



Estimation of Number of Flight Using Particle Swarm Optimization and Artificial Neural Network

Ebru Pekel Özmen^a and Engin Pekel^b

^a Ebru PEKEL ÖZMEN: Istanbul University-Cerrahpasa, Industrial Engineering Department, Istanbul.
E-mail: pk1.ebru@gmail.com.

^b Engin PEKEL: Hitit University, Industrial Engineering Department, Corum. E-mail: pekelc@hotmail.com.

KEYWORD

artificial neural network; airport; particle swarm optimization; estimation

ABSTRACT

The number of flight (NF) is one of the key factors for the administration of the airport to evaluate the apron capacity and airline companies to fix the size of the flight. This paper aims to estimate the monthly NF by performing particle swarm optimization (PSO) and artificial neural network (ANN). Performed PSO-ANN algorithm aims to minimize the proposed evaluation criterion in the training stage. PSO-ANN based on the proposed evaluation criterion offers satisfying fitness values with respect to correlation coefficient and mean absolute percentage error in the training and testing stage.

1. Introduction

The number of flight (NF) is a key factor for the management of the airport to evaluate the apron capacity. Recently, ever-increasing air traffic loads force airlines and airports to search for new borders in operations efficiency. In 2017, Airports Council International (ACI) reported over 2.22 billion passengers in domestic and international flights in Europe. It is expected that this quantity will rise by at least 50%, in global air traffic by 2025. This raise requires the construction of new airports, major capacity expansions at busy airports, a commensurate adjustment of aviation infrastructure, and the identification of operational policies and managerial directives. In particular, airports are faced with persistent challenges related to apron capacity, a key bottleneck in the air transport system (Farhadi, Ghoniem *et al.* 2014).

NF is a factor for the airline companies to determine the size of flight. If the NF is not thoroughly estimated, it may occur in undesired situations. If the NF is underestimated, the profits to earn from passengers may be decreased, or if the NF is overestimated, it will lead to unnecessary costs. Although the NF is a very important variable, there is not enough study focused on the NF in the literature. Instead of the NF, there are many studies focused on variables such as flight time and the number of passengers.

Literature is investigated by considering “Flight” and “estimation” keywords and 6244 published papers are found and examined in detail. 6244 papers consist of 3589 articles, 2575 proceedings papers, and 80 review papers. None of the published papers performed the hybridization of particle swarm optimization (PSO) and artificial neural network (ANN). (Zhang, Liu *et al.* 2016) conducted a short-term forecast of air traffic flow by using fuzzy clustering and neural network methods. They confirmed that their method is reasonable and applicable

for short-term forecasting of air traffic flow.(Suryani, Chou *et al.* 2010)developed a model to estimate air passenger demand and to assess policy scenarios for expanding runway and passenger terminal capacity to satisfy future demand. (Xie, Wang *et al.* 2014)proposed two-hybrid approaches with the least squares vector regression (LSSVR) model and seasonal decomposition for short-term forecasts of air passenger.(Kim and Shin 2016) developed a model to forecast a short-term air passenger demand using big data from search queries to identify short-term fluctuations. In another paper, authors (Kim, Choi *et al.* 2016) proposed a multiple estimation model by performing the deep learning paradigmto evaluate a detailed analysis of models in air traffic delays.

In conclusion and to the best of our knowledge, the estimation of NF where hybridization of PSO and ANN is performed has not been previously addressed in the article papers.

The main contributions of the paper are as follows:

- The paper illustrated that the application of PSO-ANN to estimate NF presents a high coefficient of determination (R^2) value and low mean absolute percentage error (MAPE).
- The estimation of NF considering PSO-ANN is performed with different parameters.
- Performed PSO-ANN algorithm aims to minimize the proposed evaluation criterion.
- Future research directions are addressed in the conclusion section.

The rest of the paper is organized as follows: Section 2 explains the case study that is evaluated in this research. Section 3 presents the methodology consisted of PSO and ANN to estimate the NF. Section 4 presents the estimation results of the NF. Section 5 summarizes the overall findings.

2. Case study

Samsun, where has over a million people on the north coast, in Turkey, is considered as the capital of the Black Sea region. The growing city has two universities, many hospitals, and light manufacturing industry. Carsamba airport is a public airport, which is away 23 km from Samsun.The data used in this study is gathered from the General Directorate of State Airports Authority (D.H.M.I) in 2019. The data is collected as monthly and covers from 2010 to 2018 years. Dataset is split into two parts for training and testing stage. The initial dataset consists of 105 instances that include three inputs and one output. Inputs consist of month, extra holiday and population of Samsun and output is NF. 80% and 20% of the dataset are kept at random for training and testing stage, respectively.

3. Methodology

ANN, PSO and hybridization of PSO-ANN are given in this section. PSO-ANN is performed to reach the best estimation values because estimation model performing ANN may not offer the best results.

3.1. ANN

ANN is developed to imitate the basic biological neural systems in the human brain. A couple of interconnected simple processing nodes exist in ANN. An input signal is collected by each node and it is processed locally through an activation or transfer function. Thus, a transformed output signal is produced. Though each function is implemented by each individual neuron quite slowly, a network may execute an amazing number of tasks efficiently (Reilly and Cooper 1995, Zhang, Patuwo *et al.* 1998). Pseudo code of ANN is illustrated in Algorithm 1. All formal steps are listed concurrently. Initial weights are set off randomly in the first iteration and the output of each hidden neuron and the error are computed. The change of the weights is calculated with regard to a specified function and the change is utilized to update the weights. Then, the next iteration is carried out with regard to the updated weights.

Algorithm 1.

```

Initialize initial weights ( $w(t = 0)$ )
 $t \leftarrow 0$ ;
Compute the output for every neuron
Compute the error at the output
 $\Delta w(t) \leftarrow \text{change\_of\_weights}()$ ;
 $w_r(t) \leftarrow \text{weight\_update}()$ ;
Evaluate( $w_r(t)$ );
 $w_r(t) \leftarrow \text{build\_next\_generation}(w_r(t), w(t))$ 
 $t \leftarrow t+1$ ;

```

3.2. PSO

PSO is one of the population based stochastic optimization technique developed by the researchers [6]. PSO algorithm performs the behaviors of bird flocking. PSO algorithm aims to have all the particles locate the optima. It consists of a collection of particles that move around the search space affected by their own best past location and the best past location of the whole swarm.

The pseudo code of the PSO algorithm is illustrated in Algorithm 2. PSO algorithm begins by generating the initial particles and the initial velocities are assigned to the initial particles. In every iteration, each particle is updated by the following two best values. One of them is the fitness value that has been achieved so far and this value is stored as the best solution. The other value is the best value that is reached by any particle in the population.

Algorithm 2.

Pseudo code of the PSO algorithm

```

Initialize each particle
Population  $\leftarrow \Theta$ 
Output P_g_best  $\leftarrow \Theta$ 
For (k=1 to Population_size)
P_velocity  $\leftarrow \text{Random\_velocity}()$ 
P_position  $\leftarrow \text{Random\_position}(\text{Population\_size})$ 
P_p_best  $\leftarrow \text{P\_position}$ 
Evaluation
P_g_best  $\leftarrow \text{P\_p\_best}$ 
End
While(termination_criteria)
For each particle
P_velocity  $\leftarrow \text{update\_velocity}(\text{P\_velocity}, \text{P\_g\_best}, \text{P\_p\_best})$ 
P_position  $\leftarrow \text{update\_position}(\text{P\_position}, \text{P\_velocity})$ 
Evaluation
P_p_best  $\leftarrow \text{P\_position}$ 
Evaluation
P_g_best  $\leftarrow \text{P\_p\_best}$ 
End
End
Return(P_g_best)

```

$v_i(t+1)$ is the new velocity for the i . particle. C_1 and C_2 parameters are weighting coefficients for the local and global best positions, respectively. $p_i(t)$ is the i particle's position at iteration t . p_i^{best} is the i particle's

best-known position and p_{gbest} is the best position known to the swarm. The rand variable generates random values between 0 and 1.

$$v_i(t+1) = v_i(t) + \left(C_1 * rand * (p_i^{best} - p_i(t)) \right) + \left(C_2 * rand * (p_{gbest} - p_i(t)) \right) \quad (1)$$

Eq. (1) shows the calculation of updating the velocity of the particle.

$$p_i(t+1) = p_i(t) + v_i(t) \quad (2)$$

Eq. (2) shows the calculation of updating the position of the particle. Particle positions are used as weights between neurons and inputs in PSO-ANN.

3.3 PSO-ANN

The ANN is actually a novel computer architecture that is related to traditional computers by imitating biological neural networks. It allows using very simple computational operations such as fundamental logic elements to deal with complex, mathematically vague, nonlinear or stochastic problems. Recently, significant efforts have been made to improve the result of the fitness functions. One of the significant efforts is the hybridization of ANN with another meta-heuristic method.

In this paper, PSO and ANN are hybridized in order to improve the estimation quality of NF. Computation methodology is applied to the network connection weights. Algorithm 3 shows the pseudo code of the hybridized PSO-ANN.

Algorithm 3.

Pseudo code of the hybridized PSO-ANN

```

Initialize initial weights ( $w(t = 0)$ )
 $t \leftarrow 0$ ;
Initialize each particle
Population  $\leftarrow \Theta$ 
Output  $P_{g\_best} \leftarrow \Theta$ 
For ( $k=1$  to Population_size)
P_velocity  $\leftarrow$  Random_velocity()
P_position  $\leftarrow$  Random_position (Population_size)
P_p_best  $\leftarrow$  P_position
Evaluation
P_g_best  $\leftarrow$  P_p_best
End
While(termination_criteria)
For each particle
P_velocity  $\leftarrow$  update_velocity(P_velocity, P_g_best, P_p_best)
P_position  $\leftarrow$  update_position(P_position, P_velocity)
Evaluation
P_p_best  $\leftarrow$  P_position
Evaluation
P_g_best  $\leftarrow$  P_p_best
Update weights (P_position)
End
End
 $t \leftarrow t+1$ ;

```

Similar to ANN, initial weights are set off randomly in the first iteration and the output of each hidden neuron and the evaluation criterion is computed in PSO-ANN. The change of the weights is calculated by considering the evaluation criterion and the change is utilized to update the weights. Then, the next iteration is carried out with regard to the updated weights. The evaluation criterion is satisfied in a given iteration and the estimation of NF is carried out by performing PSO-ANN.

3.4 Evaluation Criterion

Evaluation criterion is an important factor to reach a satisfying estimation performance. Mean square error (MSE), MAPE and correlation coefficient (R^2) are chosen as an evaluation criterion in the estimation papers. However, a different fitness function is performed to reach the best estimation results in this paper.

$$\sum_{i=1}^{i=n} O_i * (|O_i - P_i|) \tag{3}$$

Eq. (3) shows the function of the evaluation criterion.

4. Numerical Results

PSO-ANN has been experimented to estimate the NF by satisfying the best fitness performance. All characteristics of all features and numerical results of the estimation experiment are provided in this section. The method of PSO-ANN has been run on a computer that has a 32-bit Windows 7 operating system, 2.9 GHz processor, and 4-GB memory. The performed method has been implemented in the Matlab. The total time that consists of training, testing stages and the estimation results takes approximately 150 seconds. Stopping criterion is the iteration number and it is set up to 200.

4.1. Training Results

Training results of NF are obtained considering three different parameters of PSO-ANN. All combinations of three different parameters are experimented considering Eq. (3) in PSO-ANN. C_1 and C_2 are the parameters of original PSO and experimented between 1 and 4 values. Hidden neuron number is experimented between 2 and 20 values. All combinations equal to 304. C_1 , C_2 and hidden neuron number that offer the lowest evaluation criterion are 1, 2 and 16, respectively.

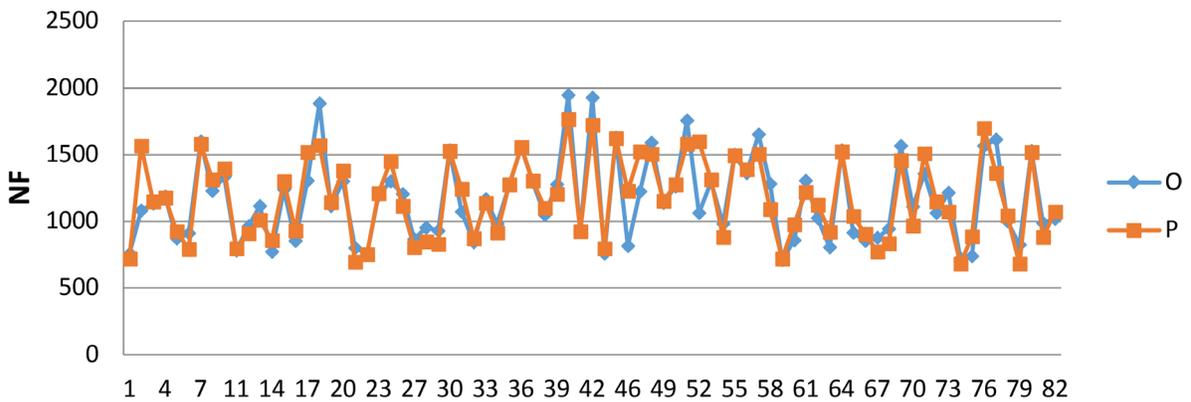


Figure 1: Estimation results of NF in the training stage.

Fig. 1 shows the estimation results of NF in the training stage. The estimation performance of NF is evaluated by calculating R^2 and MAPE. The performed PSO-ANN offers quite high R^2 (0.89) and low MAPE (0.089) values by considering Eq. (3) in the training stage. In Fig. 1, the coherent pattern between the output and the estimated value of NF shows that the proposed algorithm is well learned. The proposed PSO-ANN algorithm finds somewhat different estimation values in observation 17 and 46. However, the breakpoints in the NF are considerably predicted accurately.

4.2. Testing Results

Testing results of NF are calculated considering the lowest evaluation criterion computed in the training stage. Similar to training results, testing results show a decent estimation quality.

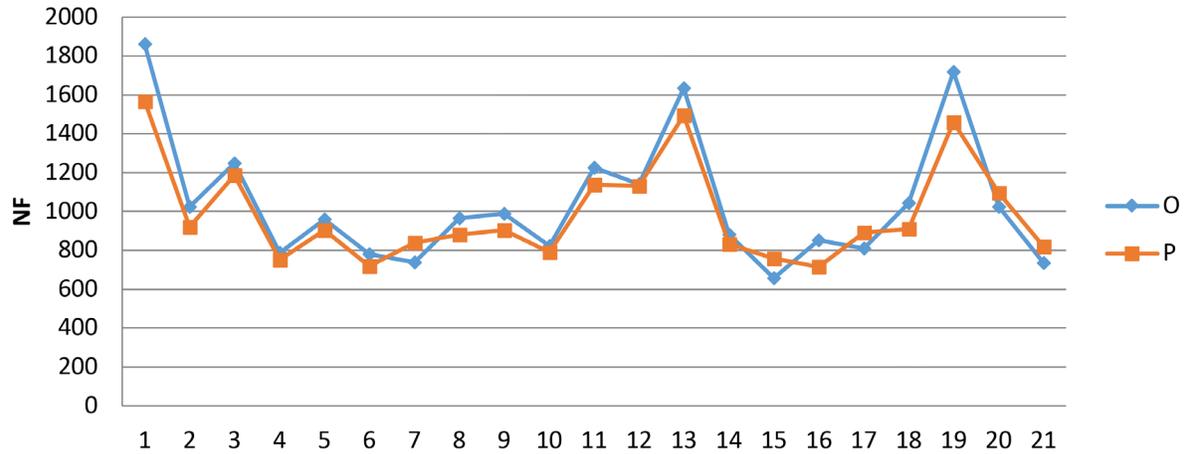


Figure 2: Estimation results of NP in the testing stage.

Fig. 2 shows the estimation results of NP in the testing stage. The proposed algorithm estimates the values and breakpoints of the NF graph thoroughly. The predicted values obtained with the implementation of the proposed algorithm and the observed values are seen in Fig. 2 to be very harmonious. Therefore, performance values of estimation are also very satisfying. The performed PSO-ANN offers quite high R^2 (0.96) and low MAPE (0.092) values by considering Eq. (3) in the testing stage.

5. Conclusion

The estimation of NF is a key factor for the management of the airport to evaluate the apron capacity and the airway companies to determine the size of flight. This paper addresses the estimation of NF by performing PSO-ANN that is not previously addressed in the article papers. Our main contributions are addressed in the following. First, the application of PSO-ANN to estimate NF presents a high R^2 value and low MAPE value. This means that the performed method satisfies two different fitness performance with their best values. Another contribution is that the performed methodology provides an accurate estimation in a short time. Besides, another important contribution is the estimation of NF by performing PSO-ANN that considers three different parameters (C_1 , C_2 and hidden neuron number).

Training and testing results of NF are obtained considering three different parameters of PSO-ANN. All combinations of three different parameters are experimented considering different evaluation criterion in PSO-ANN. C_1 , C_2 and hidden neuron number that offer the lowest evaluation criterion are 1, 2 and 16, respectively. The estimation performance of NF is also evaluated by calculating R^2 and MAPE. The performed PSO-ANN offers quite high R^2 (0.89) and low MAPE (0.089) values by considering Eq. (3) in the training stage. Similar

to training results, the performed PSO-ANN offers quite high R^2 (0.96) and low MAPE (0.092) values by considering Eq. (3) in the testing stage.

Further research needs to be conducted to improve the estimation of NF by considering different evaluation criteria and other algorithms. Design of experiment methodology maybe investigated by applying to the estimation of NF because the methodology may reduce the required the instance number of data. The reduced instance number of data may provide better fitness values. Additionally, other supervised learning methods, including kernel ridge regression, deep learning, SVM and neural network, may be investigated by applying to the estimation parameters of NF in this paper.

6. References

- Farhadi, F., A. Ghoniem and M. Al-Salem (2014). "Runway capacity management-an empirical study with application to Doha International Airport". *Transportation Research Part E: Logistics and Transportation Review* 68: 53-63.
- Kim, S. and D. H. Shin (2016). "Forecasting short-term air passenger demand using big data from search engine queries". *Automation in Construction* 70: 98-108.
- Kim, Y. J., S. Choi, S. Briceno and D. Mavris (2016). A deep learning approach to flight delay prediction. 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC).
- Reilly, D. L. and L. N. Cooper (1995). An overview of neural networks: early models to real world systems. *How We Learn; How We Remember: Toward An Understanding Of Brain And Neural Systems: Selected Papers of Leon N Cooper*, World Scientific: 300-321.
- Suryani, E., S.-Y. Chou and C.-H. Chen (2010). "Air passenger demand forecasting and passenger terminal capacity expansion: A system dynamics framework". *Expert Systems with Applications* 37(3): 2324-2339.
- Xie, G., S. Wang and K. K. Lai (2014). "Short-term forecasting of air passenger by using hybrid seasonal decomposition and least squares support vector regression approaches". *Journal of Air Transport Management* 37: 20-26.
- Zhang, G., B. E. Patuwo and M. Y. Hu (1998). "Forecasting with artificial neural networks: The state of the art". *International journal of forecasting* 14(1): 35-62.
- Zhang, M., K. Liu, H. Yu and J. Yu (2016). Short-term Forecasting Method of Air Traffic Flow based Neural Network Ensemble. 2016 2nd Workshop on Advanced Research and Technology in Industry Applications (WARTIA-16), Atlantis Press.

