https://doi.org/10.1016/j.tre.2019.10.005

A Novel Self-Organizing Constructive Neural Network for Estimating Aircraft Trip Fuel Consumption

Waqar Ahmed Khan[a], Sai-Ho Chung[a*], Hoi-Lam Ma[b], Shi Qiang Liu[c], Ching Yuen Chan[a]

[a]*Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong*
[b]*Department of Supply Chain and Information Management, The Hang Seng University of Hong Kong, Hong Kong*
[c]*School of Economics and Management, The Fuzhou University, China*

*Corresponding Author: nick.sh.chung@polyu.edu.hk (Sai-Ho Chung)

Abstract- Accurate estimation of aircraft fuel consumption is critical for airlines in terms of safety and profitability. In current practice, estimation of fuel consumption for a flight trip is usually done by engineering approaches, which mainly consider physical factors, e.g., planned weather and planned cruise level. However, the actual performance of a flight usually deviates from such estimation. Therefore, we propose a novel self-organizing constructive neural network (CNN) that features a cascade architecture and analytically determines connection weights to estimate the trip fuel of a flight. The proposed method generates non-redundant and linearly independent hidden units by an orthogonal linear transformation of operational parameters to achieve the best least-squares solution. Our findings provide insights for the aviation industry in controlling airlines' excess fuel consumption.

Keywords: aircraft fuel estimation; engineering approach; high dimensional data; machine learning; neural network.

1. Introduction

The calculation of the amount of trip fuel is essential for an aircraft to safely reach the destination. Consequently, it receives much attention in the aviation sector. Controlling excess fuel consumption has become one of the major concerns for airline operating organizations given that it contributes to increasing operating expenses (Sheng et al., 2019). Globally, during the last decade, fuel cost accounts for an average contribution of 28.2% of the total operating cost among various airline operating expenses. Therefore, it is critical to design methodologies for accurately planning the trip fuel required for each flight (IATA, 2019). The required quantity of trip fuel loaded in an aircraft depends on many operational parameters and estimation methods. Loading suboptimal trip fuel may result in the utilisation of fuel from the supplementary reservoir, whereas abundant trip fuel may increase the ramp weight. Both situations are undesirable for smooth operation of an airline. Utilising supplementary reservoir fuels, which are reserved to meet unexpected flight conditions such as bad weather, alternative airport divergence, airport congestion, and hold-on, may create uncertainty for the flight crew. Conversely, loading abundant fuel may ensure a safe journey but with an additional cost in terms of excess fuel consumption and early aircraft maintenance. In both scenarios, the actual trip fuel consumed during each flight significantly deviates from the estimated trip fuel. Fuel deviation, defined as the difference between the actual trip fuel consumed during a flight and the trip fuel estimated before that flight, may take either negative or positive values corresponding to underestimation or overestimation, respectively. A low confidence in estimation methods and the need to meet unforeseen flight issues require adding an extra amount of fuel in the discrepancy – rather than contingency or final emergency – reservoir based on experience. This makes the situation worse for airlines. First, it increases the total weight of the aircraft, requiring more thrust to balance weight and drag in combination with other atmospheric and physical factors (Irrgang et al., 2015). Taking more fuel onboard not only increases the weight of an aircraft but also affects the performance of its engines in the long run by burning more fuel per unit distance. This ultimately shortens the lifetime of engines, which need more frequent maintenance than planned (Abdelghany et al., 2005). Moreover, an aircraft may require more fuel to burn than a brand-new one and the pilots may be unaware of the actual amount of wear (deterioration) in the aircraft or may not know how the wear is calculated by particular estimation methods. As a result, the pilots will demand more fuel as a buffer in the discrepancy reservoir of the aircraft (Irrgang et al., 2015). Therefore, the objective of reaching the destination with the smallest possible amount of fuel left in the trip tank or utilised from the supplementary tank constitutes a challenging task to accomplish.

Excess fuel consumption is disadvantageous for airlines both economically and environmentally. Economically, because the increasing trend of fuel prices and a low confidence in estimation methods may lead to more costs for airline operating organizations to ensure smooth operations while meeting the growing need of passengers and cargo. Fig. 1 shows the trend of increasing jet fuel prices. In 2019, the forecast has been that fuel prices will increase by 31.18% as compared to those in 2015 (IATA, 2019). This forces airlines to adopt risk management strategies such as fuel hedging and fuel ferrying, which may not be useful in many cases (Merkert and Swidan, 2019; Sibdari et al., 2018). Fuel hedging and fuel ferrying are certainly good strategies, but they become expensive when fuel prices fall − fuel hedging − and because of the high aircraft maintenance cost associated with loading excess fuel from inexpensive stations − fuel ferrying. Figs. 2 and 3 show data on traveling passengers and cargo movement in the airline sector over the past decade. The demand, in terms of revenue passenger kilometres (RPKs), has increased 7.4% in 2018 as compared to 2017. The cargo demand, in terms of freight tonne kilometres (FTKs), has risen to its highest level in the last ten years with a growth rate of 9.7% in 2017 as compared to the level in 2016, which is more than double the 3.6% growth rate in 2016. Indeed, the growth rate has further increased 3.4% in 2018 as compared to the level in 2017 (IATA, 2019). Environmentally, higher fuel consumption causes ozone depletion by more emission of carbon dioxide (CO_2) and greenhouse gases (GHG) (Pagoni and Psaraki-Kalouptsidi, 2017; Seufert et al., 2017). Fig. 4 shows the trend of fuel consumption and carbon dioxide emission. Actually, International Air Transport Association (IATA) stresses to avoid ozone depletion by joint efforts to

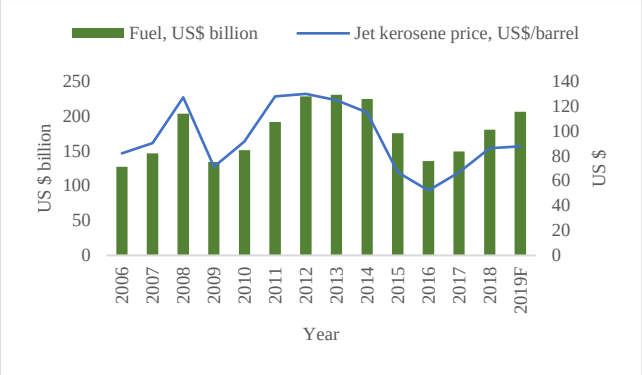

Fig. 1. Fuel cost (US$ billion) and cost per barrel (US$)

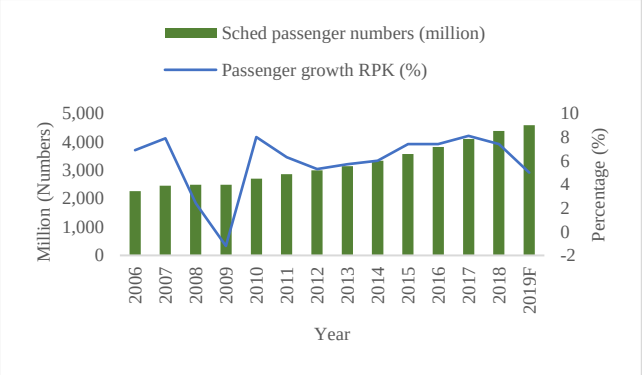

Fig. 2. Passenger demand (RPK)

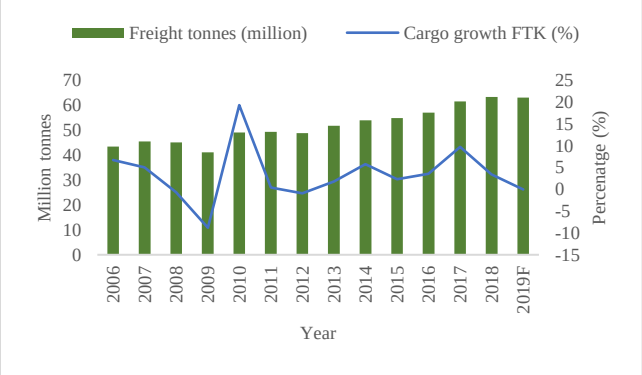

Fig. 3. Cargo demand (FTK)

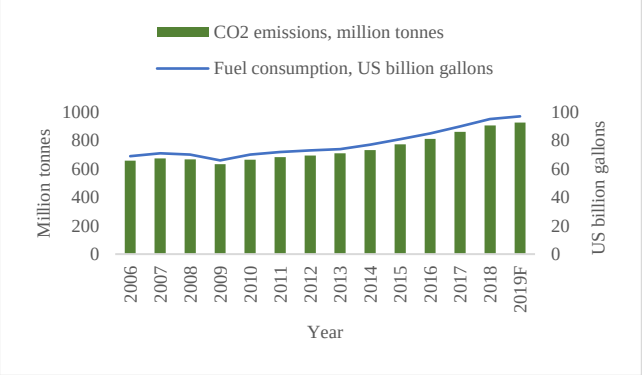

Fig. 4. Fuel consumption and carbon dioxide emission

reduce CO_2 emissions by 50% by 2050 with respect to 2005. In the future, international authorities are planning to make it compulsory for airlines to certify their aircrafts based on size and weight, according to CO_2 certification standards (IATA, 2018). The increasing awareness for environmental protection by international authorities in conjunction with growing fuel prices and boosting demand from tourism are encouraging airline operating companies to adopt competitive strategies in fuel management to control excess fuel consumption for long-term sustainability.

Early efforts to estimate fuel for each flight made use of mathematical formulations derived from energy balance expressions. The most popular of them were the Simmod simulation program, developed from the so-called advanced fuel burn model (AFBM) (Collins, 1982) by the Federal Aviation Administration (FAA) (Baklacioglu, 2016), and the base of aircraft data (BADA) developed by the European Organization for the Safety of Air Navigation (Eurocontrol) (Nuic, 2014). The energy balance of the aircraft is maintained by balancing the energy loss (drag) and gain (thrust) through control of fuel consumption. Senzig et al. (2009) highlighted that BADA is an effective method that works better for cruise phase; however, it losses accuracy for terminal areas such as takeoff and descent. Besides, the information needed to determine coefficients for energy balance methods is not always available in a real scenario and it needs to be generated from other sources, what may limit their applicability (Pagoni and Psaraki-Kalouptsidi, 2017; Trani and Wing-Ho, 1997; Yanto and Liem, 2018). For the sake of simplicity, the methods based on mathematical formulations of energy balance are named as engineering approaches (EAs) in our study. To provide an alternative and simplify the EA-based fuel estimation methods, the application of a fixed-typology backpropagation neural network (BPNN) was suggested in the literature. Schilling (1997), Trani et al. (2004), and Baklacioglu (2016) proposed fuel estimation models based on such BPNN method. The main limitation of previous works on the application of a BPNN for fuel estimation is that they only covered a small number of aircraft types with limited flight data. The reasons for this may be the weak generalization performance and slow convergence drawbacks of a BPNN based on trial-and-error approaches to select optimal hyperparameters (Kapanova et al., 2018). The generalization performance is inadequate because it converges at a local minimum if the global minimum is far away. Concerning convergence (learning speed), it becomes slow if the learning rate is small and unstable when the learning rate is large. In addition, the need for iterative tuning of the connection weights gives rise to a time-consuming algorithm (Huang et al., 2006b). The traditional BPNN is based on a fixed topology. The selection of a suitable combination of hidden units and hidden layers along with other global and local hyperparameters make this approach more dependent on expert involvement. This fact may make it intractable to solve.

To overcome the above limitation of BPNNs and its implication on trip fuel estimation, we propose a self-organising constructive neural network (CNN) featuring a cascade topology and capable of analytically calculating connection weight coefficients to achieve better generalization performance and faster learning speed. The purpose is to eliminate the need for a trial-and-error approach and reduce the number of hyperparameter adjustments and expert involvement. We consider that insufficient attempts have been reported in the literature concerning estimates of trip fuel using CNNs along with high-dimensional data for the entire trip flight phases collectively. Previously, the entire trip was divided into different flight phases, namely takeoff, climb, cruise, descent, and approach, to propose a fuel estimation model with limited flight data restricted to a small number of aircrafts or generated by a simulation planner. In our study, we analyse high-

dimensional data associated with 107 wide-body aircrafts − Airbus A330-300 (31 aircrafts) and Boeing 747-400 (9 aircrafts)/747-800 (14 aircrafts)/777-300 (53 aircrafts) − that operated a total of 19,117 real international passenger and cargo flights in eight sectors over two years. A comparative study of the proposed CNN was performed with an existing airline engineering approach (AEA) and a BPNN. The numerical results demonstrate that the trip fuel estimation by the proposed CNN achieves better results while requiring the shortest time compared to AEA and BPNN. The proposed CNN also exhibits an average improvement of 45.71% and 25.71% for a combined eight-sector model, and 67.93% and 10.15% for a sector-wise individual model, compared to AEA and BPNN, respectively. The significant improvement in trip fuel estimation creates greater confidence in the proposed CNN given that it may eliminate the need for adding more fuel based solely on experience.

The rest of this paper is structured as follows. Section 2 presents a literature review. Section 3 describes the trip fuel estimation model. Section 4 explains the existing BPNN and shows its limitation in comparison with the proposed CNN. Section 5 discusses experimental work and Section 6 concludes our study.

2. Literature Review

The trip fuel containing a maximum portion of fuel weight must be considered for improvement. A small improvement in fuel estimation can bring significant savings in fuel consumption and substantial cost benefit (Jensen et al., 2013). Irrgang et al. (2015) concluded from their work on aircraft fuel optimisation analytics that the optimal amount of fuel should be accurately determined, resulting in the minimum amount of fuel remaining at the destination airport. Taking extra fuel increases the weight of the aircraft, what results in more fuel consumption than required. Turgut et al. (2014) worked on the fuel flow during the cruise phase. They suggested that reducing the aircraft weight by 1 tonne, increasing the flying altitude by 100 ft and reducing the cruise speed by 1 knot may result in a reduction of per-hour fuel consumption of 15-21 kg, 26-28 kg, and 7.7-8.7 kg, respectively. Taking more fuel increases the weight of the aircraft, and in consequence, more fuel is consumed with higher maintenance cost (Abdelghany et al., 2005). Therefore, more efficient and effective methods are needed to make the estimation models applicable (Choi et al., 2016).

2.1. Engineering approaches (EAs)-based fuel estimation

Early attempts made use of engineering approaches (EAs) to estimate the fuel for aircrafts. EAs are based on extensive mathematical formulations and are derived from the basic concept of energy balance. They assume static constants of aircraft performance and a dynamic input of the path profile. The energy balance can be expressed in terms of the energy the aircraft gains and losses as it travels along the prescribed path profile. During each flight operation, as a result of the change in kinetic and potential energies, the aircraft suffers energy losses because of drag. These losses are adjusted by a certain amount of thrust to maintain the energy balance. Thus, an aircraft can be viewed as a system undergoing energy losses and gains that should be continuously balanced by consumption of fuel energy. Collins (1982) derived an algorithm for fuel estimation based on such energy balance approach by considering the description of aircraft configuration, weight, and path profile. The path profile depends on a change in true airspeed, altitude, and flight time. Collins (1982) explained that various aircraft-specific constants need to be determined for each aircraft to define the relationship between drag and lift/weight coefficients and to define the relationship between fuel consumption and energy gain from the thrust as a function of velocity and altitude. Experimental work demonstrated that the algorithm could result in saving more than two million US gallons of fuel annually. The accuracy check of the algorithm with Eastern Airlines demonstrated its effectiveness with a difference of less than 3%. This algorithm, known as AFBM, was developed by the Mitre corporation under the FAA and was incorporated in their simulation software Simmod to predict the fuel consumption for each flight (Baklacioglu, 2016; Schilling, 1997). Similarly, BADA, developed by Eurocontrol (Nuic, 2014), is another engineering-based fuel estimation method that estimates thrust-specific fuel consumption (TSFC) as a function of true airspeed. The fuel for different phases and engine types (Jet, Turbo, and Piston) can be calculated from coefficients defined for each aircraft type in operational performance files (OPF) of BADA. The BADA database holds a set of files explaining the performance and operating procedure coefficient of many different types of aircrafts. The coefficients are used to calculate drag, thrust, and fuel flow. As a result, climb, nominal cruise, and descent speeds for the targeted aircraft are provided. The mathematical formulations used to calculate fuel flow for EA (Simmod and BADA) models are presented in Appendix A.

The AFBM Simmod model is based on various parameters to estimate fuel flow. The parameters include engine constants, airspeed, aircraft mass, atmospheric density, extra thrust needed for ascending, and wing area (Huang et al., 2017). Trani and Wing-Ho (1997) argued that the information needed to determine fuel flow is not always available in a real scenario and a great deal of flight testing is needed to generate the parameter data, what may make the Simmod method expensive. The Simmod method has certainly improved the measurement of airline fuel efficiency (Baklacioglu, 2016; Schilling, 1997; Senzig et al., 2009; Trani et al., 2004). However, Senzig et al. (2009) reported that BADA is an effective method that has gained much more popularity than Simmod. Indeed, BADA can estimate fuel consumption with an average deviation of 3% during the cruise phase. For the terminal areas, such as takeoff and descent, BADA losses accuracy. An average difference of -22.3% was observed between the BADA model and the fuel consumption reported by an airline during the takeoff phase. Moreover, according to recent studies, aircraft fuel estimation may not be accurate in EA methods (Simmod and BADA) owing to outdated existing databases and the unavailability of data for all types of aircrafts and engines (Yanto and Liem, 2018). The unavailability of data may result in the usage of more global parameters with default values rather than local parameters, what in turn may lead to inaccurate results (Pagoni and Psaraki-Kalouptsidi, 2017). Complex mathematical computation, high testing and consultation cost, expert involvement, and estimation errors, which may become worse in certain conditions, limit EA applicability.

2.2. BPNN-based fuel estimation

In the literature, many efforts have been made to simplify EA-based fuel estimation methods with respect to BPNNs. In this study, although BPNNs are certainly one of the EAs, we assume only Collin's approach and BADA as EAs. Schilling (1997) proposed to estimate aircraft fuel by a BPNN in conjunction with the AFBM energy balance theory to simplify the computational capability and improve the estimation accuracy of the AFBM. The engine-specific constant in AFBM was determined by multivariate regression with an average accuracy error of 4% (Schilling, 1997). Their proposed aircraft fuel consumption model (AFCM) comprises two systems: 1) initial fuel estimation by AFBM; and 2) information generated by AFBM such as altitude, velocity, and fuel estimated to map into the BPNN. Two inputs with 600 flight instances were trained and a network with 2 layers and 7 hidden units in each layer was selected through a trial-and-error approach among six candidate BPNNs. The employed various activation function with a backpropagation Levenberg-Marquardt (LM) learning algorithm. Boeings 747-100 and 767-200, Bombardier Dash-7, DC 10-30, and Jetstar were aircrafts used for performance comparison. Schilling (1997) demonstrated that AFCM achieved better results compared to AFBM. AFCM was modelled based on low-level input variables such as altitude and velocity with the future recommendation of incorporating more operational parameters.

Trani and Wing-Ho (1997) pointed out that EAs are challenging to implement because of the information required to define the constants and to determine the coefficients demands a great effort in addition to testing and expertise. These disadvantages are the main constraints for their expansion. Trani et al. (2004) proposed a BPNN as an alternative approach for fuel estimation during the climb, cruise, and descent phases of a flight. Eight different sizes of BPNN candidate networks for fuel estimation were constructed according to four operational parameters, namely Mach number, weight, temperature, and altitude, covering the cruise phase in a dataset comprising 1,610 flight instances performed by Fokker F-100 aircrafts. They suggested that a BPNN trained by LM with three layers is the best approach for fuel estimation with eight hidden units in the first two layers each and one unit in the output layer. The employed various activation function with maximum stopping training epochs of 10,000. The BPNN was trained for the cruise phase with a target vector of a specific air range rather than consumed fuel. The trained BPNN model of the cruise phase was also implemented for the climb and descent flight phases. Actually, the climb and descent phases consisted of a fuel burn output vector. This implies that this research work is more focused on engine performance over distance travelled during the cruise phase. The work does not clearly distinguish the estimation in different flight phases because the specific air range is defined as the distance covered per unit of consumed fuel (Nautical Miles/ Kilogram or NM/kg), whereas fuel burn may be related to fuel consumption (kg) during flight. There exist many similarities between the suggestions of (Schilling, 1997) and (Trani et al., 2004) concerning BPNNs. These similarities are based on the same assumptions, involving trial-and-error approaches to find hidden units for two hidden layers along with different activation functions.

Baklacioglu (2016) improved the work reported in (Trani et al., 2004) and proposed a so-called genetic algorithm optimised neural network (GA-NN) to determine the network hyperparameters and number of hidden layers to achieve the best model with less effort. Such network considered two input operational parameters, i.e., altitude and airspeed, as a function of fuel flow rate, with a dataset composed of 347, 404, and 483 flight instances for the climb, cruise, and descent phases, respectively. The targeted aircraft was a medium-weight Boeing 737-800. The work indicated that GA-NNs with one hidden layer are better than those with two hidden layers and achieve improved results for the climb and descent phases compared to a previous suggested model (Trani et al., 2004).

The application of neural networks (NNs) is gaining much popularity in the airline sector to improve its various operations and enhance services. For instance, Lin and Vlachos (2018) proposed an analytical framework, known as "Importance-Performance-Impact-Analysis", to improve customer satisfaction from airlines by incorporating techniques such as BPNNs, decision-making trail, and evaluation laboratory. Cui and Li (2017) studied airline efficiency measures under "Carbon Neutral Growth from 2020" and used a BPNN to predict the CO_2 emission volume for an individual airline. Khanmohammadi et al. (2016) studied the issue of nominal variables in data and proposed to use a multilevel input layer BPNN to predict incoming flight delays for the John F. Kennedy (JFK) International Airport in New York. The list of applications of neural networks in the airline domain is long, thereby motivating to apply NN-based machine learning methods to study fuel estimation for flight trips.

2.3. Estimation methods other than NN-based machine learning methodologies

Apart from NN machine learning methodologies, researchers have also developed models to achieve a better fuel estimation. Turgut and Rosen (2012) proposed a genetic algorithm fuel consumption model for commercial flights in a Boeing B737-800 by studying the relationship between altitude and fuel flow during four different descents within the descent flight phase. This model suggests avoiding low-level flight and maintaining higher altitude if possible during the descent phase for greater fuel saving. When a delay condition occurs, aircraft holding at a higher altitude could save substantial fuel rather than holding the flight at a lower altitude. The results demonstrate that performing low-level flights at an altitude 1000 ft higher could substantially decrease fuel consumption. Chati and Balakrishnan (2017) studied the impact of takeoff weight (TOW) on aircraft fuel consumption and proposed a Gaussian process regression (GPR) statistical approach to determine TOW using observed data from takeoff round roll. The estimated TOW model was used as an input to study its effect on fuel consumption during ascent. In particular, GPR-estimated TOW was averaged over eight different types of aircraft for 874 flight instances. The predicted TOW was used as an input feature to estimate the fuel flow rate during the ascent phase. Ryerson et al. (2011) studied the impact of three performance metrics, namely schedule padding, airborne delay, and departure delay, on two aircrafts − Boeing 757-200 and 737-800 − on fuel consumption. They found that airborne delays burn 50-60 lbs/minute of fuel compared with schedule padding, which is 4.5-12 lbs/minute, and a departure delay, which is 2.3-4.6 lbs/minute. Additionally, a congested airport terminal area increases fuel consumption by 16%. Improvement in performance measurements by eliminating the above three delays by econometric methods can reduce airborne fuel consumption.

2.4. EA and BPNN limitations

We limit the scope of our study to EAs and BPNNs. EAs are disadvantageous in that they require a lot of effort to perform flight testing, and they involve expertise for complex mathematical formulation as well as cost to determine coefficients. All these aspects make EAs difficult to implement (Pagoni and Psaraki-Kalouptsidi, 2017; Senzig et al., 2009; Yanto and Liem, 2018). The efforts based on BPNNs to provide an alternative for EAs can be considered as an improvement; however, existing research works are restricted to trial-and-error approaches to determine the optimal hyperparameters, the number of hidden units and layers, and the activation functions. The major drawbacks of the fixed topology of BPNNs are that it involves iterative tuning of connection weights. This tuning may converge at local minima when global minima are far away. As a result, learning becomes slow when the learning rate is low and unstable when the learning rate is high (Huang et al., 2006b). Other issues include local and global hyperparameter adjustment and decision, for instance, setting the learning rate, connection weights, hidden units, hidden layer, and gradient learning algorithm. Too many hyperparameters and their adjustment with iterative tuning influence between each other make the network complex, which leads to the problem known as weak generalization performance and slow convergence (Huang et al., 2006a; Kapanova et al., 2018; Krogh and Hertz, 1992; Liew et al., 2016; Srivastava et al., 2014). The aforementioned drawbacks need to be resolved in a more innovative way to get an accurate trip fuel estimation for each flight with less expertise requirement and faster learning speed.

2.5. Innovation and contribution

Early attempts to estimate the trip fuel by EAs have gained popularity. However, the successful wide use of machine learning in many applications has changed the interest of the airlines operating organisation. The unavailability of data related to local parameters, complex mathematical formulations, and expertise involvement limit the applicability of EAs. The application of BPNNs as an alternative to EAs constitutes an attempt to overcome the above limitations with improved results. However, the existing BPNN-based fuel estimation models only cover a small number of aircraft types with limited flight data. The possible reasons are the random generation and iteratively tuning of the connection weights on both sides of BPNNs by gradient-based learning algorithms. This may create a complex co-adaptation by generating redundant hidden units, thereby contributing less to error convergence. The adjustment of local and global hyperparameters with the connection weights may require many experimental trials to obtain an optimal fixed-topology network. These problems increase the expertise requirement as well as cause weak generalization performance and slow convergence.

The application of traditional machine learning NNs for airline trip fuel estimation is not straightforward. Extensive knowledge is required to fit the algorithm to the model. The varying nature of airlines input parameters and the high dimensionality of data might be challenging for traditional BPNNs to accurately estimate the trip fuel with the aforementioned limitations. The same problem has been observed in our experimental work (discussed in Section 5). Similarly, in the current literature, the application of fixed-topology BPNNs for fuel estimation is limited in scope by considering a few aircraft types, flight data, and operational parameters. This means that the application of machine learning NNs on the varying nature of operational parameters and the high dimensionality of data are still an open area. To overcome the limitations of BPNNs, this study proposes a CNN featuring a self-organising cascading architecture and capable of analytically calculating connection weights. These characteristics generate linearly independent (non-redundant) hidden units having capability of maximum error reduction. The proposed CNN thus achieves a better generalization performance and converges significantly faster than traditional BPNNs.

The application of the proposed CNN with varying operational parameters and high-dimensional data for airlines trip fuel estimation can bring a breakthrough improvement compared to existing fuel estimation methods. Unlike EAs, the advantages of the proposed CNN are that it requires neither the involvement and consultation of an expert for a complex mathematical formulation nor the cost to determine the coefficients. In addition, it can work with any type of available data. Compared to BPNNs, the proposed CNN avoids the fixed-topology trail-and-error experimental works with too much adjustment of hyperparameters and iterative tuning of the connection weights. These advantages simplify the implementation on high-dimensional data with the objective of achieving better trip fuel estimation with fast convergence. A comparative study has been performed among the proposed self-organising CNN, AEA, and a BPNN to estimate airline trip fuel. To make the study more valuable, the comparison is not limited to high-dimensional data and its varying operational parameters (referred to as combined model in Section 5.3(a)). We also include the portion of data resulting from dividing high-dimensional data into a small dataset representing the same behaviour within operational parameters (referred to as individual models in Section 5.3(b)). The significant improvement in trip fuel estimation and the learning speed in both models (combined model and individual models) increase the confidence in that the proposed CNN may eliminate the need of adding more fuel based solely on experience.

3. Trip Fuel Estimation Model

In this study, we work on a dataset obtained from an international airline operating in Hong Kong. Currently, it experiences the usual problem of excess fuel consumption that increases its fuel expenses. Before each flight operation, the flight plan is prepared such that it contains details about the amount of fuel to be loaded in the tank reservoirs. Fig. 5 shows the various amounts of takeoff fuel (tof) needed for flight trips with fuel deviation, including

$$tof = f(taxi, trip\ fuel, contingency, extra, discrepancy, alternate, final\ reserve) \quad (1)$$

1) *Taxi Fuel*: Amount of fuel needed to operate an auxiliary power unit, start the engine, and cover the ground distance before starting the takeoff.
2) *Trip Fuel*: Amount of fuel needed for normal flight operation from the takeoff at the departure airport to landing at the arrival airport.
3) *Contingency Fuel*: Additional fuel loaded to meet holding and insufficient block fuel.
4) *Extra Fuel*: Additional fuel loaded to manage bad weather conditions and/or airport congestion.

5) *Discrepancy Fuel*: Additional fuel loaded according to experience to meet unforeseen conditions and/or account for aircraft deterioration.
6) *Alternate Fuel*: Additional fuel loaded to fly to an alternate airport if required.
7) *Final Reserve Fuel*: Last emergency reservoir to handle any uncertain situation.

To achieve the objectives of better fuel estimation with less expertise requirement and faster learning speed, our study focuses on analysing high-dimensional data (Choi et al., 2018; Chung et al., 2015). These high-dimensional data were provided by the airline and contain details about historical real flights operated from April 2015 to March 2017. The useful operational parameters are extracted from data that contributes to fuel consumption. The extracted data are applied to estimate the trip fuel before each flight and compare it with actual consumed fuel to measure fuel deviation in absolute percentage error. The comparative study is performed among existing AEA, a BPNN, and the proposed CNN (discussed in Section 4.2). These methods are used to estimate fuel for each flight. The one leading to a lesser fuel deviation is considered for its effectiveness.

During each flight operation, the aircraft generates a considerable amount of high-dimensional data with many operational parameters. Selection of relevant operational parameters plays an important role in model performance (Guo et al., 2018). Instead of putting all the operational parameters from the collected data into algorithms, a correlation analysis was performed among operational parameters and the consumed fuel to identify and select the most relevant operational parameters that significantly contribute to fuel consumption. Table 1 shows the selected operational parameters in such conditions. The runway direction is only a categorical variable included because it also influences fuel consumption. Statistical work shows that takeoff from a runway that is opposite in direction to the destination consumes more fuel because of the loop the aircraft must perform to face towards the destination airport.

From the selected operational parameters, we consider the fuel consumption for each flight in the form

$$\boldsymbol{Y} = f(\boldsymbol{X}) \tag{2}$$

$$\boldsymbol{Y} = f(t, m, w, temp^{a}, temp^{g}, h^{l1}, h^{l2}, h^{l3}, h^{l4}, h^{l5}, d, perf^{ac}, mac^{zfw}, mac^{tow}, mac^{law}, v, rwy). \tag{3}$$

The objective function is

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\hat{\boldsymbol{Y}}_i - \boldsymbol{Y}_i\right)^2} \tag{4}$$

or

$$MAPE = \frac{100\%}{m}\sum_{i=1}^{m}\left|\frac{\hat{\boldsymbol{Y}}_i - \boldsymbol{Y}_i}{\boldsymbol{Y}_i}\right| \tag{5}$$

such that

$$\hat{\boldsymbol{Y}} = (\emptyset(\boldsymbol{X}\boldsymbol{W} + b))\boldsymbol{\beta} + e \tag{6}$$

$$\emptyset(z) = \frac{1}{1 + e^{-z}} \tag{7}$$

where

$\boldsymbol{Y}$= Output target variable representing actual consumed fuel after flight operation in Kilogram (Kg)

$\boldsymbol{X}$ = Input variables with *m*-rows representing flight instances and *n*-columns representing operational parameters

t = Flight time duration from takeoff to landing in minutes (min)

m = Ramp weight including aircraft, passengers, crew, and usable and unusable fuel weight in Kilogram (Kg)

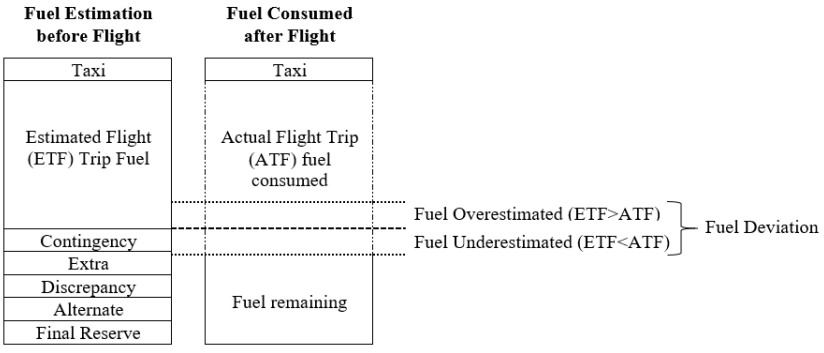

Fig. 5. Fuel estimation before flight and consumption after the flight

Table 1
Correlation analysis of selected operational parameters with trip fuel consumption

Operational Parameters	SR_H^{USI}	SR_H^D	SR_{US1}^H	SR_{US2}^H	SR_{UK}^H	SR_H^{US2}	SR_H^{UK}	SR_H^A	All Sectors
t	0.364	0.863	0.735	0.940	0.645	0.782	0.524	0.046	0.766
m	0.897	0.681	0.536	0.850	0.825	0.713	0.812	0.888	0.783
w	-0.379	-0.876	-0.741	-0.922	-0.726	-0.712	-0.616	-0.226	-0.295
$temp^a$	-0.585	0.770	0.454	0.548	0.015	-0.416	0.157	0.131	0.021
$temp^g$	-0.081	0.802	0.283	0.455	-0.053	-0.293	0.173	0.641	-0.059
h^{l1}	-0.833	-0.596	-0.551	-0.222	-0.570	-0.448	-0.111	-0.541	-0.577
h^{l2}	-0.812	-0.506	-0.604	-0.177	-0.305	-0.451	-0.412	-0.584	-0.687
h^{l3}	-0.589	0.056	-0.080	-0.041	-0.331	-0.457	-0.382	-0.506	-0.368
h^{l4}	-0.141	0.331	0.114	0.112	-0.293	-0.287	-0.171	-0.052	0.151
h^{l5}	-0.089	0.261	0.041	-0.152	0.071	-0.012	0.086	0.175	0.178
d	0.046	0.456	0.176	-0.279	0.111	0.180	0.098	0.073	0.687
$perf^{ac}$	-0.144	-0.014	-0.227	0.138	-0.131	-0.250	-0.047	-0.211	-0.317
mac^{zfw}	-0.047	-0.025	0.046	0.043	-0.033	0.067	-0.067	-0.468	-0.047
mac^{tow}	-0.469	-0.190	-0.436	-0.309	-0.240	-0.375	-0.268	-0.538	-0.344
mac^{law}	-0.029	0.205	0.056	0.042	-0.031	0.065	-0.066	-0.404	-0.118
v	0.155	-0.012	-0.165	-0.220	0.045	0.107	-0.099	0.105	0.422

w= Wind speed (a negative value means headwind whereas a positive value means tailwind) in Knots (kt)

$temp^a$= Temperature deviation at air in degree Celsius (°C)

$temp^g$= Temperature deviation at ground in degree Celsius (°C)

h^{l1}= Cruise altitude level 1 in hectofeet (hft)

h^{l2}= Cruise altitude level 2 in hectofeet (hft)

h^{l3}= Cruise altitude level 3 in hectofeet (hft)

h^{l4}= Cruise altitude level 4 in hectofeet (hft)

h^{l5}= Cruise altitude level 5 in hectofeet (hft)

d= Travel distance in nautical miles (NM)

$perf^{ac}$= Performance of aircraft engine (% compared to a new model aircraft)

mac^{zfw}= Mean aerodynamic chord at zero fuel weight (%)

mac^{tow}= Mean aerodynamic chord at takeoff weight (%)

mac^{law}= Mean aerodynamic chord at landing weight (%)

v= Speed of aircraft in cost index (CI)

rwy = Runway direction (a number between 01 and 36, which is generally the magnetic azimuth of the runway's heading in decadegrees)

$\hat{\boldsymbol{Y}}$ = Fuel estimated for flight in Kilogram (Kg)

$\boldsymbol{W}$ = Input connection weights connecting input units (operational parameters) to hidden units

b = Bias for hidden units (+1 value)

$\boldsymbol{\beta}$ = Output connection weight connecting hidden units to the output unit (fuel estimation)

$\emptyset$ = Nonlinear sigmoid activation function

e = Predefined target error

$RMSE$ = Root mean square error between estimated fuel and actual consumed fuel (Kg)

$MAPE$ = Mean absolute percent error between estimated fuel and actual consumed fuel (%)

SR_H^{USI}= Sector route from departure airport-one at the United States of America and arrival airport at Hong Kong

SR_H^D= Sector route from departure airport at Dubai and arrival airport at Hong Kong

SR_{US1}^H= Sector route from departure airport at Hong Kong and arrival airport-one at the United States of America

SR_{US2}^H= Sector route from departure airport at Hong Kong and arrival airport-two at the United States of America

SR_{UK}^H= Sector route from departure airport at Hong Kong and arrival airport at the United Kingdom

SR_H^{US2}= Sector route from departure airport-two at United States of America and arrival airport at Hong Kong

SR_H^{UK} = Sector route from departure airport at the United Kingdom and arrival airport at Hong Kong

SR_H^{A}= Sector route from departure airport at Australia and arrival airport at Hong Kong

The objective functions (4) or (5) correspond to fuel deviation and their purpose is to achieve the smallest deviation by estimating a value of $\widehat{\boldsymbol{Y}}$ that approximately represents $\boldsymbol{Y}$. The RMSE objective function (4) measures the performance of the estimation method by determining the difference between the estimated $\widehat{\boldsymbol{Y}}$ and the actual consumed $\boldsymbol{Y}$. The measurement scale is identical to that of $\boldsymbol{Y}$**.** The RMSE tends to give more importance to a higher deviation by computing its square error. Actually, this is a more useful metric for comparison given that a higher deviation is undesirable. The MAPE objective function (5) measures the performance in percentage accuracy by determining the absolute difference between the estimated $\widehat{\boldsymbol{Y}}$ and actual consumed $\boldsymbol{Y}$. The percentage accuracy helps to make a more informed decision by comparing the deviation of the estimation methods on an absolute scale. In the NN, the value of $\widehat{\boldsymbol{Y}}$, as shown in Equation (6), can be estimated by determining coefficients of connection weights $\boldsymbol{W}$ and $\boldsymbol{\beta}$. Hidden units can be generated in-between connection weights by using a nonlinear sigmoid activation function $\emptyset$, as shown in Equation (7), to truly approximate $\boldsymbol{Y}$ from the complex nonlinear $\boldsymbol{X}$.

4. Solution Approach

Feedforward NNs following the universal approximation theorem may map any complex nonlinear function more accurately compared to other statistical parametric methods (Ferrari and Stengel, 2005; Hornik et al., 1989; Kumar et al., 1995). The rapidly growing interest in NNs (Au et al., 2008; Lin and Vlachos, 2018; Ruiz-Aguilar et al., 2014), and particularly in CNNs (Chung et al., 2017), motivates the study of their application to the airline sector. CNNs are considered to be more powerful compared to standard fixed NNs (Hunter et al., 2012; Wilamowski et al., 2008). Next, we briefly explain the existing fixed-topology BPNN and its limitations, followed by the proposed CNN.

4.1. BPNN algorithm and its limitations

A BPNN is defined as a network that does not form a cycle. All the connection signals exist from the input unit $\boldsymbol{X}$ to the output unit $\boldsymbol{Y}$ only in forward direction. Note that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are connected through a number of hidden units $\boldsymbol{H}$ by a series of connection weights without any cycle or loop (Hecht-Nielsen, 1989). Fig. 6 illustrates a simple BPNN with three layers. The number of hidden units and layers in the BPNN determines its topology structure. A BPNN with one hidden layer is known as shallow type, whereas a BPNN featuring more than one hidden layer is known as deep type (Bianchini and Scarselli, 2014; LeCun et al., 2015). The weights $\boldsymbol{W}$ connecting $\boldsymbol{X}$ to $\boldsymbol{H}$ are known as input connection weights whereas the weights $\boldsymbol{\beta}$ connecting $\boldsymbol{H}$ to $\boldsymbol{Y}$ are known as output connection weights. A BPNN is initialised by randomly defined hyperparameters, for instance, learning rate, connection weights, hidden units, and hidden layers. Training a model involves two steps, i.e., forward propagation and backward propagation. In the forward propagation step, the output $\widehat{\boldsymbol{Y}}$ is estimated by taking $\emptyset$ from the product of $\boldsymbol{H}$ and $\boldsymbol{\beta}$. Suppose $\boldsymbol{X}$ be $m\times n$, $\boldsymbol{Y}$ be $m\times q$, $\boldsymbol{H}$ be $m\times p$, $\boldsymbol{W}$ be $n\times p$, and $\boldsymbol{\beta}$ be $p\times q$ matrices. The estimation of output $\widehat{\boldsymbol{Y}}$ can be expressed as

$$\widehat{\boldsymbol{Y}} = \emptyset(\boldsymbol{H}\boldsymbol{\beta}) \tag{8}$$

where

$$\boldsymbol{H} = \emptyset(\boldsymbol{X}\boldsymbol{W} + b)\boldsymbol{.} \tag{9}$$

The objective function is the sum-of-squares error (SSE)

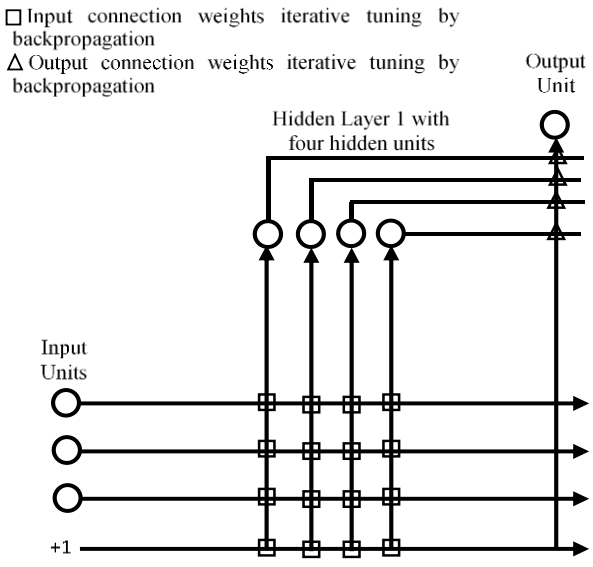

Fig. 6. A simple illustration of a BPNN architecture with four hidden units initialized in the hidden layer (shallow-type structure)

$$SSE = E = \sum_{i=1}^{m} \left(\hat{\boldsymbol{Y}}_i - \boldsymbol{Y}_i\right)^2. \quad (10)$$

If $E < e$, where e is the predefined target error, the algorithm stops. Otherwise, a backpropagation (BP) iterative learning algorithm is applied in the backward propagation step to train the model. The BP learning algorithm works on either first-order or second-order gradient information. In its simplest form, for a first-order derivative, the gradient error E with respect to the connection weights can be calculated by the chain rule:

$$\frac{\partial E}{\partial \boldsymbol{\beta}_i} = \frac{\partial E}{\partial out_i} \frac{\partial out_i}{\partial net_i} \frac{\partial net_i}{\partial \boldsymbol{\beta}_i} \quad (11)$$

where $\frac{\partial E}{\partial \boldsymbol{\beta}_i}$ is a partial derivative of error with respect to $\boldsymbol{\beta}$ for iteration i, out_i is the activation output, and net_i is the weighted sum of the inputs of $\boldsymbol{H}$.

Thus, for updating the output connection weight $\boldsymbol{\beta}_{i+1}$, we have

$$\boldsymbol{\beta}_{i+1} = \boldsymbol{\beta}_i - \eta \frac{\partial E}{\partial \boldsymbol{\beta}_i} \quad (12)$$

Similarly, for updating the input connection weight $\boldsymbol{W}_{i+1}$, we apply

$$\boldsymbol{W}_{i+1} = \boldsymbol{W}_i - \eta \frac{\partial E}{\partial \boldsymbol{W}_i} \quad (13)$$

A similar procedure is followed for second-order derivative learning algorithms. The updated new connection weights are forward propagated and E is recalculated. The connection weights are trained after each iteration to obtain a minimum gradient error. When E converges or starts increasing, the algorithm is stopped.

The major drawbacks of BPNNs limiting their applicability are their weak generalization performance and learning speed. The generalization performance is weak in the sense that it may stop at local minima of the function if global minima are far away. The learning speed (also known as convergence rate) is slow and dependent on the learning rate and BP learning algorithms. When the learning rate is small, it converges slowly, and when the learning rate is large, it becomes unstable. The BP learning can be considered time-consuming because of the repetitive tuning of the connection weights in forward and backward steps with other hyperparameters (Huang et al., 2006b). This may create a complex co-adaptation among the parameters and hidden units that demands further attention in regularisation (Srivastava et al., 2014). Moreover, with respect to the drawbacks mentioned above, the following human expertise significantly affects the generalization performance and learning speed of BPNNs:

a) What should be the structure of the network? For instance, shallow or deep?
b) How many hidden layers should be defined in a deep learning network?
c) How many hidden units should exist in each hidden layer?
d) What should be the initial guess of connection weights and learning rates?
e) Which BP learning algorithm will be more suitable? Stochastic gradient descent, Levenberg Marquart, quick prop…

The drawbacks and expertise issues need to be solved in a more innovative way to improve the generalization performance and learning speed. To overcome the above limitations, we propose a self-organising cascade topology-based CNN capable of analytically calculating connection weights and requiring less expertise by eliminating the need for complex hyperparameter initialisation and adjustment. The objective is to improve the performance of the estimation model with less expertise involvement and obtain results with fast learning speed.

4.2. Proposed CNN

To address the learning issues of BPNNs, the so-called cascade correlation learning algorithm (CasCor) was proposed to add hidden units sequentially to the network (Fahlman and Lebiere, 1990). CasCor begins with the minimal network by linearly connecting $\boldsymbol{X}$ to $\boldsymbol{Y}$ through randomly generated $\boldsymbol{\beta}$. The values of $\boldsymbol{\beta}$ are iteratively tuned (or trained) by BP quick prop (QP) to minimise E. When training converges and E is greater than e, it adds $\boldsymbol{H_k}$ $(k = 1,2, \ldots, l)$ one at a time to the network, which receives randomly generated $\boldsymbol{W}$ from all $\boldsymbol{X}$ and any pre-existing $\boldsymbol{H_{k-1}}$. The $\boldsymbol{H_k}$ is not yet connected to $\boldsymbol{Y}$. The values of $\boldsymbol{W}$ are iteratively tuned to maximise the covariance objective function between $\boldsymbol{H_k}$ and E. When the covariance objective function converges, $\boldsymbol{H_k}$ is added to $\boldsymbol{Y}$ by freezing $\boldsymbol{W}$, and $\boldsymbol{\beta}$ is once again iteratively tuned by QP. This process continues and finally stops when E converges or starts increasing. The QP quickly reaches the loss function by taking a much larger step rather than infinitesimal steps (Fahlman, 1988). The advantage of CasCor over fixed-topology BPNNs is that it can learn complex tasks more quickly as well as determine its own network topology. In addition, it is economical and has no underfitting issue.

The growing applications of CNN-based algorithms concluded that CasCor learning speed is better than that of BPNNs, but its generalization performance might be not optimal. Hwang et al. (1996) and Lehtokangas (2000) highlighted that CasCor works better with classification tasks compared to regression tasks and this may make it unsuitable for some applications. Liang and Dai (1998) made an effort to improve the generalization performance of CasCor on classification problems by using a genetic algorithm, at the additional cost of requiring more training time. Kovalishyn et al. (1998) studied the application of CasCor on the quantitative structural activity relationship and did not achieve any dramatical increase in performance compared to BPNNs. The reasons may be (i) the chaotic behaviour and

numerical instability of the iterative QP learning algorithm and (ii) no guarantee of maximum error reduction by the covariance objective function when a new hidden unit is added.

Firstly, QP during iterative tuning takes a much larger step to move towards the loss function based on information of past and current gradient. If the current gradient is in the opposite direction from the past gradient, the QP may cross the minimum of the loss function and moves towards the opposite direction of the valley, from where it needs to come back (Fahlman, 1988). This may cause the QP to act chaotically across the minimum valley of the loss function. Banerjee et al. (2011) explained that QP becomes numerically unstable if the current iteration gradient becomes closer or equal to the previous iteration gradient. In such a case, the weight difference will become zero and the QP formula will remain zero even if the gradient changes.

Secondly, Huang et al. (2012) argued that the CasCor covariance objective function may not guarantee a maximum error reduction when a new hidden unit is added. In addition, the repeatedly iterating tuning of $\boldsymbol{\beta}$ can be more time-consuming. They proposed an orthogonal least squares based cascade network (OLSCN) by reformulating the objective function based on ordinary least squares for training of $\boldsymbol{W}$, which needs to be further optimised by the second-order Newton's method. Qiao et al. (2016) explained that updating weights by the Newton's method may cause OLSCN to converge at a local minimum with an increase in computational burden. They proposed a faster cascade neural network (FCNN) to add a contribution to the CNN. The FCNN works by selecting linearly independent instances of $\boldsymbol{X}$ by the Gram–Schmidt orthogonalisation method and a hidden unit from a pool of candidate units by a modified index method.

Motivated by the cascade architecture of CasCor, we propose a new CNN by improving an original CasCor and its variants. The key idea is to analytically determine connection weights on both sides of the network rather than iterative tuning with modified cascade architecture. We propose to analytically determine $\boldsymbol{W}$ by the orthogonal linear transformation of $\boldsymbol{X}$, whereas $\boldsymbol{\beta}$ by the ordinary least squares method. The proposed CNN algorithm is named cascade principal component least squares (CPCLS) neural network. The architecture of CPCLS is illustrated in Fig. 7.

Given $\boldsymbol{X}$ and $\boldsymbol{Y}$, CPCLS is initialised by defining e and hidden units N of $\boldsymbol{H}$ such that $p \leq n$. For *input connection weight* $\boldsymbol{W}$ determination, it transforms the set of correlated $\boldsymbol{X}$ n-features orthogonally and linearly into uncorrelated $\boldsymbol{H}$ p-features by eigen-decomposition of the $\boldsymbol{X}$ covariance square $n \times n$ matrix $\boldsymbol{S}$:

$$\boldsymbol{S} = \frac{1}{m-1}(\boldsymbol{X} - \bar{\boldsymbol{x}})^T(\boldsymbol{X} - \bar{\boldsymbol{x}}) \tag{14}$$

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{15}$$

The highest eigenvalue λ generated from $\boldsymbol{S}$ corresponds to an eigenvector. These selected N eigenvectors are considered as the input connection weight, i.e., $\boldsymbol{W}$:

$$|\boldsymbol{S} - \lambda \boldsymbol{I}| = 0 \tag{16}$$

$$(\boldsymbol{S} - \lambda \boldsymbol{I})\boldsymbol{W} = 0 \tag{17}$$

We calculate $\boldsymbol{H}$ by taking the nonlinear activation $\emptyset$ from the product of $\boldsymbol{X}$ and $\boldsymbol{W}$ by adding a bias b:

$$\boldsymbol{H} = \emptyset(\boldsymbol{X}\boldsymbol{W} + b) \tag{18}$$

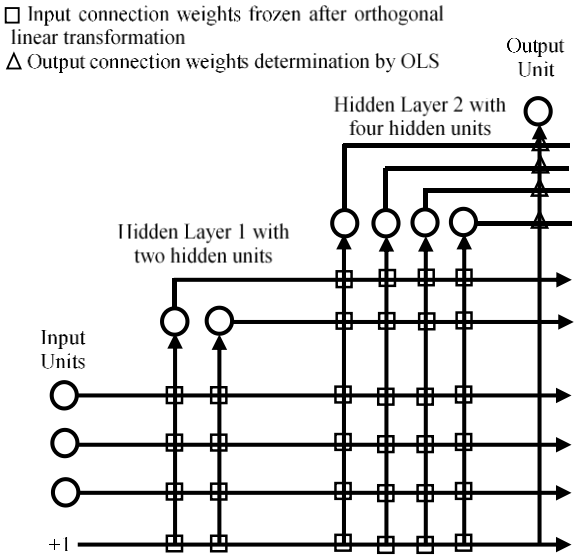

Fig. 7. CPCLS architecture with input units, two hidden layers, and output units.

The orthogonal linear transformation by eigen-decomposition of $\boldsymbol{S}$ generates $\boldsymbol{H}$, which explains the maximum variance in the data, what in turn can help to estimate trip fuel more accurately. The maximum variance in $\boldsymbol{H}$ becomes more linear in a relationship with $\boldsymbol{Y}$, and SSE can be minimised through ordinary least squares (OLS) by calculating for *output connection weight, i.e.,* $\boldsymbol{\beta}$:

$$\boldsymbol{\beta} = (\boldsymbol{H^T H})^{-1}\boldsymbol{H^T Y} \tag{19}$$

where $(\boldsymbol{H^T H})^{-1}\boldsymbol{H^T}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{H}$. We estimate the trip fuel $\hat{\boldsymbol{Y}}$ by linearly transferring $\boldsymbol{H}$ (18) through $\boldsymbol{\beta}$ (19):

$$\hat{\boldsymbol{Y}} = \boldsymbol{H\beta} \tag{20}$$

We calculate fuel deviation in (4) or (5) by subtracting the actual consumed fuel $\boldsymbol{Y}$ from the estimated trip fuel $\hat{\boldsymbol{Y}}$.

If (4) or (5) is less than e, we stop; otherwise, another hidden layer $\boldsymbol{H_k}$ with N' hidden units is added with respect to the previously added hidden layer $\boldsymbol{H_{k-1}}$ having N hidden units until the desired performance is achieved. The newly added hidden layer $\boldsymbol{H_k}$ will receive all the connections from $\boldsymbol{X}$ and from any pre-existing hidden layers, i.e., $\boldsymbol{H_{k-1}}$. In other words, the previously calculated $\boldsymbol{H_{k-1}}$ becomes a part of $\boldsymbol{X}$, such that

$$\boldsymbol{X} = (\boldsymbol{X}, \boldsymbol{H_{k-1}}) \tag{21}$$

$$N = N + N' \tag{22}$$

The steps from Equation (14) to (20) are repeated.

Equation (21) implies that only newly added $\boldsymbol{H_k}$ will connect to $\boldsymbol{Y}$ and will eliminate previous connections, i.e., $\boldsymbol{H_{k-1}}$, $\boldsymbol{H_{k-2}}$, etc. This helps to avoid linear dependencies among hidden units in the multiple hidden layers. For better illustration, consider an example in Fig. 7, where the hidden layers are initialised with N=2 and N'=2. It implies that the first hidden layer (hidden layer 1) is initialised with N=2 hidden units and the algorithm is then trained. If the error resulting from training is larger than e, then another hidden layer (hidden layer 2) is added with N'=2 hidden units with respect to hidden units of previous layers ($N = N + N'$ = 2+2= 4 hidden units in the hidden layer 2). Again, the algorithm is trained, and the error is calculated by connecting only a newly added hidden layer (hidden layer 2) to the output unit and by eliminating a previous hidden layer (hidden layer 1) from the output connections weight. These steps continue until the targeted value of e is achieved.

The advantages of CPCLS that makes it superior compared to CasCor and its variants is the generation of non-redundant and linearly independent multiple hidden units in the layers. The linearly independent hidden units generated by an orthogonal linear transformation of the operational parameters help to achieve the best least squares solution. Similarly, to avoid the linear dependency among hidden layers, a cascade architecture is improved by connecting only the newly added hidden layer to the output, with all previous connections being eliminated. These characteristics help CPCLS to converge faster and ensure maximum error reduction when the new hidden layer is added. The achievement of estimation results in the shortest possible computational time along with minimum hyperparameter initialisation, such as $\boldsymbol{H}$ and e, makes CPCLS more favourable. Other properties include no iterative tuning, no trial-and-error to find the best hyperparameters, and less human intervention. The effectiveness of the CPCLS algorithm compared to AEA and BPNN is demonstrated in the following experimental work section.

5. Experimental Work

To estimate a trip fuel with less deviation, requiring less expertise, and achieving faster learning speed, real high-dimensional data were provided by an airline that historically performed international flights in various regions over two years. The dataset consists of several operational parameters that may or may not contribute to fuel consumption. For better estimation, the useful operational parameters were first extracted, as reported in Table 1. In our study, the extracted operational parameters are represented by $\boldsymbol{X}$ and the objective is to predict the trip fuel $\hat{\boldsymbol{Y}}$ that truly represents the actual consumed trip fuel $\boldsymbol{Y}$. Experimental work was performed by considering two cases. In the first case, all sectors in the dataset were considered as a combined model. In the second case, each sector was considered as an individual model. A sector is defined as an airway route starting from the departure airport and ending in the arrival airport. Three estimation methods were used for the comparative study, namely AEA, BPNN, and the proposed CPCLS. For AEA, the estimated and actual consumed fuel information were retrieved from available data. CPCLS is compared to AEA in terms of fuel deviation, whereas its comparison with BPNN is in terms of both fuel deviation and learning speed. All the experimental work was carried out in Anaconda Spyder Python v3.2.6 programming language. The dataset was normalised in the range [0,1]. One hot encoding was performed on categorical variables to represent it in a binary vector. The nonlinear sigmoid activation function was used in the hidden layers to make it compatible with the dataset. For both cases, the dataset was split into a 50:50 ratio for training and testing. The stopping criteria for the algorithms were defined as error convergence or start increasing.

5.1. Data description

The set of real data that we obtained from an international airline operating in Hong Kong consists of 19,117 international passenger and cargo flights operated among eight sectors covering different geographical regions. The flights were performed from April 2015 to March 2017 by 107 wide-body aircrafts distributed as Airbus A330-300 (31 aircrafts) and Boeing 747-400 (9 aircrafts)/747-800 (14 aircrafts)/777-300 (53 aircrafts). The data consist of detailed information on operational parameters related to 1) carrier with aircraft performance and operations, 2) date/time of departure and arrival per sector with crew allocation, 3) weather and atmospheric conditions, 4) sector and aircraft local and global conditions, and 5) fuel requirements at different tanks and consumption levels. Among the collected data, instead

of putting all operational parameters into algorithms, we work in a more novel way by extracting the relevant operational parameters contributing to fuel consumption, as summarised in Table 1.

5.2. Theoretical relationship between extracted operational parameters and fuel consumption

The relationship between each extracted operational parameter and fuel consumption is clear: *flight time* and *distance travel* are directly related to fuel consumption because the more the aircraft is airborne, the more the fuel it will consume. Flight duration and covered distance significantly contribute to aircraft operational expenses. The *cost index (CI)* is used to adjust the speed of aircraft to a trade-off between higher operational expenses or more fuel saving (Edwards et al., 2016). The operational expenses can include crew time, leasing rate, plan maintenance, landing and take-off fees, and ground services. All of them play major roles in deciding whether to keep the aircraft airborne to make the airline profitable. The CI ranges from 0 to 999 and is expressed as

$$CI = \frac{Flying\ Cost}{Fuel\ Cost} \tag{23}$$

When fuel is expensive, the CI will be lower, which means a slower speed of the aircraft. Operating an aircraft at low speed results in a higher climb rate due to excessive engine thrust and, simultaneously, it is recommended to fly at a higher altitude to lower fuel consumption. Similarly, when fuel is cheap, the CI will be higher, and a faster aircraft will operate to be less airborne to reduce the amount of operational cost at the cost of more consumed fuel. Note that the CI shortens or lengthens the airborne phase by changing the aircraft speed depending on fuel prices and operational expenses to cut overall expenses.

The *ramp weight* (m) can be expressed according to the following combination (Pagoni and Psaraki-Kalouptsidi, 2017):

$$m = zfw + tof \tag{24}$$

Zero fuel weight (zfw) is the total weight of an aircraft including crew, passengers, and unusable fuel, minus the total weight of the usable fuel. During flight operation, the zfw remains constant whereas tof weight decreases over time because of its continuous consumption. This reduces the value of m for an aircraft and ultimately decreases fuel consumption over time.

A suitable choice of *wind direction* has a significant influence on fuel consumption. Headwind that blows in a direction opposite to the aircraft is more favourable during takeoff and landing. The aerofoils can generate more lift during takeoff and more induced drag during landing. Tailwind that blows in the direction of the aircraft flight path is helpful during the cruise phase as the aircraft travels faster and saves fuel at the same ground speed (Irrgang et al., 2015). The ground speed is determined by the vector sum of aircraft speed, wind speed, and direction. Headwind subtracts from the ground speed, while tailwind is added to the ground speed.

The *altitude* can be used to assess the aerodynamic performance of an aircraft at certain atmospheric conditions. Flying at higher altitude can significantly save fuel consumption because of less drag (Turgut and Rosen, 2012). Simultaneously, it may face a problem of low-density air for fuel combustion that may result in more fuel flow to the engine. Therefore, the altitude is adjusted for a non-standard temperature known as density altitude and can be expressed as

$$DA = PA + (118.8\ ft/\ ^{o}\mathrm{C})(OAT - ISA\ Temp) \tag{25}$$

$$PA = h + (30ft/millibar)(1013millibar - QNH) \tag{26}$$

where DA is the density altitude in feet, PA is the pressure altitude in feet, QNH is the atmospheric pressure in millibar, OAT is the outside air temperature in degree Celsius, and $ISA\ Temp$ is the international standard atmospheric temperature in degree Celsius. Aircrafts are designed for specific optimal altitude, what minimises fuel consumption. A change in that altitude may cause aircrafts to burn more fuel (Diao and Chen, 2018).

The weight and balance of aircrafts can be expressed in terms of the percentage of the *mean aerodynamic chord* (Dancila et al., 2013).

$$MAC\% = \frac{\left(\frac{\left(\sum_{i=1}^{N} w_i a_i\right)}{m} - LEMAC\right)}{TEMAC - LEMAC} \tag{27}$$

where $\frac{\left(\sum_{i=1}^{N} w_i a_i\right)}{m}$ is defined as the centre of gravity (CG) with w denoting component weights and a denoting the arm value, defined both as the moment arm (wa). $LEMAC$ and $TEMAC$ are the leading-edge mean aerodynamic chord and tail-edge mean aerodynamic chord, respectively. Improper distribution of aircraft weight may shift its CG forward and may need more tail lift force for stable flight. The tail force ultimately increases the aircraft angle of attack, what may cause aircrafts to face more induced drag. Therefore, the aircraft weight balance (at the centre of aircraft) should be maintained at zero fuel weight during takeoff and landing phases to avoid generation of excess induced drag and reduce fuel consumption.

The *location and direction of the runway* have clear effects on fuel consumption (Singh and Sharma, 2015). Runway directions are identified by a number within the range 1 to 36, which are magnetic azimuths in decadegrees. A runway number 09 means pointing toward the east (90°), 18 means pointing toward the south (180°), 27 means pointing toward the west (270°), and 36 means pointing toward the north (360°). Furthermore, if there is more than one runway in the same direction, then the location is differentiated with a letter as follows: L for left, C for centre, and R for right. It is more favourable to take off from a runway that faces the headwind direction. However, it can

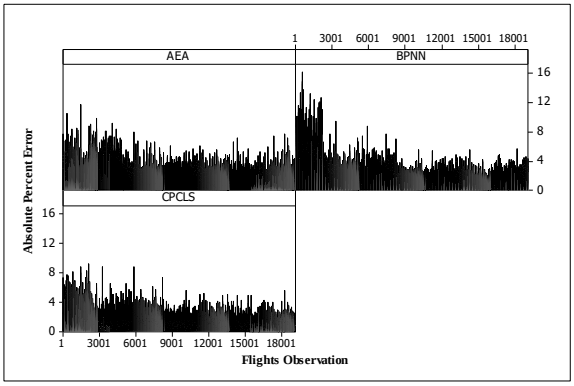

Fig. 8. Fuel deviation (combined model)

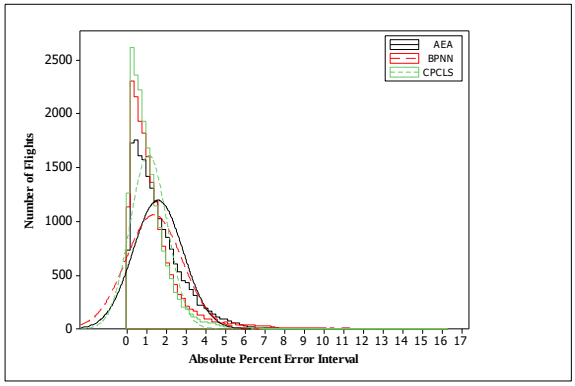

Fig. 9. Fuel deviation interval (combined model)

also be a source of excess fuel consumption. If that favourable runway is in the opposite direction to the allocated flight path, the aircraft must take a loop to head towards the allocated path.

5.3. Results of trip fuel estimation models and discussion

The main objective of our study is to develop a method capable of accurately and efficiently estimating trip fuel; however, a comparison with existing methods plays an important role in demonstrating effectiveness. To study the effect of operational parameters, the fuel is estimated for the following two types of models:

a) All sectors combined for fuel estimation (referred to as a combined model)
b) Sector-wise fuel estimation (referred to as individual models)

a) Combined model:

In this model, the trip fuel is estimated by considering all eight sectors of the 19,117 flights as a combined model. The selection of operational parameters and their data trend play a major role in examining and analysing the estimation method. The significant difference in operational parameter ranges and standard deviation may make estimating trip fuel accurately a challenging task. The runway direction parameter is a categorical variable with 26 different runway directions for eight sectors. One hot encoding pre-processing technique is applied on each runway direction to represent it in a binary vector. It translates runway direction features into 26 subfeatures as a binary value with 1 representing existence and 0 representing no existence. All the flight phases such as takeoff, climb, cruise, descent, and approach are considered as one entire flight trip. The CPCLS is initialised with 5 hidden units in the first layer and 15 added hidden units in the next hidden layers with respect to previously added hidden units. Among different trial-and-error approaches, BPNN is selected, with 10 hidden units and 0.5 learning rate as the best hyperparameters for the network.

Table 2 shows the trip fuel deviation resulting from AEA, BPNN, and CPCLS. The values of Mean, StDev, Min, Q1, Med, Q3, Max, Ran, and IQR under the column absolute percent error (APE) respectively refer to the mean, standard deviation, minimum, first quartile, median, third quartile, maximum, range, and interquartile range in percentage; RMSE refers to root mean-squares error; time refers to the computational learning time in seconds needed for estimation. For the sake of readability, the mean of APE is termed as MAPE. Descriptive statistics are employed to understand and compare the central tendency and the variability/dispersion of the fuel deviation estimated by the methods. Mean measures the central tendency by computing an average value and StDev measures the dispersion by computing how much data is far from the mean. Both measurements are important to compare the results generated by the studied methods. The method with less mean and StDev is favourable because it indicates that the model has estimated fuel with less deviation and dispersion. The disadvantages of mean and Stdev are that they are only useful to measure the central tendency and dispersion of the whole dataset. It might occur that some portion of data may exhibit a higher improvement compared to others. Therefore, to make the comparison reliable for a better conclusion of the findings, improvements in various quartiles (Q1, Med, Q3, IQR) along with the minimum (min) and maximum (max) fuel deviation for each method are also studied. In Table 2, the first finding is that CPCLS estimates trip fuel with a fuel deviation

Table 2:
Trip fuel deviation by AEA, BPNN, and CPCLS (combined model)

Sector	**Method**	**APE**									**RMSE**	**TIME (s)**
		Mean	**StDev**	**Min**	**Q1**	**Med**	**Q3**	**Max**	**Ran**	**IQR**		
All sectors	AEA	1.53	1.28	0.00	0.57	1.21	2.16	11.66	11.66	1.59	0.0183	---
	BPNN	1.32	1.44	0.00	0.42	0.92	1.68	16.14	16.14	1.26	0.0135	745.00
	CPCLS	**1.05**	**0.95**	0.00	**0.37**	**0.81**	**1.44**	**9.31**	**9.31**	**1.07**	**0.0115**	**8.21**

Table 3:
Performance improvement comparison (combined model)

Sector	AEA/BPNN (%)	AEA/CPCLS (%)	BPNN/CPCLS (%)
All Sectors	15.91	45.71	25.71

of 1.05% compared to AEA (1.53%) and BPNN (1.32%). A CPCLS Stdev of 0.95% implies that data are more clustered towards the mean compared to AEA (1.28%) and BPNN (1.44%). Regarding convergence rates, CPCLS trains a model in 8.21 s whereas BPNN in 745 s. The computational time for AEA is not mentioned because the information is not available in the historical dataset provided by the airline. The fuel deviation results are split into Q1 (25th percentile), Q2 or Median (50th percentile), and Q3 (75th percentile) to compare the improvement of estimation methods in different quartiles. IQR, not affected by outliers and often considered a better measurement of a spread than range, describes the 50% of values by taking the difference between Q1 and Q3. For each quartile and interquartile, the second finding is that the improvement of CPCLS is evident compared to AEA and BPNN. For Q1, median, Q3, and IQR, the fuel estimated by CPCLS features a mean deviation of 0.37%, 0.81%, 1.44%, and 1.07% compared to AEA (0.57%, 1.21%, 2.16%, and 1.59%, respectively) and BPNN (0.42%, 0.92%, 1.68%, and 1.26%, respectively). The maximum fuel deviation predicted by CPCLS and AEA is 9.31% and 11.66% respectively, whereas for BPNN it is 16.14%, which make it unfavourable for this model. Although the overall BPNN MAPE is better than that of AEA, what validates the importance and significance of the BPNN, the third finding is that the maximum deviation approaches 16.14% compared to AEA (11.66%). This may create a problem in a real application by a wrong estimation of fuel for certain flights. For ease of readability, the lowest and best values for each metric is highlighted in bold with an underline in Table 2.

Table 3 shows the performance improvement comparison in percentage in terms of MAPE. Thus, AEA/BPNN, AEA/CPCLS, and BPNN/CPCLS refer to a gain in the improvement of BPNN compared to AEA, CPCLS compared to AEA, and CPCLS compared to BPNN, respectively. The fourth finding is that the CPCLS-based estimation model shows an improvement of 45.71% and 25.71% compared to both AEA and BPNN, respectively, whereas BPNN shows an improvement of only 15.91% compared to AEA.

Fig. 8 shows the AEA, BPNN, and CPCLS fuel deviation in APE for each flight. The horizontal axis represents a number of operated flights whereas the vertical axis is APE for each flight. The comparison illustrates that the fuel estimated from CPCLS for each flight exhibits less fuel deviation compared to AEA and BPNN. Similarly, Fig. 9 shows the histogram and distribution comparison of AEA, BPNN, and CPCLS. The horizontal axis represents an APE interval (bins) whereas the vertical axis represents a number of flights for each interval. As shown in Fig. 9, the fifth finding is that CPCLS shifts the flight trend more towards the lower fuel deviation interval. For instance, the histograms show that CPCLS estimated the fuel deviation of 6,236 flights in the range 0.00-0.50%, what demonstrates an improvement of 47.77% and 11.35% compared to the 4,220 flights from AEA and 5,600 flights from BPNN, respectively. For other histogram bins, the improvement shown by CPCLS is more significant. In addition, Fig. 9 indicates the sixth finding, i.e., that CPCLS has more cluster distribution (lower StDev) compared to disperse distribution of AEA and BPNN. However, Fig. 8 shows that CPCLS and BPNN estimated a number of 2,000 flights with less improvement in fuel deviation as compared to the remaining flights. An in-depth analysis shows our seventh finding, i.e., these flights are operated in SR_H^D and SR_H^A sectors, which belong to short-range flights, whereas the remaining flights are operated in the SR_H^{USI}, SR_{US1}^H, SR_{US2}^H, SR_{UK}^H, SR_H^{US2}, and SR_H^{UK} sectors and belong to long-range flights on long-haul routes. The large portion of data belongs to the long-range route flights. This forces NN-based algorithms to determine network connection weights by considering long-range route flights rather than short-range route flights. This is the main reason for which it is hard for BPNNs to truly approximate the complex operational parameters and their data instances. It can be easily visualised from Fig. 8 that the BPNN estimation for the mentioned 2000 flights is even worse than both AEA and CPCLS. Therefore, to study the effect of the short-range and long-range route on estimation models, our work is extended by estimating the trip fuel for each sector individually. The purpose of sector-wise fuel estimation (individual model) is to study the effect of the extracted operational parameters of fuel consumed within each sector.

b) Individual models

The types of operational parameters and their data instances have a significant effect on fuel consumption. For example, an aircraft traveling from airport A to airport B might face headwinds, and during the return, it might face tailwinds. The difference in both conditions may cause different fuel consumption for the same route. The real data consist of 2658, 1252, 3080, 2493, 3199, 2423, 3116, and 896 flights operated in sectors SR_H^{USI}, SR_H^D, SR_{US1}^H, SR_{US2}^H, SR_{UK}^H, SR_H^{US2}, SR_H^{UK}, and SR_H^A, respectively. The aircraft flying in sectors SR_H^{USI}, SR_{UK}^H, SR_H^{US2}, SR_H^A faced headwind compared to sectors SR_H^D, SR_{US1}^H, SR_{US2}^H, which faced tailwind. High-performance aircrafts are operated in sectors SR_{US2}^H, SR_{UK}^H, SR_H^{US2}, and SR_H^{UK}. The short-range route flights in sectors SR_H^D and SR_H^A are operated at a lower altitude level compared to long-range route flights. The flight time varies according to an increase in distance for all sectors, whereas the mean aerodynamic chord and aircraft speed are more dependent on aircraft weights. The runway directions for SR_H^{USI} are 07, 15, 25, and 33; for SR_H^D, are 12 and 30; for SR_{US1}^H are 07 and 25; for SR_{US2}^H are 07 and 25; for SR_{UK}^H are 07 and 25; for SR_H^{US2} are 06, 07, 24 and 25; for SR_H^{UK} are 09 and 27; and for SR_H^A are 03, 06, 21 and 24. Location of more than one runway with same directions are differentiated with a letter as follows: L for left, C for centre, and R for right. A one hot encoding pre-processing technique is applied sector-wise on the runway direction categorical variable to convert it to a binary vector. This varying behaviour of operational parameters may significantly influence the accuracy of machine learning algorithms.

Table 4 shows the sector-wise trip fuel deviation results for AEA, BPNN, and CPCLS. The first finding is that the fuel estimation by CPCLS produces fuel deviation MAPEs of 1.11%, 1.30%, 0.95%, 0.70%, 0.80%, 0.76%, 0.88%, and 1.03% compared to AEA fuel deviation MAPEs of 2.84%, 2.73%, 1.18%, 0.99%, 1.36%, 1.27%, 1.12%, and 1.52%, and that of BPNN, i.e., 1.21%, 1.38%, 1.13%, 0.85%, 0.90%, 0.78%, 0.93%, and 1.08% for sectors SR_H^{USI}, SR_H^D, SR_{US1}^H, SR_{US2}^H, SR_{UK}^H, SR_H^{US2}, SR_H^{UK} and SR_H^A, respectively. StDev measurements show that CPCLS dispersion from mean is less for all sectors compared to AEA and BPNN. Apart from sector SR_H^A, the second finding is that the quartile and interquartile measurements for all other sectors indicate that CPCLS achieves better estimation for all quartiles compared to AEA and BPNN. BPNN achieves a better estimation of 0.39% in Q1 for SR_H^A; however, this estimation is very close to 0.40% from CPCLS. Moreover, the maximum fuel deviation estimated by CPCLS is lessen than those of AEA and BPNN. For

SR_H^{USI}, BPNN was able to achieve a better maximum fuel deviation value of 9.50% compared to that of CPCLS (9.94%) and AEA (9.77%). An in-depth study shows that CPCLS was able to estimate SR_H^{USI} for all flights within an APE of 7%; however, only one flight shows as APE of 9.94%. The IQR, not effected by the outliers, explain this issue. The IQR measure of the CPCLS (1.17%) is significantly better compared to the BPNN (1.26%). The average network learning time of the CPCLS is 1.53s, which is hundred times faster than the BPNN of 104.19s. The total estimation time of CPCLS and BPNN for individual models is 12.22s and 833.52s which is approximately equal to the combined model of 8.21s and 745s respectively. The third finding is that the slightly increase in the time is because of a discovering trend in the individual dataset which were previously difficult for NN because the majority of the dataset belongs to the long-range route flights.

Table 5 shows the performance improvement comparison in percentage in term of MAPE. The AEA/BPNN, AEA/CPCLS, and BPNN/CPCLS refers to a gain in the improvement of the BPNN compared to the AEA, the CPCLS compared to the AEA, and the CPCLS compared to the BPNN respectively. The CPCLS shows maximum improvement of 155.86% for SR_H^{USI} and minimum improvement of 24.21% for SR_{US1}^H, with an average improvement of 67.93% compared to the AEA. The maximum and minimum improvements provide an important insight, fourth finding, about the influence of the operational parameters. In both improvements, the flight route is the same. The SR_H^{USI} sector is flight traveling from airport-one located in the United States to Hong Kong and SR_{US1}^H sector is flight travelling from Hong Kong to the airport-one in the United States. This explains that the existing AEA mathematical calculation cannot differentiate among the varying behaviour of operational parameters. The use of global parameters for the both sectors of the same route rather than local parameters with the old database in mathematical equations may give better estimation for SR_{US1}^H compared to SR_H^{USI}. The BPNN leads to an average 53.57% better estimation compared to the AEA but less accurate than the CPCLS. The better performance of the BPNN in sector-wise individual fuel estimation models 53.57% (Table 5) compared to combined fuel estimation model 15.91% (Table 3) suggests that it may perform better with a chunk of data comprising the same behaviour of the operational parameters. Comparing to the AEA, the CPCLS shows an improvement from 45.71% (combined model) to 67.93% by performing sector wise individual fuel estimation. The fifth finding is that the difference in improvement for the CPCLS is 22.22% as compared to the BPNN of 37.66%. This implies that the CPCLS estimation capability with high dimensional data is much better than the traditional BPNN.

Figs. 10-17 shows the fuel deviation in APE for each flight in the sectors by the AEA, BPNN, and CPCLS. A horizontal axis is a number of flights operated and the vertical axis is APE for each flight. The comparative study in the figures helps to understand the improvement made by the CPCLS compared to the AEA and BPNN. The sixth finding is that the CPCLS and BPNN were able to achieve better estimation for short-range route flights of SR_H^D and SR_H^A. In the combined model, for the first 2000 flights related to short-range route flights, the performance of the BPNN was much worse compared to the AEA, and for CPCLS the performance was better than AEA but not compared to others long-range route flights. In the current estimation models, Figs. 11 and 17, the BPNN achieves better estimation than the AEA, and the CPCLS achieves better estimation compared to both AEA and BPNN. Similarly, for all other sectors, Figs. 10, 12-16, the APE of each flights shows better results for the CPCLS compared to the AEA and BPNN. As mentioned earlier that sector SR_H^{USI} and SR_{US1}^H follow the same route but travel in the opposite direction. Comparative study of Fig 10 and 12, shows that the AEA estimated

Table 4
Trip fuel deviation by the AEA, BPNN and CPCLS (Individual Models)

Sector	Method	APE								RMSE	TIME (s)
		Mean	StDev	Min	Q1	Med	Q3	Max	Ran	IQR		
SR_H^{USI}	AEA	2.84	1.46	0.00	1.83	2.73	3.75	9.77	9.77	1.92	0.0362	---
	BPNN	1.21	1.03	0.00	0.45	0.95	1.71	**9.50**	**9.50**	1.26	0.0243	105.81
	CPCLS	**1.11**	**0.93**	0.00	**0.42**	**0.88**	**1.58**	9.94	9.93	**1.17**	**0.0225**	**1.87**
SR_H^D	AEA	2.73	1.69	0.00	1.37	2.57	3.91	10.40	10.40	2.54	0.0864	---
	BPNN	1.38	1.18	0.00	**0.50**	1.10	1.93	9.96	9.95	1.43	0.0326	23.83
	CPCLS	**1.30**	**1.09**	0.00	**0.50**	**1.04**	**1.82**	**9.49**	**9.49**	**1.32**	**0.0303**	**0.91**
SR_{US1}^H	AEA	1.18	0.96	0.00	0.45	0.95	1.64	7.95	7.95	1.19	0.0362	---
	BPNN	1.13	1.01	0.00	0.41	0.90	1.56	9.63	9.63	1.15	0.0314	91.17
	CPCLS	**0.95**	**0.82**	0.00	**0.35**	**0.75**	**1.33**	**8.88**	**8.88**	**0.98**	**0.0265**	**3.48**
SR_{US2}^H	AEA	0.99	0.83	0.00	0.38	0.80	1.38	7.10	7.10	1.00	0.0365	---
	BPNN	0.85	0.72	0.00	0.32	0.67	1.18	5.34	5.34	0.86	0.0324	92.32
	CPCLS	**0.70**	**0.55**	0.00	**0.28**	**0.56**	**0.98**	**4.53**	**4.52**	**0.70**	**0.0261**	**2.44**
SR_{UK}^H	AEA	1.36	0.98	0.00	0.59	1.19	1.96	6.09	6.09	1.37	0.0621	---
	BPNN	0.90	0.73	0.00	0.33	0.74	1.27	5.28	5.28	0.93	0.0410	176.73
	CPCLS	**0.80**	**0.64**	0.00	**0.31**	**0.67**	**1.15**	**5.81**	**5.81**	**0.84**	**0.0364**	**0.83**
SR_H^{US2}	AEA	1.27	0.97	0.00	0.53	1.08	1.79	7.63	7.63	1.26	0.0461	---
	BPNN	0.78	0.64	0.00	0.30	0.64	1.09	5.83	5.83	0.79	0.0319	124.18
	CPCLS	**0.76**	**0.63**	0.00	**0.29**	**0.61**	**1.06**	**5.24**	**5.24**	**0.77**	**0.0313**	**1.01**
SR_H^{UK}	AEA	1.12	0.90	0.00	0.46	0.89	1.56	6.62	6.62	1.11	0.0458	---
	BPNN	0.93	0.77	0.00	0.35	0.76	1.32	6.60	6.60	0.98	0.0352	147.00
	CPCLS	**0.88**	**0.71**	0.00	**0.32**	**0.73**	**1.24**	**4.33**	**4.33**	**0.92**	**0.0330**	**0.52**
SR_H^A	AEA	1.52	1.17	0.00	0.67	1.31	2.11	11.66	11.66	1.44	0.0600	---
	BPNN	1.08	0.97	0.00	**0.39**	0.83	1.46	10.13	10.12	1.07	0.0440	72.48
	CPCLS	**1.03**	**0.94**	0.00	0.40	**0.80**	**1.40**	**9.26**	**9.26**	**1.00**	**0.0423**	**1.16**
Average	AEA	1.62	1.12	0.00	0.78	1.44	2.26	8.40	8.40	1.48	0.0512	---
	BPNN	1.03	0.88	0.00	0.38	0.82	1.44	7.78	7.78	1.06	0.0341	104.19
	CPCLS	**0.94**	**0.79**	0.00	**0.36**	**0.75**	**1.32**	**7.18**	**7.18**	**0.96**	**0.0311**	**1.53**

Table 5:
Performance accuracy comparison (Individual Models)

Sector	AEA/BPNN (%)	AEA/CPCLS (%)	BPNN/CPCLS (%)
SR_H^{USI}	134.71	155.86	9.01
SR_H^D	97.83	110.00	6.15
SR_{US1}^H	4.42	24.21	18.95
SR_{US2}^H	16.47	41.43	21.43
SR_{UK}^H	51.11	70.00	12.50
SR_H^{US2}	62.82	67.11	2.63
SR_H^{UK}	20.43	27.27	5.68
SR_H^A	40.74	47.57	4.85
Average	53.57	67.93	10.15

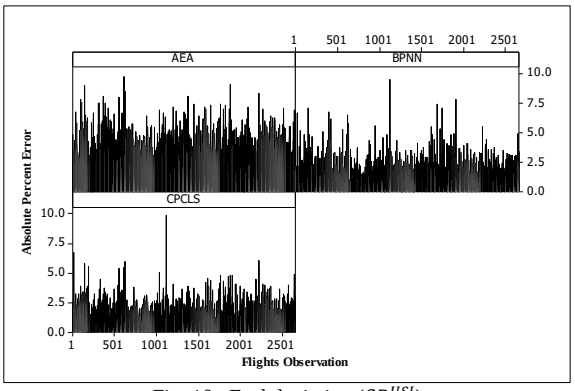

Fig. 10. Fuel deviation (SR_H^{USI})

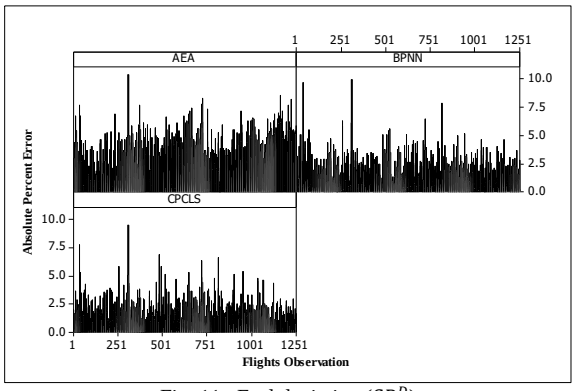

Fig. 11. Fuel deviation (SR_H^D)

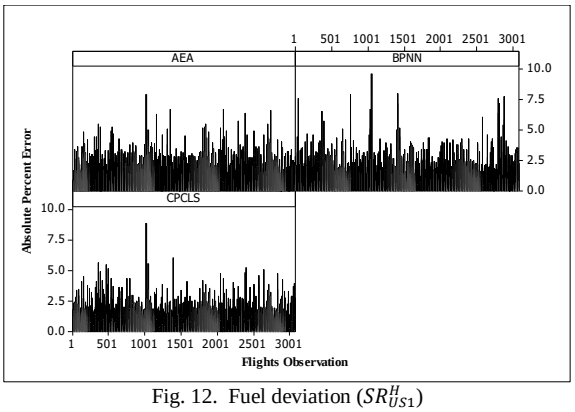

Fig. 12. Fuel deviation (SR_{US1}^H)

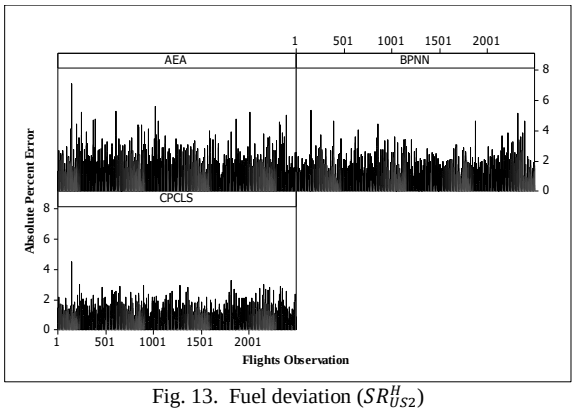

Fig. 13. Fuel deviation (SR_{US2}^H)

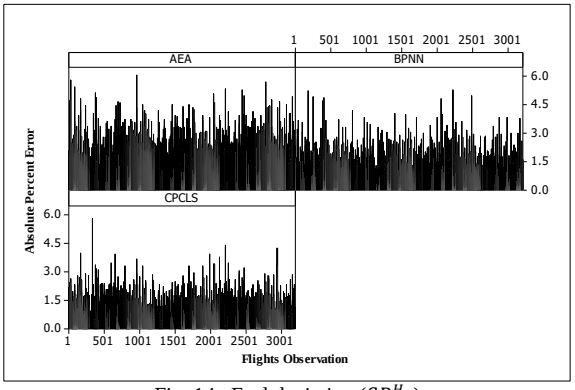

Fig. 14. Fuel deviation (SR_{UK}^H)

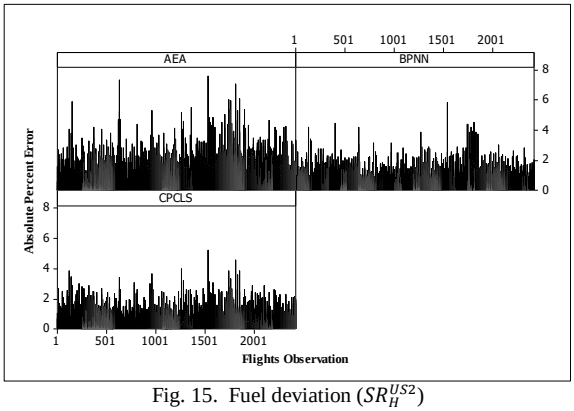

Fig. 15. Fuel deviation (SR_H^{US2})

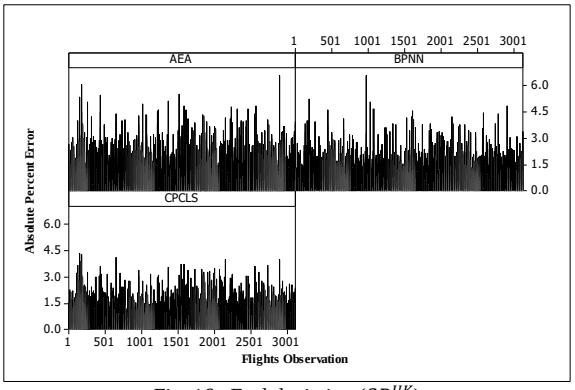

Fig. 16. Fuel deviation (SR_H^{UK})

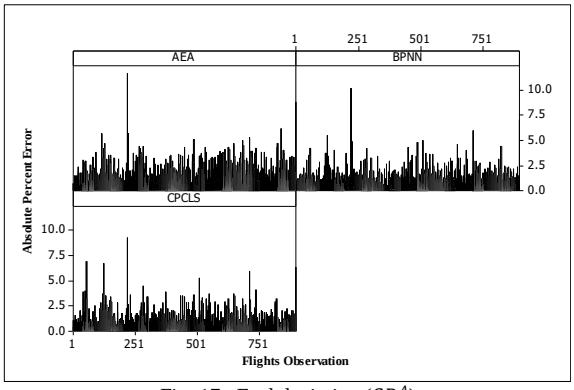

Fig. 17. Fuel deviation (SR_H^A)

APE for flights is much higher for SR_H^{USI} (MAPE 2.84%) compared to SR_{US1}^H(MAPE 1.18%). The reason is using a same global parameter for both sectors rather than local parameters. Based on historical data, the seventh finding is that the CPCLS minimized this issue and shows APE for both sectors (Figs. 10 and 12) almost identical. For other sectors having same route, for instance SR_H^{US2},SR_{US2}^H and SR_{UK}^H, SR_H^{UK}, the figures show the same issue of usage of global hyperparameter and old database by the AEA. The SR_H^{US2} shows higher APE compared to SR_{US2}^H, and SR_{UK}^H shows higher APE compared to SR_H^{UK}. This demonstrate the main limitation, as highlighted in literature review section 2, of the EA based methods.

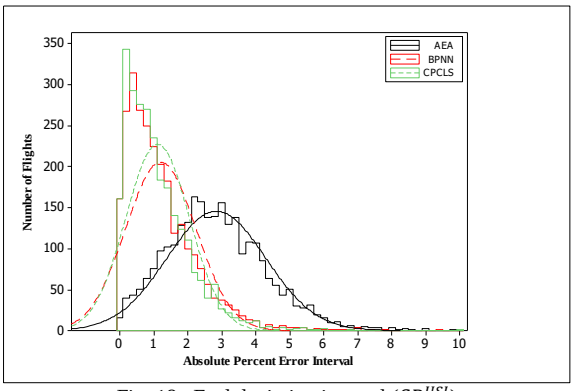

Fig. 18. Fuel deviation interval (SR_H^{USI})

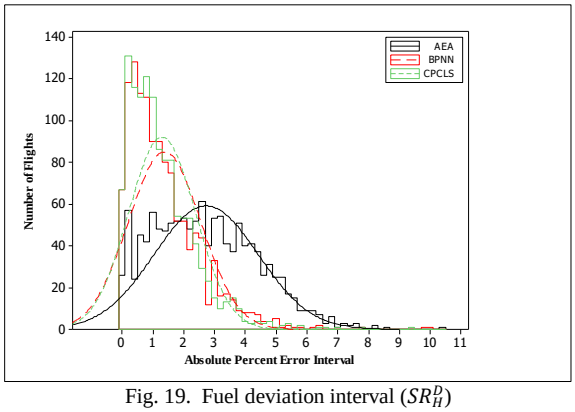

Fig. 19. Fuel deviation interval (SR_H^{D})

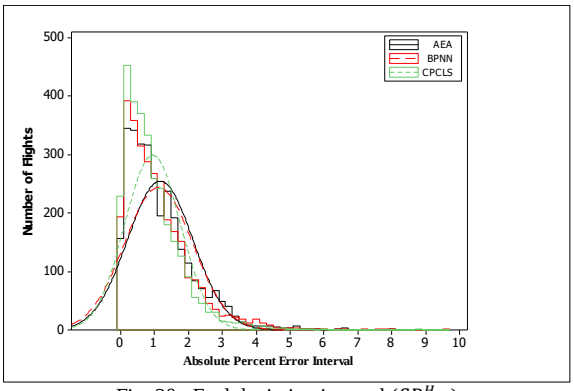

Fig. 20. Fuel deviation interval (SR_{US1}^{H})

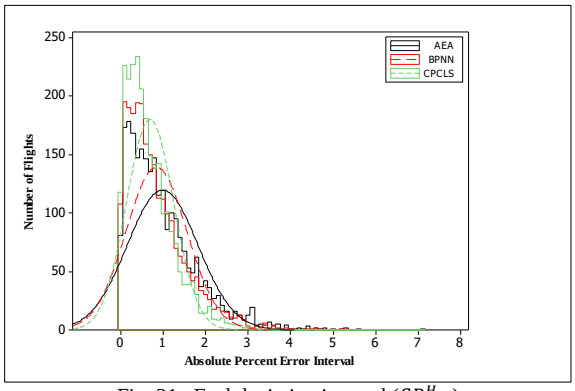

Fig. 21. Fuel deviation interval (SR_{US2}^{H})

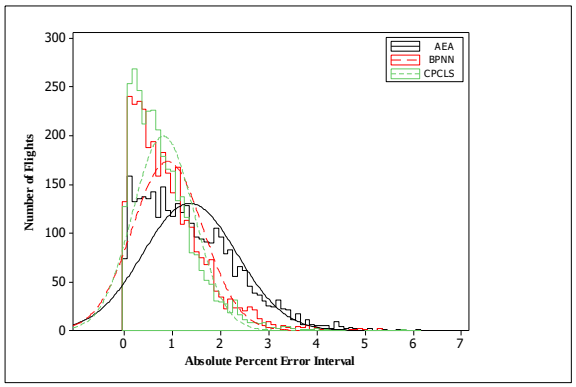

Fig. 22. Fuel deviation interval (SR_{UK}^{H})

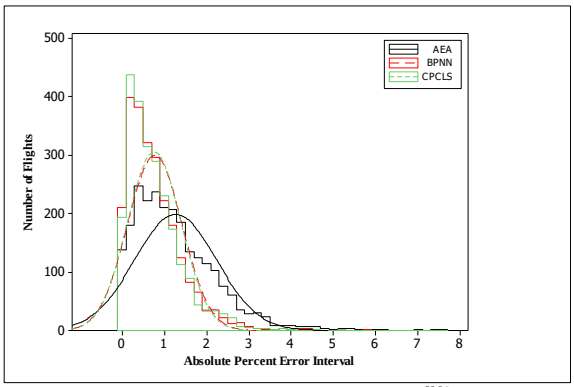

Fig. 23. Fuel deviation interval (SR_H^{US2})

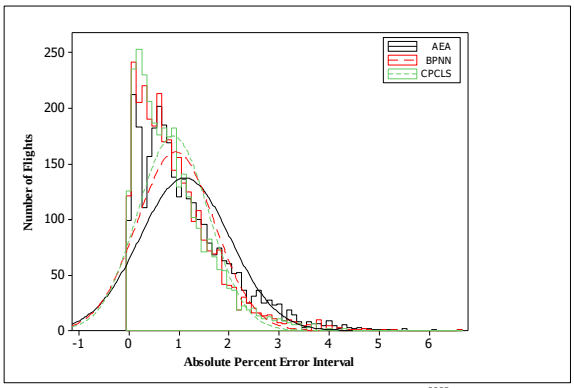

Fig. 24. Fuel deviation interval (SR_H^{UK})

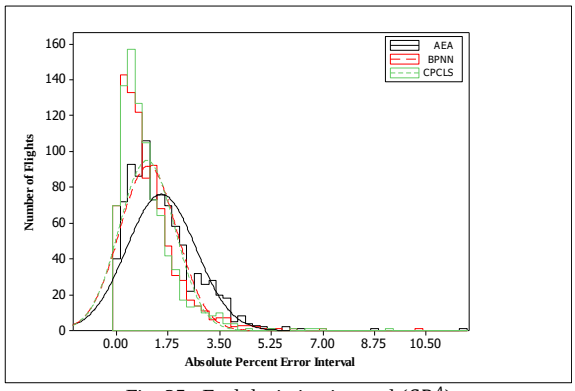

Fig. 25. Fuel deviation interval (SR_H^{A})

Figs. 18- 25 shows the histogram and distribution comparison of the AEA, BPNN, and CPCLS. The histograms indicate that the CPCLS fuel deviation is spread on less APE bins with a major focus on estimating flights contributing to less deviation bins. For most sectors, the eighth finding is that the distribution of CPCLS is more clustered as compared to the AEA and BPNN. For instance, in Figs. 18 and 19, the bell shape of SR_H^{USI} and SR_H^{D} sectors show that the AEA method fuel deviation exists mainly in the bins 1.50-3.50%, whereas, the CPCLS distribution peak shape shows that fuel deviation exists in 0.00-1.00%. In Figs. 20, 23, 25 for SR_{US1}^{H}, SR_H^{US2}, SR_H^{A}, the distribution of BPNN is almost identical to the AEA or CPCLS. For all sectors, the histogram and distribution trend of the CPCLS illustrates that it

attempts to estimate trip fuel for flights with less deviation compared to the AEA and BPNN. The histogram and distribution of the AEA show much dispersed fuel deviation. The possible reason for the AEA high fuel deviation is less confidence in method because of the availability and applicability of global parameters rather than the local parameters.

The experimental work demonstrates the effectiveness of the CPCLS. The estimation of the CPCLS results in a fuel deviation mean absolute error percentage (MAPE) of 1.05% and 0.94% with an improvement of 45.71% and 67.93%, and BPNN estimated fuel with deviation of 1.32% and 1.03% with an improvement of 15.91% and 53.57% compared to the AEA fuel deviation of 1.53% and 1.62% for combined and individual models respectively. In comparison with the BPNN, the CPCLS achieved an improvement of 25.71% and 10.15% for combined and individual models respectively. Similarly, CPCLS trained the network a hundred times faster compared to the traditional BPNN. In the CPCLS, the analytical calculation of connection weights by eigendecomposition of the input covariance squares matrix tries to generate linearly independent hidden units explaining the maximum variance in the operational parameters. The eigendecomposition eliminates a problem of linear dependence of operational parameters and the hidden units. These characteristics make it more suitable to accurately estimate trip fuel with less expertise requirement and faster learning speed. Sector-wise (individual) fuel estimation models validates the earlier work of fuel estimation by the BPNN with a limited amount of data operational parameters and flights, divided among different flight phases. However, during the study, it is observed that the estimation accuracy of the BPNN starts decreasing with an increase in the data structure. The varying nature of the operational parameter behavior made it difficult for the BPNN to create hidden units with the capability of operational parameters linearly separable. The sole purpose of this work is not a comparison of the CPCLS with AEA and BPNN. The main contribution is to propose and apply the self-organizing CNN algorithm that gives better estimation accuracy with less expertise requirement and faster learning speed. The significant improvement of 67.93% by the proposed CPCLS method in existing airline fuel estimation (i.e., AEA) will directly benefit in eliminating excess fuel consumption.

6. Conclusions

This paper addresses airline trip fuel deviation from estimated values, what results in excessive fuel consumption. A low confidence in existing estimation methods leads to extra fuel loaded in the aircraft. This ultimately increases the weight of the aircraft, which requires more thrust (fuel) to get balanced. Estimating the trip fuel accurately and faster with less expertise requirement along with high-dimensional data has become an important research topic. Early attempts made use of EAs to estimate the trip fuel for each flight. Their complex mathematical computation, high consultation and testing cost, their reliance on global parameters rather than local parameters, and high estimation errors limit the applicability of EAs. As an alternative, the application of BPNNs was suggested. Traditional BPNNs depend on trial-and-error approaches to find the best hyperparameters. This may require a lot of effort and user expertise to search for the best combinations of hyperparameters. Additionally, the iterative tuning of connection weights by backpropagation learning algorithms may make learning complicated, inefficient, and time-consuming. This may in turn cause BPNNs to converge to a suboptimal solution. The varying nature of the operational parameters and high dimensionality of data might make it more difficult for BPNNs to accurately estimate trip fuel because of the abovementioned problems. Therefore, in this study, a novel CNN named as CPCLS, is proposed with the objective of achieving better trip fuel estimation with less expertise requirement and faster learning speed. A comparative study with the existing AEA and a traditional BPNN to estimate trip fuel proves that our work is unique in studying the sectors both in combination and individually by considering all flight phases as a complete flight trip. Unlike AEA and BPNN, the proposed CPCLS does not require complex mathematical formulation, a trial-and-error approach, gradient derivative techniques, or too many hyperparameter initialisations. This makes it superior to estimate trip fuel accurately with less expertise requirements and faster learning speed.

The experimental work on the varying nature of operational parameters and high dimensionality of data demonstrates the effectiveness of CPCLS, which, with low fuel deviation and more stable results, demonstrates that can be effectively employed for estimating trip fuel for each flight. The better and faster estimation results of the self-organising CPCLS is because of its analytical calculation of network connection weights explaining maximum variance in the operational parameters. Thus, CPCLS is a promising machine learning tool for estimating flight trip fuel. In practical applications, the proposed CPCLS will be beneficial to airlines by accurately planning the amount of flight trip fuel. This may avoid the need for extra fuel to be loaded in the aircraft and helps in better management control. Flight planners and pilots will be more confident with trip fuel estimation and will avoid the requirement for extra fuel in the discrepancy tank. The weight of aircrafts will decrease, what will not only reduce the fuel consumption but also will increase the lifetime of the engines. The flight planner may suggest and select a suitable airways route by simulating different combinations of operational parameters to reach the destination with less fuel requirements. Furthermore, controlling excess fuel consumption may benefit in contributing to less environmental pollution, preventing the scarce jet fuel from natural resources from being wasted, as well as surviving with growing fuel prices, less unplanned aircraft maintenance, fewer fuel surcharges, and more competitiveness by fulfilling the demand of passengers and cargo.

Appendix A. Simmod and BADA fuel estimation models

The energy balance can be expressed as energy gain and loss by the aircraft as it travels along the path profile

$$energy\ gain - energy\ loss = energy\ change \tag{A.1}$$

$$E_T - E_D = \triangle KE + \triangle PE \tag{A.2}$$

During each flight operation, the change in kinetic $\triangle KE$ and potential $\triangle PE$ energies, the aircraft suffers energy losses because of drag E_D is adjusted by the amount of thrust E_T to maintain the energy balance. Thus, aircraft can be considered a system with energy losses and gains that should be continuously balanced by consumption of fuel energy. Collins (1982) derived an algorithm for fuel estimation based

on the aircraft configuration, weight and path profile. The path profile may be described by considering a change in true airspeed, altitude and time, with the following expressions for fuel estimation

$$\widehat{\boldsymbol{Y}} = \frac{t\bar{v}_t F_N}{v_t P} + t(LAM1) \tag{A.3}$$

Such that

$$F_N = \frac{R_1 K_1}{2}\bar{\rho} S_w \overline{v_t^2} + \frac{2R_2 K_2 m^2}{\bar{\rho} S_w \overline{v_t^2}} + \frac{m}{gt}(v_{t2} - v_{t1}) + \frac{m}{tv_t}(h_2 - h_1) \tag{A.4}$$

$$P = K_{10} e^{K_{11}\bar{v}_t} + K_{12}\left(\frac{h_1^2 + h_1 h_2 + h_2^2}{3}\right) + K_{13}\left(\frac{h_1 + h_2}{3}\right) + K_{14} \tag{A.5}$$

$$LAM1 = \begin{cases} 0 & if\ F_N \le K_7 \\ K_8 F_N + K_9 & otherwise \end{cases} \tag{A.6}$$

where $\widehat{\boldsymbol{Y}}$ is fuel estimation, t is time, v is true velocity, $\bar{v}$ is average velocity, F_N is thrust, P is thrust energy, $LAM1$ is a relationship between fuel flow and thrust, $\bar{\rho}$ is average atmospheric density, S_w is wing area, m is weight, g is gravitational acceleration and h is altitude. In the Equations (A.4), (A.5) and (A.6), Collins (1982) explained that K_1, K_2, R_1, R_2, K_7, K_8, K_9, K_{10}, K_{11}, K_{12}, K_{13}, K_{14} and S_w are aircraft specific constant and need to be determined for each aircraft. Constants K_1, K_2 and S_w determine the relationship between drag and lift/weight coefficient and constants K_{10}, K_{11}, K_{12}, K_{13} and K_{14} determine the relationship between fuel consumption and energy gain from the thrust as a function of velocity and altitude. The drag increases with the change in a configuration such as gear up or gear down and can be expressed as

$$R_1 = GU_1 F^3 + GU_2 F^2 + GU_3 F + 1\ (gear\ up) \tag{A.7}$$

$$R_1 = GD_1 F^3 + GD_2 F^2 + GD_3 F + 1\ (gear\ down) \tag{A.8}$$

$$R_2 = FDM_1 F^3 + FDM_2 F^2 + FDM_3 F + 1 \tag{A.9}$$

where F is flap angle, D is drag and M is Mach number. Similarly, The Base of Aircraft Data (BADA) developed by Eurocontrol (Nuic, 2014) is another engineering-based fuel estimation method that estimates thrust specific fuel consumption (TSFC) η as a function of true airspeed v_{TAS}. The fuel for different phases and engines can be expressed as

$$\eta = \begin{cases} C_{f1}\left(1 + \frac{v_{TAS}}{C_{f2}}\right), & Jet\ Engine \\ C_{f1}\left(1 + \frac{v_{TAS}}{C_{f2}}\right)\left(\frac{v_{TAS}}{1000}\right), & Turboprop\ Engine \end{cases} \tag{A.10}$$

$$\widehat{\boldsymbol{Y}}_{nom} = \eta Thr, \qquad Jet\ and\ Turbo\ Engine \tag{A.11}$$

$$\widehat{\boldsymbol{Y}}_{nom} = C_{f1}, \qquad Piston\ Engine \tag{A.12}$$

where $\widehat{\boldsymbol{Y}}_{nom}$ is nominal fuel estimation and C_{f1}, C_{f2} are fuel flow coefficients for the specific aircraft type and are derived in operational performance files (OPF) of BADA. Equations (A.10), (A.11) and (A.12) are used for fuel estimation in all flight phases except during idle descent and cruise. The following mathematical calculations need to be performed to estimate the fuel consumption for idle descent phase of flight

$$\widehat{\boldsymbol{Y}}_{min} = C_{f3}\left(1 + \frac{H_p}{C_{f4}}\right), \qquad Jet\ and\ Turbo\ Engine \tag{A.13}$$

$$\widehat{\boldsymbol{Y}}_{min} = C_{f3}, \qquad Piston\ Engine \tag{A.14}$$

where $\widehat{\boldsymbol{Y}}_{min}$ is minimal fuel estimation and C_{f3}, C_{f4} are fuel flow coefficient. When an aircraft switches from idle descent and reaches the approach ap and landing ld phase, the thrust is increased. The calculation for fuel flow at approach and landing phase can be expressed as

$$\widehat{\boldsymbol{Y}}_{ap/ld} = MAX(\widehat{\boldsymbol{Y}}_{nom}, \widehat{\boldsymbol{Y}}_{min}), \qquad Jet\ and\ Turbo\ Engine \tag{A.15}$$

Cruise phase fuel is estimated by using η, cruise thrust Thr and the cruise fuel flow factor C_{fcr} for jet and turbo engine, and fuel coefficients C_{f1} and C_{f3} for piston engine

$$\widehat{\boldsymbol{Y}}_{cr} = \eta \text{ x } Thr \text{ x } C_{fcr}, \qquad Jet\ and\ Turbo\ Engine \tag{A.16}$$

$$\widehat{\boldsymbol{Y}}_{cr} = C_{f1} \text{x } C_{f3}, \qquad Piston\ Engine \tag{A.17}$$

The BADA manual contains detailed mathematical expressions needed to calculate the thrust Thr for each flight phase of the climb, take-off, cruise, descent, approach and landing.

Acknowledgment

This work was supported by a grant from the Research Committee of The Hong Kong Polytechnic University under the account code RLKA and the National Natural Science Foundation of China with Grant Nos. 71871064 and 18BGL003.

References

Abdelghany, K., Abdelghany, A., Raina, S., 2005. A model for the airlines' fuel management strategies. J. Air Transp. Manag. 11, 199–206. https://doi.org/10.1016/j.jairtraman.2004.10.002

Au, K.F., Choi, T.M., Yu, Y., 2008. Fashion retail forecasting by evolutionary neural networks. Int. J. Prod. Econ. 114, 615–630. https://doi.org/10.1016/j.ijpe.2007.06.013

Baklacioglu, T., 2016. Modeling the fuel flow-rate of transport aircraft during flight phases using genetic algorithm-optimized neural networks. Aerosp. Sci. Technol. 49, 52–62. https://doi.org/10.1016/j.ast.2015.11.031

Banerjee, P., Singh, V.S., Chatttopadhyay, K., Chandra, P.C., Singh, B., 2011. Artificial neural network model as a potential alternative for groundwater salinity forecasting. J. Hydrol. 398, 212–220.

Bianchini, M., Scarselli, F., 2014. On the complexity of neural network classifiers: A comparison between shallow and deep architectures. IEEE Trans. Neural Networks Learn. Syst. 25, 1553–1565. https://doi.org/10.1109/TNNLS.2013.2293637

Chati, Y.S., Balakrishnan, H., 2017. Statistical modeling of aircraft engine fuel, in: Twelfth USA/Europe Air Traffic Management Research and Development Seminar (ATM 2017).

Choi, T.M., Chiu, C.H., Chan, H.K., 2016. Risk management of logistics systems. Transp. Res. Part E Logist. Transp. Rev. 90, 1–6. https://doi.org/10.1016/j.tre.2016.03.007

Choi, T.M., Wallace, S.W., Wang, Y., 2018. Big data analytics in operations management. Prod. Oper. Manag. 27, 1868–1883. https://doi.org/10.1111/poms.12838

Chung, S.H., Ma, H.L., Chan, H.K., 2017. Cascading delay risk of airline workforce deployments with crew pairing and schedule optimization. Risk Anal. 37, 1443–1458. https://doi.org/10.1111/risa.12746

Chung, S.H., Tse, Y.K., Choi, T.M., 2015. Managing disruption risk in express logistics via proactive planning. Ind. Manag. Data Syst. 115, 1481–1509. https://doi.org/10.1108/IMDS-04-2015-0155

Collins, B.P., 1982. Estimation of aircraft fuel consumption. J. Aircr. 19, 969–975. https://doi.org/10.2514/3.44799

Cui, Q., Li, Y., 2017. Airline efficiency measures under CNG2020 strategy: An application of a Dynamic By-production model. Transp. Res. Part A Policy Pract. 106, 130–143.

Dancila, B.D., Botez, R., Labour, D., 2013. Fuel burn prediction algorithm for cruise, constant speed and level flight segments. Aeronaut. J. 117, 491–504. https://doi.org/10.1017/S0001924000008149

Diao, X., Chen, C.-H., 2018. A sequence model for air traffic flow management rerouting problem. Transp. Res. Part E Logist. Transp. Rev. 110, 15–30. https://doi.org/10.1016/j.tre.2017.12.002

Edwards, H.A., Dixon-Hardy, D., Wadud, Z., 2016. Aircraft cost index and the future of carbon emissions from air travel. Appl. Energy 164, 553–562. https://doi.org/10.1016/j.apenergy.2015.11.058

Fahlman, S.E., 1988. An empirical study of learning speed in back-propagation networks. School of Computer Science, Carnegie Mellon University, Pittsburgh PA 15213.

Fahlman, S.E., Lebiere, C., 1990. The cascade-correlation learning architecture, in: Advances in Neural Information Processing Systems. pp. 524–532.

Ferrari, S., Stengel, R.F., 2005. Smooth function approximation using neural networks. IEEE Trans. Neural Networks 16, 24–38. https://doi.org/10.1109/TNN.2004.836233

Guo, X., Grushka-Cockayne, Y., De Reyck, B., 2018. Forecasting Airport Transfer Passenger Flow Using Real-Time Data and Machine Learning. SSRN Electron. J. https://doi.org/10.2139/ssrn.3245609

Hecht-Nielsen, 1989. Theory of the backpropagation neural network, in: International Joint Conference on Neural Networks. IEEE, pp. 593–605 vol.1. https://doi.org/10.1109/IJCNN.1989.118638

Hornik, K., Stinchcombe, M., White, H., 1989. Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366. https://doi.org/https://doi.org/10.1016/0893-6080(89)90020-8

Huang, C., Xu, Y., Johnson, M.E., 2017. Statistical modeling of the fuel flow rate of GA piston engine aircraft using flight operational data. Transp. Res. Part D Transp. Environ. 53, 50–62. https://doi.org/10.1016/j.trd.2017.03.023

Huang, G.-B., Chen, L., Siew, C.K., 2006a. Universal approximation using incremental constructive feedforward networks with random

762 hidden nodes. IEEE Trans. Neural Networks 17, 879–892.

763 Huang, G.-B., Zhu, Q.-Y., Siew, C.-K., 2006b. Extreme learning machine: theory and applications. Neurocomputing 70, 489–501.

764 Huang, G., Song, S., Wu, C., 2012. Orthogonal least squares algorithm for training cascade neural networks. IEEE Trans. Circuits Syst. I
765 Regul. Pap. 59, 2629–2637.

766 Hunter, D., Yu, H., Pukish III, M.S., Kolbusz, J., Wilamowski, B.M., 2012. Selection of proper neural network sizes and architectures—
767 A comparative study. IEEE Trans. Ind. Informatics 8, 228–240.

768 Hwang, J.-N., You, S.-S., Lay, S.-R., Jou, I.-C., 1996. The cascade-correlation learning: A projection pursuit learning perspective. IEEE
769 Trans. Neural Networks 7, 278–289.

770 IATA, 2019. Industry facts and statistics [WWW Document]. Online. URL
771 https://www.iata.org/pressroom/facts_figures/fact_sheets/Pages/index.aspx (accessed 8.5.19).

772 IATA, 2018. Fact Sheet Climate Change & CORSIA [WWW Document]. Online. URL
773 https://www.iata.org/pressroom/facts_figures/fact_sheets/Documents/fact-sheet-climate-change.pdf (accessed 5.7.18).

774 Irrgang, M.E., Kaul, C.E., Hall, A.E., Klerk, A.D., Elham, Boozarjomehri, 2015. Aircraft fuel optimization analytics. US 2015/0279218
775 A1.

776 Jensen, L., Hansman, R.J., Venuti, J.C., Reynolds, T., 2013. Commercial airline speed optimization strategies for reduced cruise fuel
777 consumption, in: 2013 Aviation Technology, Integration, and Operations Conference. pp. 1026–1038.

778 Kapanova, K.G., Dimov, I., Sellier, J.M., 2018. A genetic approach to automatic neural network architecture optimization. Neural
779 Comput. Appl. 29, 1481–1492.

780 Khanmohammadi, S., Tutun, S., Kucuk, Y., 2016. A new multilevel input layer artificial neural network for predicting flight delays at
781 JFK airport. Procedia Comput. Sci. 95, 237–244.

782 Kovalishyn, V. V, Tetko, I. V, Luik, A.I., Kholodovych, V. V, Villa, A.E.P., Livingstone, D.J., 1998. Neural network studies. 3. variable
783 selection in the cascade-correlation learning architecture. J. Chem. Inf. Comput. Sci. 38, 651–659.

784 Krogh, A., Hertz, J.A., 1992. A simple weight decay can improve generalization, in: Advances in Neural Information Processing
785 Systems. pp. 950–957. https://doi.org/https://dl.acm.org/citation.cfm?id=2987033

786 Kumar, A., Rao, V.R., Soni, H., 1995. An empirical comparison of neural network and logistic regression models. Mark. Lett. 6, 251–
787 263.

788 LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

789 Lehtokangas, M., 2000. Modified cascade-correlation learning for classification. IEEE Trans. Neural Networks 11, 795–798.

790 Liang, H., Dai, G., 1998. Improvement of cascade correlation learning. Inf. Sci. (Ny). 112, 1–6.

791 Liew, S.S., Khalil-Hani, M., Bakhteri, R., 2016. An optimized second order stochastic learning algorithm for neural network training.
792 Neurocomputing 186, 74–89.

793 Lin, Z., Vlachos, I., 2018. An advanced analytical framework for improving customer satisfaction: A case of air passengers. Transp. Res.
794 Part E Logist. Transp. Rev. 114, 185–195. https://doi.org/10.1016/j.tre.2018.04.003

795 Merkert, R., Swidan, H., 2019. Flying with(out) a safety net: Financial hedging in the airline industry. Transp. Res. Part E Logist.
796 Transp. Rev. 127, 206–219. https://doi.org/10.1016/j.tre.2019.05.012

797 Nuic, A., 2014. User manual for the Base of Aircraft Data (BADA) revision 3.12, European Organisation for the Safety of Air
798 Navigation.

799 Pagoni, I., Psaraki-Kalouptsidi, V., 2017. Calculation of aircraft fuel consumption and CO2 emissions based on path profile estimation
800 by clustering and registration. Transp. Res. Part D Transp. Environ. 54, 172–190. https://doi.org/10.1016/j.trd.2017.05.006

801 Qiao, J., Li, F., Han, H., Li, W., 2016. Constructive algorithm for fully connected cascade feedforward neural networks. Neurocomputing
802 182, 154–164. https://doi.org/10.1016/j.neucom.2015.12.003

803 Ruiz-Aguilar, J.J., Turias, I.J., Jiménez-Come, M.J., 2014. Hybrid approaches based on SARIMA and artificial neural networks for
804 inspection time series forecasting. Transp. Res. Part E Logist. Transp. Rev. 67, 1–13. https://doi.org/10.1016/j.tre.2014.03.009

805 Ryerson, M.S., Hansen, M., Bonn, J., 2011. Fuel consumption and operational performance, in: 9th USA/Europe Air Traffic
806 Management Research and Development Seminar. pp. 1–10.

807 Schilling, G.D., 1997. Modeling aircraft fuel consumption with a neural network. Dr. Diss. Virginia Tech.

808 Senzig, D.A., Fleming, G.G., Iovinelli, R.J., 2009. Modeling of terminal-area airplane fuel consumption. J. Aircr. 46, 1089–1093.

https://doi.org/10.2514/1.42025

Seufert, J.H., Arjomandi, A., Dakpo, K.H., 2017. Evaluating airline operational performance: A Luenberger-Hicks-Moorsteen productivity indicator. Transp. Res. Part E Logist. Transp. Rev. 104, 52–68. https://doi.org/10.1016/j.tre.2017.05.006

Sheng, D., Li, Z.-C., Fu, X., 2019. Modeling the effects of airline slot hoarding behavior under the grandfather rights with use-it-or-lose-it rule. Transp. Res. Part E Logist. Transp. Rev. 122, 48–61. https://doi.org/10.1016/j.tre.2018.11.006

Sibdari, S., Mohammadian, I., Pyke, D.F., 2018. On the impact of jet fuel cost on airlines' capacity choice: Evidence from the U.S. domestic markets. Transp. Res. Part E Logist. Transp. Rev. 111, 1–17. https://doi.org/10.1016/j.tre.2017.12.009

Singh, V., Sharma, S.K., 2015. Fuel consumption optimization in air transport: a review, classification, critique, simple meta-analysis, and future research implications. Eur. Transp. Res. Rev. 7, 12. https://doi.org/10.1007/s12544-015-0160-x

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958. https://doi.org/10.1214/12-AOS1000

Trani, A., Wing-Ho, F., Schilling, G., Baik, H., Seshadri, A., 2004. A neural network model to estimate aircraft fuel consumption, in: AIAA 4th Aviation Technology, Integration and Operations (ATIO) Forum. American Institute of Aeronautics and Astronautics, Reston, Virigina, p. 6401. https://doi.org/10.2514/6.2004-6401

Trani, A.A., Wing-Ho, F., 1997. Enhancements to SIMMOD : A neural network post-processor to estimate aircraft fuel consumption Phase I final report.

Turgut, E.T., Cavcar, M., Usanmaz, O., Canarslanlar, A.O., Dogeroglu, T., Armutlu, K., Yay, O.D., 2014. Fuel flow analysis for the cruise phase of commercial aircraft on domestic routes. Aerosp. Sci. Technol. 37, 1–9. https://doi.org/10.1016/j.ast.2014.04.012

Turgut, E.T., Rosen, M.A., 2012. Relationship between fuel consumption and altitude for commercial aircraft during descent: Preliminary assessment with a genetic algorithm. Aerosp. Sci. Technol. 17, 65–73. https://doi.org/10.1016/j.ast.2011.03.007

Wilamowski, B.M., Cotton, N.J., Kaynak, O., Dundar, G., 2008. Computing gradient vector and Jacobian matrix in arbitrarily connected neural networks. IEEE Trans. Ind. Electron. 55, 3784–3790. https://doi.org/10.1109/TIE.2008.2003319

Yanto, J., Liem, R.P., 2018. Aircraft fuel burn performance study: A data-enhanced modeling approach. Transp. Res. Part D Transp. Environ. 65, 574–595. https://doi.org/10.1016/j.trd.2018.09.014