capacity.

The service offered to the passengers is measured by the mean queue length, which is the average number of passengers in a queue at a given time, the mean waiting time, which is the average time a passenger spends in a queue, or the commitment that a certain percentage of passengers wait less than a given threshold, correlating with satisfaction (Alnowibet *et al*., 2022). Therefore, tools are essential to realize macro predictions, such as statistical analysis, regression, time series, neural networks, and data mining (Gu *et al*., 2022).

As an example of prediction tools, Wu *et al*. (2014) propose a modeling framework to capture the flow of passengers and its relationship to the different factors affecting its performance, thus enabling decision-makers Borille *et al*. (2022) presented a generic framework for the assessment of airport passenger terminal circulation facilities using queueing network. In addition, Alnowibet *et al*. (2022) examined an international departure queue system and benchmarked it against the optimum wait time suggested in the literature for future decision support.

However, statistical analyses of complex queue problems have intractable or complicated likelihood functions (Ebert *et al*., 2019). The facility to simulate data and model the queuing problem allows using Likelihood-free inference (LFI), which involves inferring parameter values given observed data and a simulator model. One popular class of LFI methods is Approximate Bayesian Computation (ABC). Previous researchers such as Jiang *et al*. (2018) and Järvenpää & Corander (2023) applied ABC to theoretical queue problems. Ebert et al. (2019) applied ABC with maximum mean discrepancy to a real airport terminal queue problem.

This paper presents an approach for analyzing an airport disembark passenger queueing processes with intractable likelihood parameters. The use of ABC with a weighted Local-linear regression is a promising method, as it allows for the estimation of the marginal posterior distributions of the unknown parameters. This method has the potential to revolutionize future decision supports in the aviation industry.

2. QUEUEING THEORY AND APPROXIMATE BAYESIAN COMPUTATION

2.1 Introduction to Queueing Theory

Queueing theory is a branch of applied mathematics used to predict line behavior (also known as queues). The main characteristics that define the queue system are (*i*) the arrival process *A*, (*ii*) the service process *S*, (*iii*) the number of servers *c*, (*iv*) the capacity of the queue *K*, (*v*) the calling population *N* and, (*vi*) the queue's discipline *D*. These characteristics are usually summarized using the extended Kendall's notation $A/S/c/K/N/D$.

In the Queueing theory, each customer $\boldsymbol{n}=(1,\ 2,\ldots,N)$ arrives at the system at a specific time $\boldsymbol{a}=(a_1,a_2,\ldots a_N)$ in an arrive process *A* and requires time $\boldsymbol{s}=(s_1,s_2,\ldots s_N)$ in the service process *S* with a server $\boldsymbol{c}=(a_1,a_2,\ldots a_c)$. The number of servers *c* may be any number from 1 to ∞. It is usually assumed that servers operate in parallel and are independent. The queue *q* of the system is formed for new customers arriving at the system when all servers are busy. The waiting times in the queue are denoted as $\boldsymbol{t}=(t_1,t_2,\ldots t_N)$, and the system's departure times are $\boldsymbol{d}=(d_1,d_2,\ldots d_N)$. In theory,

$$d_j = a_j + w_j + s_j. \quad (1)$$

However, Equation 1 belies the complexity of the simulation problem since ***a*** and ***w*** are non-trivial functions that depend on the class of queueing system considered. Moreover, queueing models are usually easy to simulate but have no intractable or very difficult likelihoods, requiring likelihood-free methods to infer the models.

2.2 Approximate Bayesian Computation (ABC)

The Approximate Bayesian Computation (ABC) framework a significant part of Bayesian statistics, uses probability as a measurement of uncertainty and considers the model parameters as random variables (Gelman *et al*., 2013). In the Bayesian analysis, a priori knowledge of the parameter vector $\boldsymbol{\theta}$ is captured by the prior distribution, denoted $p(\boldsymbol{\theta})$. Adding new data $\boldsymbol{x}_{\text{obs}}$, the prior distribution is actively updated using Bayes' theorem:

$$p(\boldsymbol{\theta} \mid \boldsymbol{x}_{\text{obs}}) = \frac{p(\boldsymbol{x}_{\text{obs}} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\boldsymbol{x}_{\text{obs}})}, \quad (2)$$

where $p(\boldsymbol{x}_{\text{obs}} \mid \boldsymbol{\theta})$ is the likelihood function and $p(\boldsymbol{\theta} \mid \boldsymbol{x}_{\text{obs}})$ the posterior distribution. The likelihood it is used to transform the prior to the posterior.

Computational frameworks such as Markov chain Monte Carlo (MCMC) rely on evaluating the likelihood function (Csilléry *et al*., 2010). ABC allows approximation of the posterior distribution of unknown model parameters and the posterior probabilities of models using only forward model simulations. The ABC approaches bypass exact likelihood calculations, though usually only in an approximate and computationally intensive, using summary statistics and simulations. Summary statistics $\tilde{\boldsymbol{x}}$ are values calculated from the data to represent the maximum amount of information in the simplest possible form. The ABC relies on the ability to generate synthetic data to fit the model.

An efficient sampling scheme is required for a successful algorithm. ABC has its roots in the rejection algorithm, a simple technique to generate samples from a probability distribution (Gelman *et al*., 2013). A distance to measure the dissimilarity between the observed and synthetic data $\mathcal{D}(\|\cdot,\cdot\|)$. In this problem, the distance $\mathcal{D}$ is chosen to be the Euclidean Distance. Consider some *n* observed data $\boldsymbol{x}_{obs} = (x_1, \ldots x_n)$, summary statistics $\tilde{\boldsymbol{x}} = (x_1, \ldots x_m)$ and assume that the underlying model does not lead to a tractable likelihood function. Sampling from an approximated version of the posterior can be done likelihood-free through acceptance–rejection with a tolerance level ϵ. The posterior distribution is given by:

$$\pi_{\epsilon}(\boldsymbol{\theta} \mid \boldsymbol{x}) \propto \pi(\boldsymbol{\theta}) \int_{\mathbb{R}^t} \mathbb{I}_{\{\mathcal{D}(\boldsymbol{x},\tilde{\boldsymbol{x}})<\epsilon\}}\, p(\tilde{\boldsymbol{x}} \mid \boldsymbol{\theta})\, \mathrm{d}\tilde{\boldsymbol{x}}. \quad (3)$$

Consider:

$$\mathbb{I}_{\{\mathcal{D}(\boldsymbol{x},\tilde{\boldsymbol{x}})<\epsilon\}} = \begin{cases} 1, & \text{if } \mathcal{D}(\boldsymbol{x},\tilde{\boldsymbol{x}}) < \epsilon, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

2.3 Application of ABC in Airport Queueing Problems

In this study, ABC was applied to analyze passenger disembarkation queueing processes in an international airport, where the likelihood functions are intractable. The use of ABC with weighted local-linear regression proved to be a promising method, enabling the estimation of marginal posterior distributions of unknown system parameters.

In the present research, a local linear regression method corrects for the imperfect match between accepted $\boldsymbol{x}^{acc}$ and $\tilde{\boldsymbol{x}}$ (Blum, 2009,2017). Each accepted parameter pair $\left(\boldsymbol{x}_i^{acc}, \tilde{x}_i\right)$ is

associated with a weight $w_i \propto K\left(\left\|\tilde{x}_i - \boldsymbol{x}_i^{acc}\right\|\right)$ where K is a univariate statistical kernel. In the present research, the Epanechnikov kernel is used.

A Local-linear regression model is fitted in the neighborhood of $\boldsymbol{x}^{acc}$ as:

$$x_i = \alpha + \left(\tilde{\boldsymbol{x}} - \boldsymbol{x}^{acc}\right)\beta, \quad i = 1, \ldots, m \tag{5}$$

where α is an intercept, β is a vector of regression coefficients. In the Loc-Linear ABC algorithm the parameters (α,β) are inferred by minimizing the weighted least-squares criterion:

$$\sum_{i=1}^{m}\left\{x_i - \alpha + \left(\tilde{\boldsymbol{x}} - \boldsymbol{x}^{acc}\right)\right\}^2 \boldsymbol{w}. \tag{6}$$

More details to Loc-Linear ABC method can be found in Blum (2009).

3. PASSENGER CIRCULATRION IN AN INTERNATIONAL AIRPORT PROBLEM

Passengers have a clear flow pattern in an airport terminal. Figure 1 provides a schematic of the arrival terminal passenger process, depicting the corresponding functional zones of the airport. The services encountered by passengers between these nodes determine the time spent in various terminal areas, potentially leading to queue formation. Airline passengers are typically required to queue at several steps during their processing at the departure, transfer, and arrival systems.

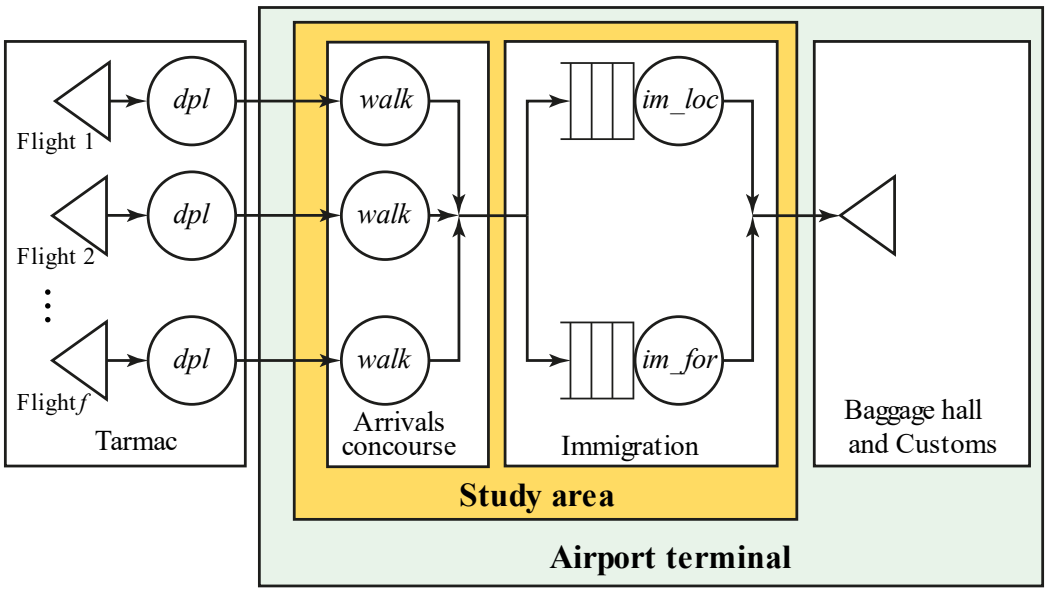


Figure 1 Passenger terminal arrivals process zones diagram. Source: Authors.

The study area of this paper focuses on the arrival terminal of an international airport. In this example, there are three key nodes (Figure 1) in this model, at which the number of passengers passing through per minute (passing rate) are monitored, i.e., (i) the disembarkation subsystem, (ii) the arrivals concourse subsystem, and (iii) the immigration processing subsystem, and subsequently (iv), the baggage hall.

The process described in Figure 1 starts with passengers disembarking from their flight at a designated gate. Each flight has a specific arrival time, and passengers enter the arrivals concourse at a time associated with the deplaning process.

Passengers then proceed from their gate to the immigration subsystem, where they are processed based on specific nationality and eligibility criteria. Each gate has a distinct distance to the immigration area, and the time at which passengers arrive at the queueing system is recorded. After being served by immigration staff, passengers exit the system.

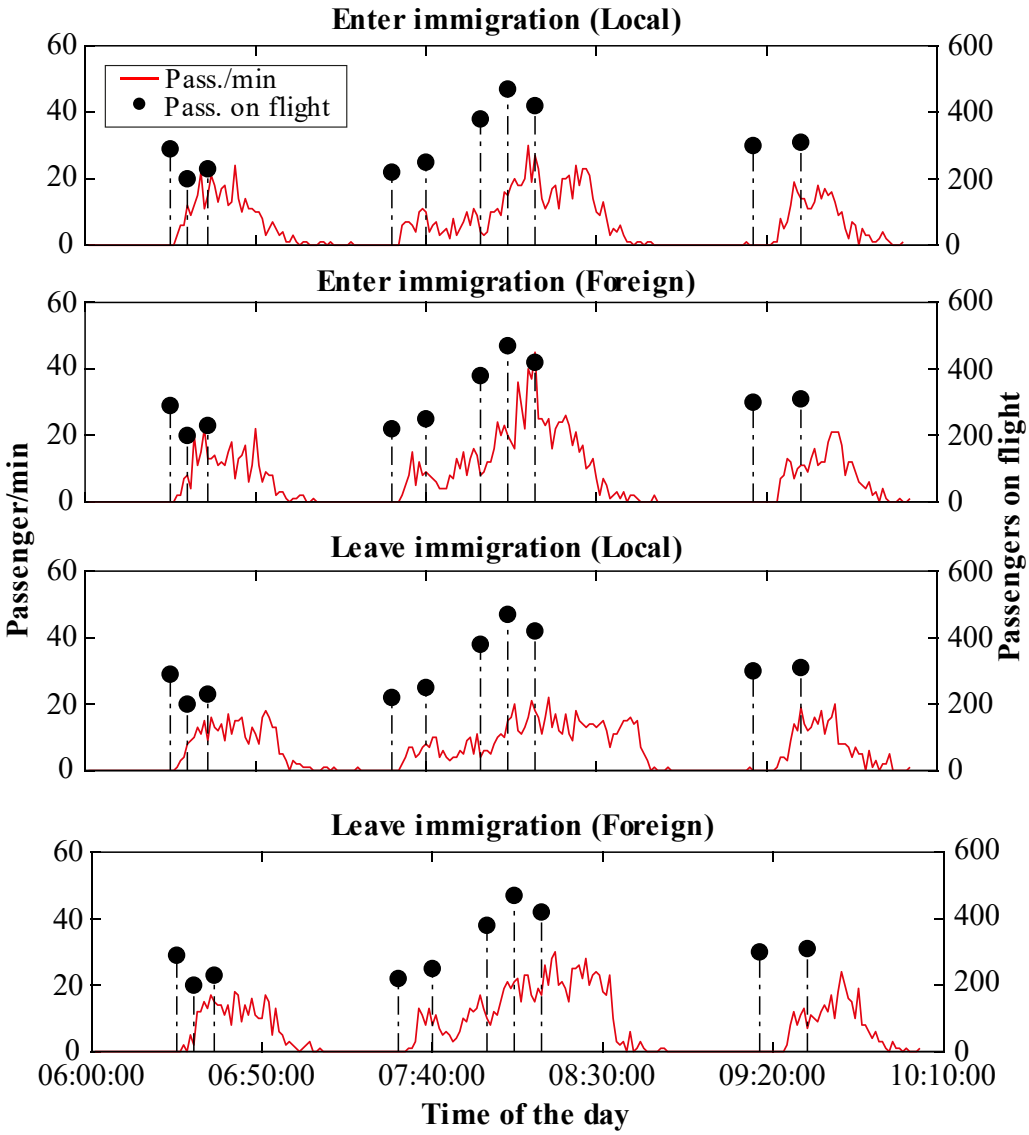


Figure 2 Passing rate in the terminal associated with the number of passengers per flight.

Table 2 Flight information

Flight number	*Arrive time (min)*	*Total passengers*	*Total local passengers*	*Gate*	*Distance gate/immigration (m)*
1	25	290	145	1	30
2	30	200	120	2	90
3	36	230	92	3	280
4	90	220	110	2	90
5	100	250	75	4	440
6	116	380	190	2	90
7	124	470	141	1	30
8	132	420	252	3	280
9	196	300	180	4	440
10	210	310	124	2	90

The synthetic data in Figure 2 consists of the schedule of 10 flights arriving on a specific day. There are 3070 passengers with a passing rate of passengers entering and leaving a set of subsystems within the arrivals terminal. The red continuous line represents the observed passing rate in the day, and the dashed vertical lines are the number of passengers on each flight arranged by flight arrival time. Table 1 summarizes the flights information's.

The services of each passenger between these nodes result from the time passengers spend in the airport terminal areas, and eventually, queues can form. Airline passengers queue up at several steps during passenger processing at the departure, transfer, and arrival systems.

In the arrival process presented in Figure 2, passengers disembark from the arriving flight $\mathbf{f} = \left(d_1, d_2, \ldots, d_f\right)$ at the gate associated with the flight with a total of $\mathbf{p}_u = \left(p_1, p_2, \ldots, p_f\right)$. Each flight has a arrival time $\mathbf{a}_f = \left(a_1, a_2, \ldots a_f\right)$, and each passenger enters the arrivals concourse from their arriving flight at time $a_{f,u}^{dpl} = a_f + t_{f,u}^{dpl}$ associated to the deplane time $t_{f,u}^{dpl}$. The passenger then walks from their gate to the immigration sub-system to be processed at a time $t_{f,u}^{im}$. At this point, passengers take either the local or the foreign route through immigration subject to specific eligibility nationality criteria. Each gate has an access distance to the immigration area. The times at which customers arrive ate the queueing system are $a_{f,u}^{imm} = a_{f,u}^{dpl} + t_{f,u}^{im}$. Immediately after a passenger finishes being served by immigration staff, they are considered to have left the immigration system.

It is assumed that along with the variability between flights, there is also significant variability in arrival profiles daily. Consequently, these data are considered insufficient for disembarkation modeling purposes and are unconsidered in this research.

The travel time of pedestrians traversing airport corridors varied significantly in relation to the walking speed in the corridors. Previous research such as Young (1999) and Lopez-Carmona (2023), revealed that, under free-flow conditions, airport pedestrians' speeds behave approximately normally distributed $v_u \sim N\left(\mu_{ac}, \sigma_{ac}\right)$, with mean μ^{ac} and standard deviation σ^{ac}. About previous research as Järvenpää & Corander (2023) and Jiang *et al*. (2018), the service times $s_{f,u}$ are approximately exponentially distributed $s_{f,u} \left| r_{f,u} \sim Exp\left(\lambda_{rf,u}\right)\right.$ with rate parameter λ_{loc} or λ_{for} depending on the route assignment $r_{f,u}$ based on the nationality. The departure times $a_{f,u}^{imm}$ are computed in a proposed model with each queue follow the *M*/*G*/c, configuration (according to Kendall's notation, *M* for Markovian process and *G* for General distribution).

Based on the previous information, it is not feasible to model this queue problem with a constant passing rate to immigration. The waves of passengers from different flights overlap, and the data only counts passengers/min, not allowing identification of the mix of flights that led to a particular inflow of passengers.

Therefore, the focus shifts to the estimation of the unknown parameters, which remain constant across different scenarios. These parameters play a crucial role in predicting the system's behavior for future flight schedules. The service rate and the pedestrian walking parameters $\lambda_{\text{target}} = \left[\lambda_{loc}, \lambda_{for}, \mu_{ac}, \sigma_{ac}\right]$ are of particular interesse . To create the synthetic scenario dataset, summary statistics is defined as $\lambda_{\text{or}} = \left[1.4, 0.9, 74.0, 14.0\right]$. A total of 15 servers in each queue immigration system is adopted. The simulation model and the virtual gate simulation are realized in the MATLAB software.

4. SIMULATION STUDY

In this section, a simulation study is conducted to investigate the performance of the ABC Loc-linear Method for model selection and parameter estimation of the disembark process in airports. First, a marginal posterior distribution of the unknown parameters is approximated. Then, the actual

values used to generate the synthetic dataset are compared. Finally, the marginal posterior distributions of the target parameters are used to predict queue information of the disembark system.

The ABC-Loclinear algorithm is used to deal with parameter estimation. It is implemented with the following hyper-parameters:

- The tolerance level ϵ=0.01.
- The total of the observed population is set to n=20000.
- The 5-dimensional baseline summary statistics obtained by Figure 2 is the mean of passengers/min $\tilde{\boldsymbol{x}} = (5.91, 7.08, 5.89, 7.05)$.
- Uniform informative priors for the parameters are imposed for the service rate $\lambda_{loc} \sim U(0,3)$, $\lambda_{for} \sim U(0,3)$. According to Alam *et al.* (2022), pedestrians follow a walking speed of 0.9 – 1.5 m/sec, resulting a prior $\mu_{ac} \sim U(54,90)$. In association, is adopted a prior $\sigma_{ac} \sim U(5,25)$.

Figure 3 displays a comparison of the marginal prior and posterior distributions of the selected target model parameters. As depicted, the selected prior distribution to represent the user's knowledge is transformed significantly after completing the Bayesian inference steps . The posterior means (cyan line in Figure 3) with a 95% Confidence Interval are $\lambda_{for} = 1.38[1.26, 1.46]\ min^{-1}$ and $\lambda_{for} = 0.93[0.86, 1.00]\ min^{-1}$ for the rate parameters. For the pedestrian speed the mean and standard deviation parameters are $\mu_{ac} = 73.0[58.0, 86.0]\ m/s$ and $\sigma_{ac} = 13.76[6.29, 21.0]\ m/s$, which are reasonable.

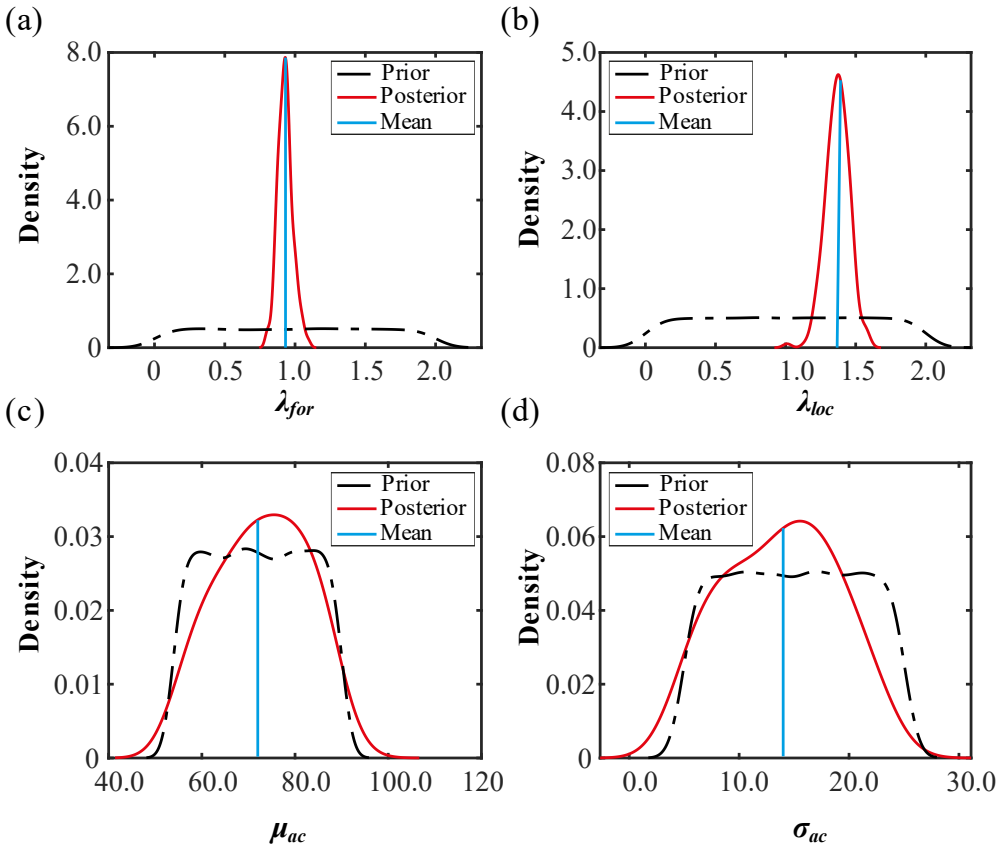


Figure 3 Marginal posterior distributions computed using ABC for parameters λ_{target}

These values can be compared with the original parameters that generate the synthetic dataset. Tabel 2 compares the λ_{target} with the λ_{or}. According to Figure 3 and Table 2, the likelihood-free method provides reasonable estimates overall, showing an ABC Loc-linear estimator to be competitive with a maximum parameter difference of 3.33%. The actual value lies within the 95% CI and is close to the posterior mean. The posterior variances of the service rate parameters are minimal.

Table 2 Comparison between λ_{target} and λ_{or}

Parameter	λ_{target}	λ_{or}	*diff*
λ_{loc}	1.38	1.40	1.42%
λ_{for}	0.93	0.90	3.33%
μ_{ac}	73.10	74.0	1.21%
σ_{ac}	13.76	14.0	1.71%

Prediction intervals of performance measures for decision support tools are easily constructed using draws from ABC posterior samples. By replicating the experiment 1000 times, it is possible to obtain the estimated queue properties of the system. Figure 4 depict the actual passing rate of the immigration process, comparing with the 95% confidence interval predicted with the λ_{target} ABC posteriors. Considering the simulation results, there is a good agreement with the actual and predicted passing rate values over the analyzed time. These simulations also allow to obtain the queue immigration parameters. For the queue immigration the mean queue length in the analyzed period is q_{loc}=24 and q_{for}=26 people with a mean wait time of t_{loc}=1.2 min and t_{for}=1.9 min, for local and foreign queues respectively. Also, it is possible to estimate a 95% CI for the maximum queue length and wait time in the queue. The maximum queue length is $q_{loc,\max} = 115[75,152]$ and $q_{for,\max} = 124[74,174]$ for local and foreign queues respectively. The maximum queue wait time is $t_{loc,\max} = 5.7[3.2,8.0]$ min and $t_{for,\max} = 9.2[5.0,13.5]$ min for local and foreign queues respectively.

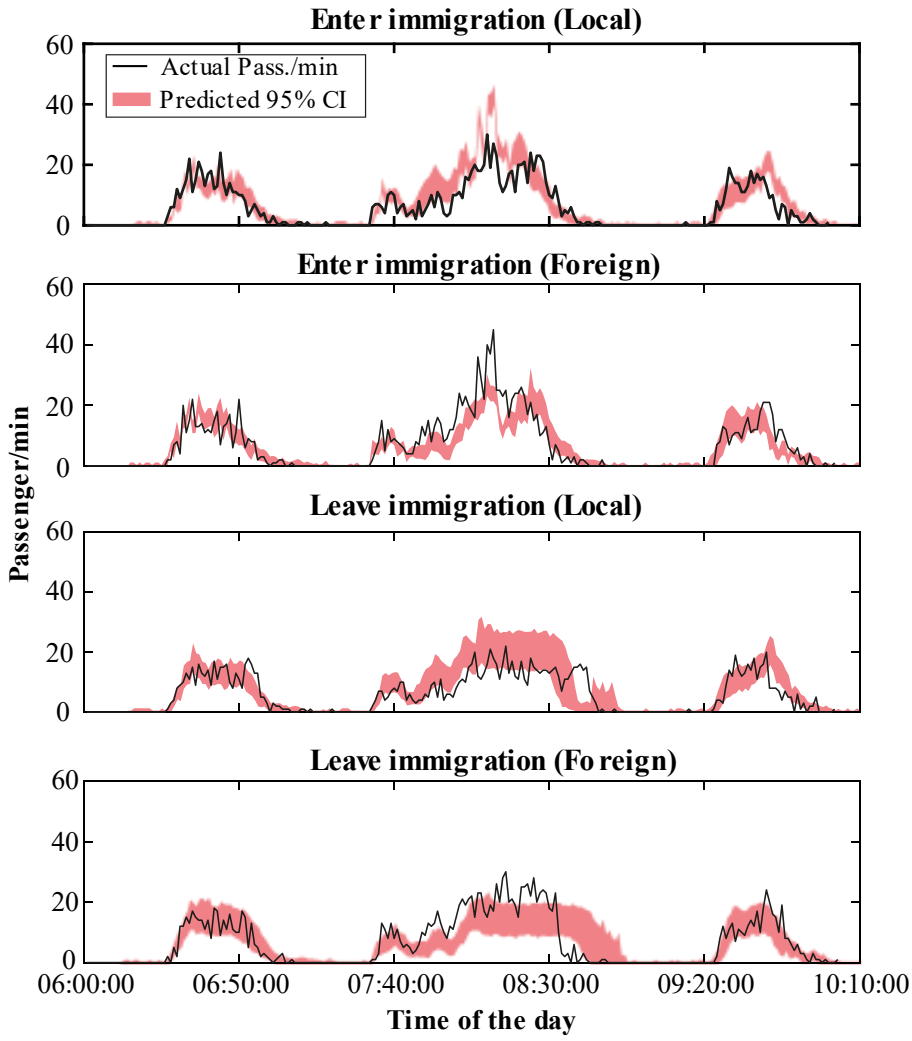


Figure 4 Observed passing rate over the day over local and foreign immigration system.

5. CONCLUSION

The primary objective of this study was to accurately forecast the parameters that directly impact queuing dynamics within the disembarkation terminal of an international airport. However, given the absence of a priori information necessary for traditional statistical inference, the ABC Loc-linear method was introduced. This approach was specifically designed to approximate the posterior predictive distribution of a complex queue model that had previously been deemed intractable.

The proposed framework relied solely on flight characteristics associated with immigration passing rates to generate the ABC summary statistics. Despite the lack of detailed information regarding service fees or the specific formation of queues within the system, the available data was effectively converted into summary statistics. This conversion allowed the transformation of prior distributions—despite their initial lack of precise factual basis—into reliable posterior distributions. These distributions, with a 95% confidence interval, accurately reflected the reality of the previously unknown variables in the proposed model.

The successful estimation of the queue system parameters marks a significant achievement, as it provides a deeper understanding of the factors influencing queue formation and behavior. This understanding is critical for future analyses, as these predictions directly affect the terminal's capacity. By gaining insights into the efficiency and capacity of the queue system, it becomes possible to mitigate inefficiencies, such as inadequate service facilities or an insufficient number of service points, which could otherwise lead to severe bottlenecks and compromise the overall capacity of the terminal.

Looking forward, the potential for applying the ABC framework to other airport processes involving queue formation is highly promising. Additionally, integrating other methodologies alongside ABC could further deepen our understanding of these systems. This not only highlights the fact that there is no single, universally applicable method for approximating the posterior predictive distribution of intractable models using ABC, but it also opens up exciting avenues for future research and practical applications.

6. REFERENCES

Alam, M. D. J., Habib, M. A., & Holmes, D., (2022) Pedestrian movement simulation for an airport considering social distancing strategy. *Transportation Research Interdisciplinary Perspectives, 13*. 100527. doi:https://doi.org/10.1016/j.trip.2021.100527.

Alnowibet, K. A., Khireldin, A., Abdelawwad, M., & Mohamed, A. W., (2022). Airport terminal building capacity evaluation using queuing system. *Alexandria Engineering Journal, 61*(12), 10109-10118. doi: https://doi.org/10.1016/j.aej.2022.03.055.

Blum, M.G.B., & François, O. (2010). Non-linear regression models for Approximate Bayesian Computation. *Stat Comput*, *20*, 63–73 https://doi.org/10.1007/s11222-009-9116-0

Blum, M.G.B., & François, O. (2017). Regression approaches for Approximate Bayesian Computation. *arXiv:Methodology*, *170701254*. doi:https://doi.org/10.48550/arXiv.1707.01254

Borille, G. M. R., Silva, E. J., Nascimento, M. V., Bandeira, M. C. G. S. P., Rendhol, A. L. M., Silva, L. M., Ansélmo, C. S., & Gomes, R. A. (2022). Airport passenger building assessment of circulation facilities for smooth traffic and efficient design. *Simulation Modelling Practice and Theory*,119, 102555. doi:https://doi.org/10.1016/j.simpat.2022.102555.

Csilléry, K., Blum M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, 25(7), 410-418. doi:https://doi.org/10.1016/j.tree.2010.04.001

Ebert, A., Dutta, R., Mengersen, K., Mira, A., Ruggeri, F., & Wu, P. (2021). Likelihood-free parameter estimation for dynamic queueing networks: case study of passenger flow in an international airport terminal. *Journal of the Royal Statistical Society, Series C. 70*(3),770-792. doi: https://doi.org/10.1111/rssc.12487

Gelman A., Carlin J.B., & Stern H.S. (2013). *Bayesian Data Analysis*, (3 ed.). New York, USA: CRC Press.

Gu, X., Xie, J., Huang, C., Ma, K., & Liu, J. (2022). Prediction of the spatiotemporal passenger distribution of a large airport terminal and its impact on energy simulation. *Sustainable Cities and Society, 78*, 103619. doi:https://doi.org/10.1016/j.scs.2021.103619.

Järvenpää, M., & Corander, J. (2023). On predictive inference for intractable models via approximate Bayesian computation. *Stat Comput* 33(42) https://doi.org/10.1007/s11222-022-10163-6

Jiang, B. (2018). Approximate Bayesian Computation with Kullback-Leibler Divergence as Data Discrepancy. *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics* (pp. 1711-1721). Lanzarote, Spain.

Rodríguez-Sanz, A., Marcos, A. F., Pérez-Castán, J. A., Comendador, F. G., Valdés, R. A., & Loreiro, A. P. (2021) Queue behavioral patterns for passengers at airport terminals: A machine learning approach. *Journal of Air Transport Management*, 90, 101940. doi:https://doi.org/10.1016/j.jairtraman.2020.101940.

Wu, P. P., Pitchforth, J., & Mengersen, K., (2014). A Hybrid Queue-based Bayesian Network framework for passenger facilitation modelling. *Transportation Research Part C: Emerging Technologies, 46*, doi:https://doi.org/10.1016/j.trc.2014.05.005.