JIANPING ZHANG, Ph.D.[1]
(Corresponding author)
E-mail: 834212837@qq.com
LIWEI DUAN, Ph.D. candidate[1,2]
E-mail: L.Duan@qq.com
JING GUO, Bachelor[3]
E-mail:jingg@vip.163.com
WEIDONG LIU, M.Sc.[1]
E-mail: atmlwd@163.com
XIAOJIA YANG, M.Sc.[1]
E-mail: atmyxj@163.com
RUIPING ZHANG, M.Sc.[4]
E-mail: snowzrp@163.com
[1] The Second Research Institute of Civil Aviation Administration of China, No. 17, South Section 2, 2nd Ring Road, Chengdu, Sichuan Province, 610041, China
[2] School of Transportation and Logistics, Southwest Jiaotong University, No. 111, North Section 1, 2nd Ring Road, Chengdu, Sichuan Province, 610031, China
[3] Civil Aviation Administration of China, No. 155, Dongsi West Avenue, Dongcheng District, Beijing, 100710, China
[4] Southwest Regional Air Traffic Management Bureau of Civil Aviation of China, Shuangliu International Airport, Chengdu, Sichuan Province, 610202, China

Traffic Engineering
Original Scientific Paper
Submitted: Nov. 19, 2015
Accepted: July 14, 2016

A GENETIC ALGORITHM-BASED BP NEURAL NETWORK METHOD FOR OPERATIONAL PERFORMANCE ASSESSMENT OF ATC SECTOR

ABSTRACT

To assess operational performance of air traffic control sector, a multivariate detection index system consisting of 5 variables and 17 indicators is presented, which includes operational trafficability, operational complexity, operational safety, operational efficiency, and air traffic controller workload. An improved comprehensive evaluation method, is designed for the assessment by optimizing initial weights and thresholds of back propagation (BP) neural network using genetic algorithm. By empirical study conducted in one air traffic control sector, 400 sets of sample data are selected and divided into 350 sets for network training and 50 sets for network testing, and the architecture of genetic algorithm-based back propagation (GABP) neural network is established as a three-layer network with 17 nodes in input layer, 5 nodes in hidden layers, and 1 node in output layer. Further testing with both GABP and traditional BP neural network reveals that GABP neural network performs better than BP neural work in terms of mean error, mean square error and error probability, indicating that GABP neural network can assess operational performance of air traffic control sector with high accuracy and stable generalization ability. The multivariate detection index system and GABP neural network method in this paper can provide comprehensive, accurate, reliable and practical operational performance assessment of air traffic control sector, which enable the frontline of air traffic service provider to detect and evaluate operational performance of air traffic control sector in real time, and trigger an alarm when necessary.

KEY WORDS

air traffic control sector; operational performance; multivariate detection index system; genetic algorithm; back propagation neural network; comprehensive evaluation;

1. INTRODUCTION

A sector is a volume of airspace for air traffic control (ATC), normally positioned with a team of air traffic controllers [1, 2]. Research on operational performance assessment of ATC sector is meaningful for all the main aspects of air traffic management (ATM), composed of air traffic service, air traffic flow management (ATFM), and airspace management. The effective results of operational performance assessment of ATC sector should be the basis and precondition for the adjustment of ATC operation and ATFM strategies, and also the optimization of ATC airspace structure.

As the primary task of ATM, it has taken a long time to study how to prevent aircraft collision and aircraft-obstacle collision. Before 1990s, the phenomena of traffic congestion and flight delay in ATC airspace were not serious since the air traffic flow was not relatively heavy, and the studies on operational performance assessment of ATC sector focused on aircraft separation and collision risk analysis. Reich proposed possibility distribution functions to estimate aircraft collision risk, and discussed aircraft separation minima to be applied in longitudinal, lateral and vertical dimensions in airspace, respectively. Related analytical techniques were applied in the North Atlantic region [3-5]. Follow-up studies introduce the improved collision risk models for different categories, such as confliction, air miss, and collision, based on Reich's model [6-9]. On the basis of Reich's model and the improved models, the concept of Target Level of Safety (TLS) in airspace is proposed by International Civil Aviation Organization (ICAO) [10]. Besides, the studies on aircraft lateral separation minima in the North Atlantic region [11], aircraft vertical separation minima in Europe [12-14], aircraft longitudinal separation minima in the Asia-Pacific region [15], as well as collision risk analysis for Performance-based Navigation (PBN) procedures [16-19], are all based on Reich's study. The mentioned aircraft separation and collision risk studies contribute greatly to aircraft separation minima determination in the aspects of structured airspace and evolved flight procedure. However, they are not the total efforts to prevent aircraft collision and aircraft-obstacle collision. In the aspect of daily operation, a number of automation systems have been applied to automate air traffic conflict detection and resolution since the early 1990s, taking Traffic Alert and Collision Avoidance System (TCAS) and Ground Proximity Warning System (GPWS) in cockpit and conflict predictors in ATC automation system for examples, considering that human failures or errors are unavoidable and highlighted in the meanwhile [20].

With continuous growth of air traffic flow and expansion of air routes since the 1990s, traffic congestion and flight delay have happened frequently and seriously with more complex air traffic distribution, while air traffic controller (ATCO) workload increased under the circumstance. Thus, studies about the operational performance assessment of ATC sector expanded to airspace capacity and the related fields, including airspace utilization, complexity of air traffic flow, and ATCO workload as well. Most studies proposed that ATCO workload threshold can be used to access ATC sector capacity, in terms of either the number of simultaneously present aircraft or the number of aircraft traversing the sector per unit of time [1, 21-24]. The literature in this field is mostly related to the analysis of the factors affecting the complexity of the controller's task and workload [22]. Moreover, the capacity of a sector presents variability and unpredictability, which are also explained by those factors, summarized as three categories, airspace physical structure, air traffic situation, and operational constraints [1]. Airspace utilization reveals the relationship between airspace capacity and traffic flow distribution, and it can be reflected in space-based or time-based utilization [25-26], traffic congestion status [27-28], and statistics of flight delays [29-31]. Furthermore, the complexity of air traffic flow is proposed for the recognition of air traffic situation in airspace, which is more accurate, comprehensive and dynamic than airspace capacity and its utilization. As the most representative assessment technique for the complexity of air traffic flow, dynamic density concept has been measured by detecting aircraft altitude, speed and heading changing times [32-37]. Even though ATCO workload is not directly related to operational performance of ATC sector, it is a kind of essential investment to implement ATC operation, and has great influence on the operational performance. ATCO workload is mainly measured by time consuming of visible task and converted time consuming of invisible task, to be evaluated in quantitative way [38-39]. Based on previous studies, the concept of operational performance of ATC sector could be defined as the operational quality and related level from different angles such as aircraft separation and collision risk, airspace capacity and its utilization, air traffic complexity and ATCO workload.

It is worth noting that recent studies have highlighted that the operational performance assessment of ATC sector has been involved in a multivariate problem, and affected by many factors which interfere with each other [40-42]. To improve operational performance of ATC sector, a multi-objective decision strategy should be considered. In [2], to ensure safety and efficiency of the aircraft, the desirable objectives of ATC are listed as follows: maximizing the runway throughput, minimizing the approach time of aircraft before landing, minimizing air traffic controllers' workload, maximizing fairness among the aircraft, minimizing the aircraft taxi-in/taxi-out time, minimizing the arrival/departure delay, minimizing deviations from an appropriate balance between arrivals and departures [2]. In [21], the analysis shows that congestion and delay reduction by significant capacity increase may be not applicable in a short term, while better traffic demand management and available capacity allocation could be more effective from both strategic and tactical perspectives [21]. Most of the existing studies, however, as far as we can see, are limited in only one aspect, or based on a single side. It is clear that one-factor-only studies cannot detect and evaluate operational performance of ATC sector systematically. On the other hand, most studies aim at theoretical, not empirical research. A novel method is needed to assess the operational performance of ATC sector at the frontline of air traffic service provider as empirical use.

Thus, to support multi-objective decision making for the improvement of operational performance of ATC sector for empirical and practical purpose, the contributions of this paper with respect to the related state-of-the-art could be listed as follows. An integrated and universal multivariate detection index system is established for operational performance of ATC sector, introducing 5 variables including operational trafficability (traffic throughput and containing ability), operational complexity, operational safety, operational efficiency, and ATCO workload, and 17 indicators calculated with real-time operation data which can be acquired with high engineering feasibility. A novel method to comprehensively evaluate operational performance of ATC sector in more reliable and accurate way is provided with Back Propagation (BP) neural network, in which the initial weights and thresholds are optimized by Genetic Algorithm (GA). Both the index system and the method are validated to have the availability to be integrated into application system for the frontline of air traffic service provider, to assist the operation personnel in the judgment and decision making about operation strategies from both strategic and tactical perspectives.

The remainder of this paper is as follows: Section 2 presents an integrated and universal multivariate detection index system, in which both the indicators under each variable and the definition for each indicator are listed. A comprehensive evaluation method based on GABP neural network is provided in Section 3, in which detailed method progress is discussed. Next, Section 4 presents an empirical study which is conducted in one ATC sector in China, and results analysis and discussion are carried out. Last but not least, Section 5 contains the conclusions and future research expectations.

2. MULTIVARIATE DETECTION INDEX SYSTEM

The operational performance of ATC sector is affected by various factors which have been investigated by the existing studies from single sides, i.e. aircraft collision prevention, airspace capacity and its utilization, complexity of air traffic flow, and ATCO workload. Each study focus has connected some variables and the related indicators. Since there is a number of indicators related to operational performance assessment of ATC sector showing up in previous works, the question is how to select and validate the indicators in the paper. In [35] a systematic method is presented, based on data mining to figure it out [35]. On the one hand, it is considered to extract the meaningful principal components as the selected indicators from a large scope of metrics based on linear correlation finding through Principal Component Analysis (PCA). On the other hand, it could be to select the most relevant indicators investigating the link between the indicator set and ATC sector configurations using neural networks, assuming there are non-linear interactions. The method mentioned is suitable for theoretical or experimental research. But for empirical and practical purpose in the paper, the selected indicators are supposed to be easily measured based on the feasibility of real-time operation data acquisition, and validated with numerous on-site investigations from senior ATCOs besides theoretical or experimental references. Based on that, this paper puts forward an integrated and universal multivariate detection index system systematically for operational performance of ATC sector, including 5 variables and 17 indicators.

In detail, the studies on aircraft collision prevention could trigger a variable named operational safety, under which the indicators of Short-Term Conflict Alert (STCA) and Minimum Safe Altitude Warning (MSAW) can be detected and recorded from conflict predictors in ATC automation system in the operation site of air traffic service provider. The studies on airspace capacity and its utilization could involve two variables named *operational trafficability* and *operational efficiency*. The variable *operational trafficability* could include the indicator of *air traffic flow* reflecting airspace capacity, and the indicators of *flight miles* for space-based airspace utilization and *flight time* for time-based airspace utilization, while the variable *operational efficiency* could include the indicators of *sector saturation* and *sector queue length* for traffic congestion status, and the indicators of *delay percentage*, *delay time* and *average delay time* as the statistics of flight delays. The studies on complexity of air traffic flow could connect a new variable named operational complexity with the indicators of *aircraft climbing frequency*, *aircraft descending frequency, aircraft speed changing frequency*, and *aircraft heading changing frequency*, and the variable *operational trafficability* again with the indicator of *air traffic flow density* as well. Meanwhile, the studies on ATCO workload as a variable could be indicated by *occupation percentage of air-ground communication radio* and *call times of air-ground communication radio*. For each variable as mentioned, the indicator definitions and notations are listed in *Table 1*.

The indicators and related measurement methods are compared with other options as a form of selection and validation based on the following principles through expert investigation method by an expert team including related scientific researchers, senior ATCOs, and experienced engineers.

Principle 1: Engineering feasibility

The feasibility of real-time operation data acquisition is positioned to lay the base of indicator selection. All those 17 indicators in total within the multivariate detection index system are measured based on hourly statistical real-time operation data, collected from ATC automation system, automatic message switching system, and very high frequency (VHF) voice

Table 1 – Multivariate detection index system for operational performance of ATC sector assessment

Variable	Indicator	Definition	Notation
Operational Trafficability	Air Traffic Flow	Total flight number in sector per hour	X_1 [flight]
	Flight Miles	Total flight miles in sector per hour	X_2 [km]
	Flight Time	Total flight time in sector per hour	X_3 [min]
	Air Traffic Flow Density	Ratio of total flight number and sector area per hour	X_4 [flight/km^2]
Operational Complexity	Aircraft Climbing Frequency	Total aircraft climbing times in sector per hour	X_5 [time]
	Aircraft Descending Frequency	Total aircraft descending times in sector per hour	X_6 [time]
	Aircraft Speed Changing Frequency	Total aircraft speed changing times in sector per hour	X_7 [time]
	Aircraft Heading Changing Frequency	Total aircraft heading changing times in sector per hour	X_8 [time]
Operational Safety	STCA Frequency	Total STCA times in sector per hour	X_9 [time]
	MSAW Frequency	Total MSAW times in sector per hour	X_{10} [time]
Operational Efficiency	Sector Saturation	Ratio of total flight number and sector capacity per hour	X_{11} [%]
	Sector Queue Length	Total waiting flight number in sector per hour	X_{12} [flight]
	Delay Percentage	Ratio of delayed flight number and total flight numbers in sector per hour	X_{13} [%]
	Delay Time	Total delay time of flights in sector per hour	X_{14} [min]
	Average Delay Time	Average delay time of total delayed flights in sector per hour	X_{15} [min]
ATCO Workload	Occupation percentage of air-ground communication radio	Occupation percentage of air-ground communication radio between ATCO and pilots in sector per hour	X_{16} [%]
	Call times of air-ground communication radio	Call times of air-ground communication radio between ATCO and pilots in sector per hour	X_{17} [time]

communication system. The ATC automation system combines and then processes the data of surveyed flight tracks exported from surveillance facilities such as primary and secondary radar, and output integrated flight track information. The automatic message switching system is used to send and receive air traffic service messages via Aeronautical Fixed Telecommunication Network (AFTN). By connecting to this system, air traffic service message contents can be reformatted and analysed for flight plan data. The VHF voice communication system connects ATCO and pilots by wireless communication, and both duration and time of the radio calls can be acquired for ATCO workload assessment after decoding. The integrated flight track information, flight plan information, and voice communication information between ATCO and pilots compose source data for the calculation of all the 17 indicators above. Since the ATC automation system with intrinsic conflict predictors results in STCA and MSAW outputs, the indicators of STCA and MSAW are selected for their high weight in engineering feasibility [20]. Similarly, the indicators of *occupation percentage of air-ground communication radio* and *call times of air-ground communication radio* are selected because the results are acquired easily from the VHF voice communication system.

Principle 2: Scientific typicality

The variables are designed after the previous scientific studies have been sorted out. Some indicators and related measurement methods are selected if they have strong scientific identifications in prior studies, which are suitable for operational performance assessment. For example, the indicators of *air traffic flow density, aircraft climbing frequency, aircraft descending frequency, aircraft speed changing frequency,* and *aircraft heading changing frequency* have high typicality since they are the focus in the studies on complexity of air traffic flow [32-37], while the indicators of *sector saturation* and *sector queue* *length* are also very typical in the studies on traffic congestion status in air traffic flow management [27-28, 40, 42].

Principle 3: Operational practicability

Some indicators and related measurement methods are selected if they have been authorized to play an important role in operational performance metrics. Both the Federal Aviation Administration of the United States (FAA) and the European Organization for the Safety of Air Navigation (EUROCONTROL) have identified the key performance indicators (KPIs) for ATM-related operational performance [25]. Among those KPIs, the indicators of *air traffic flow, delay percentage, delay time, average delay time, flight miles, flight time* are selected and shaped with appropriate measurement methods based on the applicability of ATC sector [25, 29]. In practice, the total delay can be divided into unavoidable delay which has already existed at the entrance of the sector and cannot be recovered any more, and consecutive delay which represents the delay required to solve potential aircraft conflicts or sequencing problem computed from the entrance to the exit of the sector [30-31]. Since there is nothing the ATCOs can do to shorten the unavoidable delay, the delay-related indicators of operational performance are defined as consecutive delay, and measured by comparing the real flight time with the normal flight time.

It should be noted that the index system of *Table 1* is designed for ATC sector mainly in en-route control and terminal control phases. In aerodrome ground control phase, neither the sector concept nor the index system is applicable. In this phase, the focus should be on: (1) occupancy status of runway, taxiway and parking position, i.e., operational trafficability, which could be measured with the occupying aircraft number, occupancy duration or distance; (2) hotspot status in the whole layout of aerodrome, i.e., operational complexity and safety, which could be measured with conflict detection tool, saturation degree of runway, taxiway and parking position; (3) delay status, i.e., operational efficiency, which could be measured with detailed capacity utilization ratio, detailed delay percentage or time, in terms of gate traffic delay, taxi delay and runway delay; and (4) ATCO workload, which is involved with more complex task analysis and measurement. For a special part of en-route control and terminal control phases, it is also necessary to tailor the index system for applicability purpose. Taking a part of final approach as example, some indicators and related measurements on separation between the leading and trailing aircraft should be involved and play an important role in the balance analysis between operational safety and efficiency [40].

3. METHOD DESIGN

3.1 BP Neural Network

BP neural network has been widely applied in decision-making system since it was introduced in 1986. Zhang et al. put forward a comprehensive evaluation method on ATC operation performance based on BP neural network, proving that BP neural network could be used in ATC operation performance classification [42]. The structure of BP neural network based on the introduced multivariate detection index system for operational performance assessment of ATC sector can be shown as in *Figure 1*.

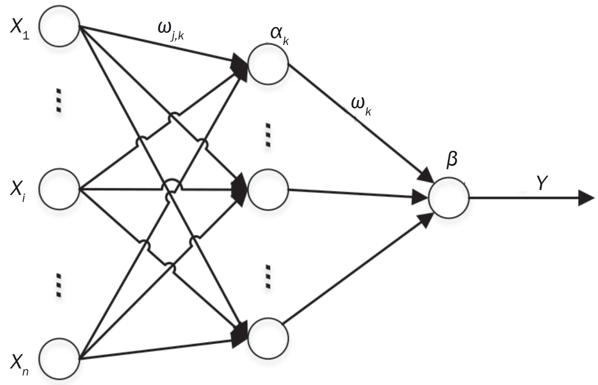

Figure 1 – Structure of BP neural network for operational performance assessment of ATC sector

There are three layers in the BP neural network, i.e., input layer, hidden layer and output layer. In *Figure 1*, X_i (i=1,..,n) is the input as detection indicator, and Y is the output as the ranking result of the operational performance of ATC sector through comprehensive evaluation. α_k and β are the thresholds of the k-th(k=1,...,H) node in the hidden layer and the single node in output layer, respectively. $\omega_{i,k}$ and ω'_k are the weights for input-to-hidden, and hidden-to-output, respectively. The BP neural network could be modelled as follows [43]:

$$I_{1,k} = \sum_{i=1}^{n} \omega_{i,k} \cdot X_i - \alpha_k \tag{1}$$

$$L_k = f_h(I_{1,k}) \tag{2}$$

$$I_2 = \sum_{k=1}^{H} \omega'_k \cdot L_k - \beta \tag{3}$$

$$Y = f_o(I_2) \tag{4}$$

where f_h and f_o are the activation functions of the hidden layer and output layer.

With initial weights and thresholds, the BP neural network will be trained with training input data, to update weights and thresholds until the training errors (difference between trained output and expected output) are small enough to be accepted. Thus, it can be seen that initial weights and thresholds will decide whether or not the training process can result in acceptable training errors within suitable time.

3.2 GABP Neural Network Method

It is clear that BP neural network is built on multi-node architecture which can structurally support comprehensive evaluation on operational performance of ATC sector with multivariate detection index system. Besides, both weights and thresholds are two important criteria for the final evaluation results, since they are the basis for network training. However, initial weights and thresholds are normally assigned randomly in BP network, resulting in the training convergence process being slow, and the BP network easily stuck in local optimum [44-45].

GABP neural network is proposed by the idea that the Genetic Algorithm is capable of global optimization of initial weights and thresholds through selection, crossover and mutation operation. Besides, GA could speed the convergence process of traditional BP neural network, and make it easier to find the global optimum and avoid the local optimum. The idea of using GA to optimize BP neural network has been testified in previous studies [46]. Furthermore, as pointed by one referee, there could be other optional metaheuristics to optimize the BP neural network. To select the proper algorithm, we have compared the performance of BP neural network optimized by Simulated Annealing (SA) algorithm and GA in a primary test, respectively. Results show that GA is better in terms of convergence process and CPU time. So, GA is adopted in our case to optimize the traditional BP neural network (more detailed and thorough comparison among more metaheuristics would be done in our future studies).

To optimize the BP neural network with GA, the initial weights and thresholds will be coded as the chromosome in GA, by selection, crossover and mutation operation, and a group of those chromosomes will be imported into a BP neural network, then the network will be iterated with training input data. In return, the training errors will be used to construct the fitness function in GA, and the chromosome with the best fitness value will be the optimized initial weight and threshold for further BP neural network training.

Here are the steps for GABP neural network method:

Step 1: Establish the training sample set

The training sample set contains training input data and related expected output data. The training input data are from historical database including all 17 detection indicators of operational performance of an ATC sector in a period of time (one hour) as a sample, denoted as $X_1,...,X_{17}$; while expected output data are based on ATC expert panel's ranking for those historical samples, from 1 to 5, indicating that the operational performance is at an excellent, very good, good, average, or bad level, respectively. The panel is made up of 5 senior ATCOs, who are qualified through skill assessment and have more than 10 years of relevant working experience. The ATCOs' ranking for the operational performance is made by their subjective judgments facing the 5 variables (operational trafficability, complexity, safety, efficiency, and ATCO workload) based on observed clues from the historical video and audio records of the operation process of each sample. Normally, it is extremely difficult for each expert to make the same ranking for every sample. If there are 4 ATCOs in the panel or more, giving the same ranking for a sample, namely consistent ranking, the sample is selected as a typical sample. Those typical samples with 17 detection indicators and consistent ranking will compose the training sample set. Besides, on the premise of sufficient number of typical samples, the proportion of training samples with different rank, i.e., 1,2,3,4,5, could be allocated as 10%, 20%, 40%, 20%, and 10%, respectively.

Step 2: Normalize the training sample data

The training sample data should be normalized to eliminate the dimension differences among indicators before further training. Let $x_{j,i}$ and $x'_{j,i}$ be the j-th(j=1,2,...,N) original and normalized training sample data for the i-th(i=1,2,...,17) indicator. The normalization formula is the *mapminmax* function in MATLAB which allows the normalized $x'_{j,i}$ ranging from -1 to 1 [47].

$$x'_{j,i} = 2 \cdot (x_{j,i} - \min(x_{1,i},...,x_{j,i},...,x_{N,i}))/ \\ (\max(x_{1,i},...,x_{j,i},...,x_{N,i}) - \min(x_{1,i},...,x_{j,i},...,x_{N,i})) - 1 \tag{5}$$

Step 3: Initialize BP neural network

BP neural network, in the case, is shaped as 17-input and single-output architecture. While the number of nodes in the hidden layer, H, is normally decided by rule-of-thumb, which is: $H < \sqrt{A+B} + C$ (in the case, A=17 and B=1 are the number of nodes in input and output layers, respectively. C is an integer from 1 to 10 [42]. Besides, for the p-th time of network training, let $L_{j,k,p}$ be the output value of the k-th(k=1,2,...,H) node in the hidden layer with regard to the j-th(j=1,2,...,N) training sample data, and $Y_{j,p}$ be the output value in output layer. Then, both $L_{j,k,p}$ and $Y_{j,p}$ can be derived with the *tansig* activation functions. The *tansig* is a hyperbolic tangent sigmoid transfer function in neural network, *tansig*(*net*)=2/(1+exp(-2 · *net*))-1, where *net* is the linear activation of the neuron [47].

$$L_{j,k,p} = tansig\left(\sum_{i=1}^{17} \omega_{i,k,p} x'_{j,i} - \alpha_{k,p}\right) \tag{6}$$

$$Y_{j,p} = tansig\left(\sum_{k=1}^{H} \omega'_{k,p} L_{j,k,p} - \beta_p\right) \tag{7}$$

$\omega_{i,k,p}$ and $\omega'_{k,p}$ are the weights for input-to-hidden and hidden-to-output during the p-th network training process. $\alpha_{k,p}$ and β_p are the thresholds of the k-th(k=1,...,H) node in the hidden layer and the node in output layer, respectively.

Step 4: Encode the population of GA

The population of GA consists of certain chromosomes which contain initial weights and thresholds of BP neural network. i.e. the chromosome can be coded in the real number as follows:

$$\omega_{1,1,p} \ldots \omega_{17,1,p} \omega_{1,2,p} \ldots \omega_{17,2,p} \ldots \omega_{1,H,p} \ldots \omega_{17,H,p} \alpha_{1,p} \ldots \alpha_{H,p} \omega'_{1,p} \ldots \omega'_{H,p} \beta_p$$

The length of GA chromosome is thus limited by the total number of weights and thresholds, and the number of chromosomes in GA is denoted as Q.

Step 5: Define fitness function

The fitness function of chromosome l(l=1,...,Q) after the d-th iteration is the reciprocal of absolute error between trained output $Y^l_{j,d}$ and expected output $E^l_{j,d}$ [48]:

$$F_{l,d} = \frac{1}{\sum_{j=1}^{N} abs(Y^l_{j,d} - E^l_{j,d})} \tag{8}$$

Step 6: Select

Rolette is used to select individuals from the population to produce new offspring. For each chromosome l, the probability of being selected is $\delta_{l,d} = \frac{F_{l,d}}{\sum_{l=1}^{Q} F_{l,d}}$ [48]

Step 7: Cross

Since the chromosomes are coded as real numbers in the case, so the crossover is done by the so-called real number crossover. Let $g_{l,m,d}$ and $g_{l',m,d}$ be the genes to be crossed, and they are located in the mth position of chromosomes l and l'. Then after crossover the new genes are [48]:

$$\left.\begin{aligned} g'_{l,m,d} &= g_{l,m,d}(1-r) + g_{l',m,d} r \\ g'_{l',m,d} &= g_{l',m,d}(1-r) + g_{l,m,d} r \end{aligned}\right\} \tag{9}$$

$g'_{l,m,d}$ and $g'_{l',m,d}$ are new genes corresponding to $g_{l,m,d}$ and $g_{l',m,d}$; r is a random number from 0 to 1.

Step 8: Mutate

By mutation, $g'_{l,m,p}$ in chromosome l will be mutated to $g''_{l,m,d}$ in certain probability determined by *formulas 10* and *11* [48].

$$g''_{l,m,d} = \begin{cases} g'_{l,m,d} + (g'_{l,m,d} - g_{max}) \cdot f(d) & a > 0.5 \\ g'_{l,m,d} - (g'_{l,m,d} - g_{min}) \cdot f(d) & a \le 0.5 \end{cases} \tag{10}$$

$$f(d) = b\left(\frac{1-d}{D_{max}}\right)^2 \tag{11}$$

Both a and b are random numbers from 0 to 1; g_{max} and g_{min} are the values of upper bound and lower bound for $g'_{l,m,d}$, respectively. And d and D_{max} are current and maximum iterations, respectively.

Step 9: Find the best chromosome

If the maximum iteration D_{max} is reached, then stop running and find the best chromosome with the biggest fitness value to identify the best initial weights and thresholds for GABP neural network training. If not, go back to Step 6 for further iteration.

Step10: Calculate training errors

With the best initial weights and thresholds from GA, BP neural network will be trained with the training sample set. The training errors will be the differences between trained output and expected output.

Step 11: Terminate training

Update weights and thresholds according to the training errors. When any of the targets on training errors' convergence accuracy, training epochs, and other validation check is reached, terminate training.

Step 12: Find the best-trained GABP neural network

By training GABP neural network with different available number of nodes in the hidden layer, find the most suitable number of nodes in the hidden layer which minimizes the total training errors for the best-trained GABP neural network.

Step 13: Evaluate operational performance of ATC sector

Let $x_{t,i}$ and $x'_{t,i}$ be the real-time detected data and normalized data for the ith (i=1,2,...,17) indicator, and normalize $x_{t,i}$ to $x'_{t,i}$ with *Formula 12* [47]:

$$\begin{aligned} x'_{t,i} = {} & 2 \cdot (x_{t,i} - \min(x_{1.i}, \ldots, x_{j,i}, \ldots x_{N,i}))/ \\ & (\max(x_{1.i}, \ldots, x_{j,i}, \ldots x_{N,i}) - \\ & - x_{t,i} - \min(x_{1.i}, \ldots, x_{j,i}, \ldots x_{N,i})) - 1 \end{aligned} \tag{12}$$

Then importing $x'_{t,i}$ to the best-trained GABP neural network, the operational performance rank of ATC sector will be evaluated by GA optimized BP neural network.

Step 14: Alarm

It is feasible to integrate software module implementing the best-trained GABP neural network into application system for real-time usage in the operation site. When the rank result of operational performance of ATC sector is 5, indicating that the performance is at a bad level, an alarm could be a trigger to alert the

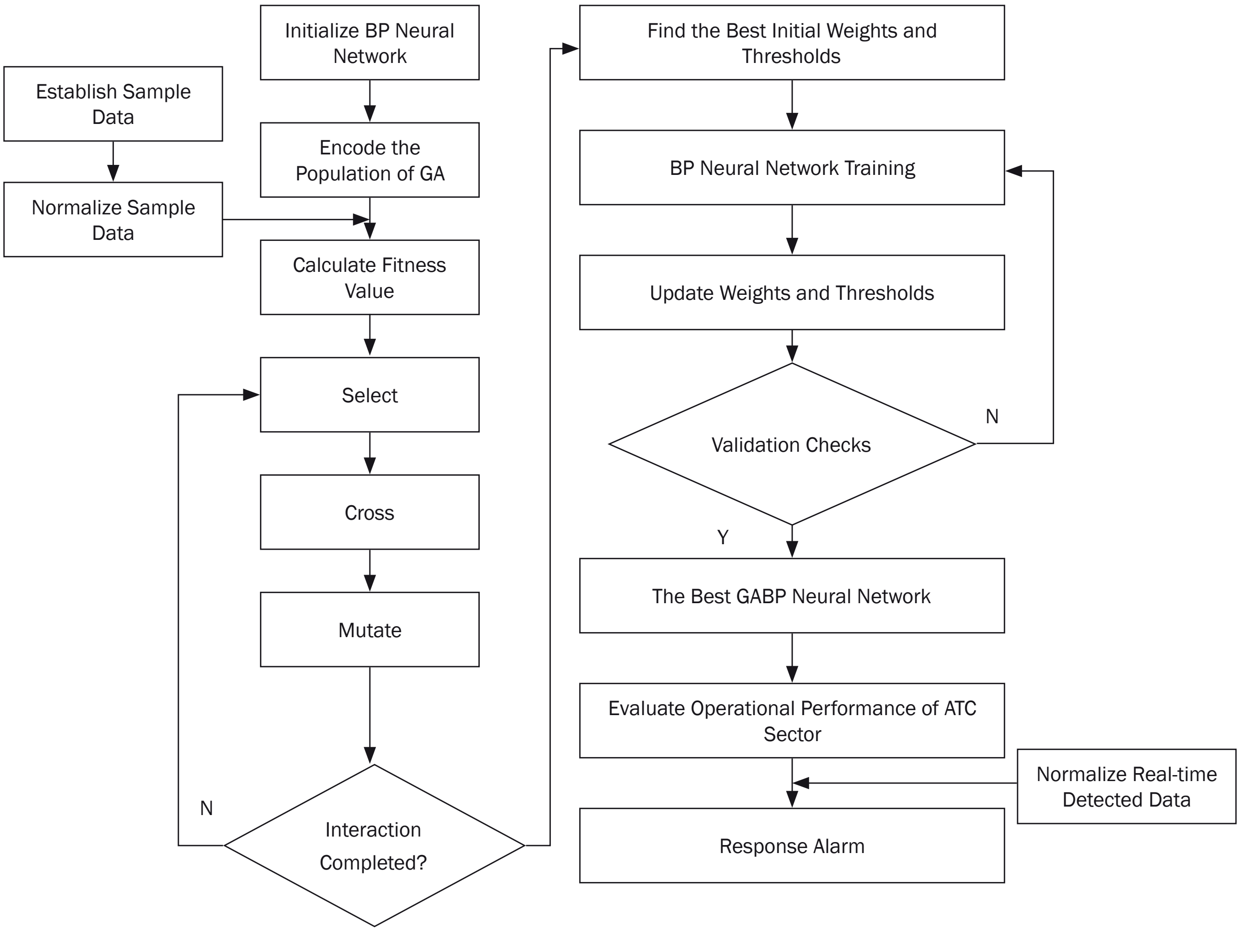

Figure 2 – Flow chart of GABP neural network training process

operation personnel to adjust the operation strategies before it is too late.

Figure 2 presents the flow chart of GABP neural network training process.

4. EMPIRICAL STUDY

To validate the GABP method, an empirical study is conducted in one ATC sector. There are 948 sets of sample data collected in total. A panel of 5 senior ATCOs gives their sample ranks facing the 5 variables (operational trafficability, complexity, safety, efficiency, and ATCO workload) based on observed clues from the historical video and audio records, including the time and space distribution of air traffic flow, the available sector airspace status impacted by meteorological hazards and military operation, the situation of vectoring and holding flights, the aircraft separation biases, the outputs of MSAW and STCA, the duration and frequency of air-ground communication radio calls, and etc., which consumes quite a long time (a huge workload in other words). Finally, 400 sets of sample data remain for further analysis as typical samples, in which 350 sets of sample data are randomly selected for network training, and the other 50 sets of sample data are for network testing. Both the training sample set and the testing sample set have the designed proportion of Step 1. *Table 2* presents an example of the selected sample data.

As mentioned earlier, the number of nodes in the hidden layer is determined by $H < \sqrt{A+B} + C$, which limits the value of H to be [5, 14]. By choosing different values of H with BP neural network, the most suitable number of nodes in the hidden layer which minimizes the total training errors is 5. So, there are 17 nodes in input layers, 5 nodes in the hidden layer, and 1 node in the output layer within the architecture of GABP neural network.

According to the method process, the GABP method is programmed with MATLAB 2014a. The initial value of BP neural network learning rate is 0.05, which is moderate considering that smaller learning rates will slow down the learning process and make it hard to be converged, and larger rates will speed up the learning process and make it unstable. The target on convergence accuracy of training errors is 0.00001, to acquire the better iteration results with higher accuracy. The target on training epochs is 500, and the number of validation checks is 6, which are large enough to get a stable, reliable result with reasonable training and validation process. Besides, the number of chromosomes in GA is 10, and the maximum iteration

Table 2 – Example of sample data

Indicator	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
data	60	2892	512	0.0186	23	4	93	301	94
Indicator	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}	rank
data	92	76.92%	0	8.57%	8	2.67	42%	95	1

number is 50. Too many chromosomes and iteration steps will increase the optimization burden and calculation time. The crossover rate and mutation rate are 0.4 and 0.1, at a moderate level, respectively. Small crossover rates will lead the iteration into local optimum, while a large crossover rate will update the individuals with a high frequency and lose the advantages in existing chromosomes. Large mutation rates will update the chromosomes in a very random way, and small mutation rate will generate new individuals with low possibility. So both too small or large crossover rate and mutation rate will weaken the optimization process.

Meanwhile, since the ranks of operational performance of ATC sector are all integers from 1 to 5, the final results by GABP will be rounded before being exported.

On the other hand, to compare the accuracy of GABP with BP neural network, we process with traditional BP neural network training again with the same training sample set and find the best-trained BP neural network; then import the 50 sets of typical sample data normalized with *Formula 12* into both the best-trained GABP and BP neural networks and export ranking results for testing. The testing errors, i.e., mean error, mean square error, maximum error, minimum error, and error probability, are all presented in *Table 3*.

Table 3 – Testing errors of the best-trained GABP and BP neural network

Testing errors	Mean	Mean Square	Maximum	Minimum	Error probability
GABP	0.18	0.15	1	0	18%
BP	0.32	0.22	1	0	32%

It is clear that the testing errors of GABP neural network are better than those of BP neural network in total. The mean error and mean square error of GABP neural network decrease by 44% and 32% with respect to BP neural network. There are 16 errors among all 50 sets of testing data by BP neural network, while there are only 9 errors out of 50 sets detected in GBAP neural network, so the error probability of BP neural network is 1.78 times of GABP neural network. This indicates that GABP neural network performs better than BP neural network and can assess operational performance of ATC sector with high accuracy. Besides, the evaluation result by GABP neural network is steadier than BP neural network, since the initial weights and thresholds are optimized by GA in GABP neural network, while in BP neural network, both the weights and thresholds are randomly assigned. Thus, GABP neural network can be further generalized for unknown data, which will qualify GABP neural network method with better reliability. *Figure 3* and *Figure 4* present the testing results and testing errors of GABP neural network.

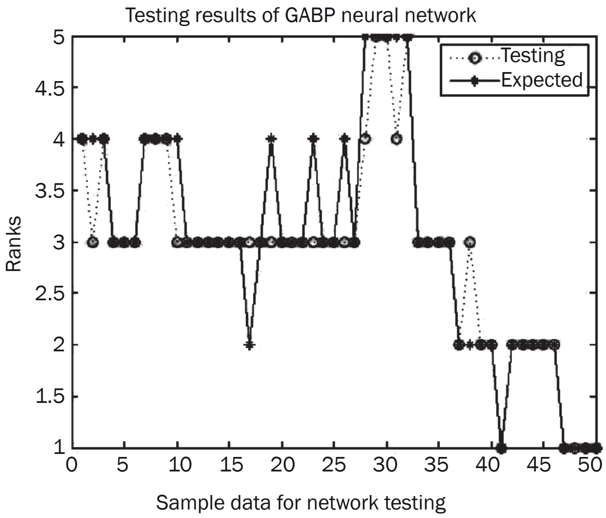

Figure 3 – Testing results of GABP

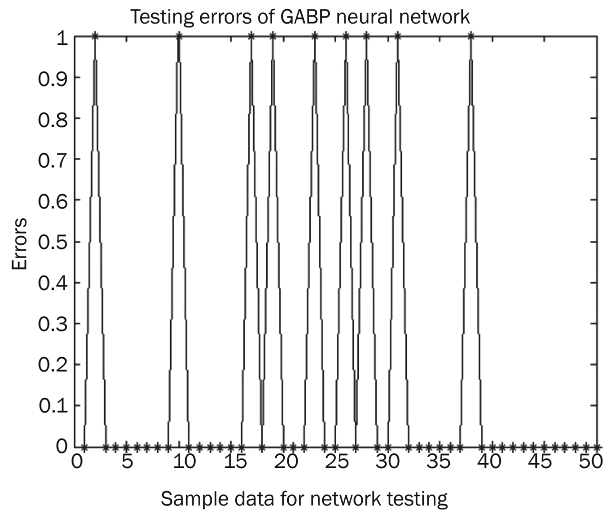

Figure 4 – Testing errors of GABP

With the best-trained GABP neural network, the operational performance assessment of ATC sector can be made by importing the normalized real-time detected data. Alarms could be triggered when the evaluation results reach the value of 5.

In fact, an application system named Air Traffic Control Operational Performance Monitoring System (ATCOPMS), which has been developed and deployed at the Air Traffic Control Center of the Southwest Regional Air Traffic Management Bureau of Civil Aviation of China in Chengdu, could ensure the usability of the GABP neural network method by software update importing the best-trained GABP neural network for respective ATC sector. Data pools could be built up to collect real-time operation data, including integrated flight track information from ATC automation system, air traffic service message from automatic message switching system, and sky-talk records between ATCO and pilot from VHF voice communication system. With these data as input, ATCOPMS can output ATC operational performance indicator detection results within seconds, based on the real-time operation situation of all 17 indicators. Moreover, the comprehensive evaluation results on ATC operational performance could be output within minutes by some artificial intelligence methods including traditional BP neural network, according to the real-time operation situation [42]. *Figure 5* presents the network topology layout of the system. Triggered by the data mining based on ATCOPMS, the Air Traffic Control Center in Chengdu works out an operation management decision-making mechanism, probing the weaknesses objectively based on the operational performance indicator detection results and related comprehensive evaluation results. It then finds the reasons (solving problems in other words) based on the playback of related video and audio records from various angles, such as operation model, airspace status and sector configuration, military restriction, meteorological hazard, expertise level of ATCO. The mechanism is validated in Chengdu to be effective in the optimization of ATC operation strategies, taking on-duty roster for example, and airspace planning and utilization, taking flexible airspace usage with military side for example [40]. Also, more positive contributions of the system could be predictable after the software updating for the application of the GABP neural network method.

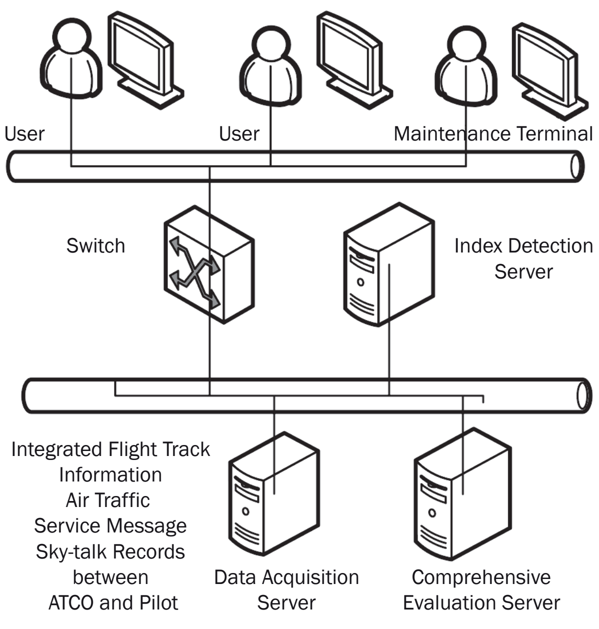

Figure 5 – Network topology layout of ATCOPMS

5. CONCLUSIONS

Most of the existing studies on operational performance assessment of ATC sector are limited to a single side and cannot apply to real-time assessment at the frontline of air traffic service provider. To solve these issues, this paper constructs a quantitative detection index system with 5 variables and 17 indicators through expert investigation method based on the principles of engineering feasibility, scientific typicality, and operational practicability, which covers most of the related factors influencing operational performance of ATC sector, and provides an improved comprehensive evaluation method based on BP neural network optimizing the initial weights and thresholds by GA. On the one hand, by sorting out the existing studies, the multivariate detection index system accommodates operational trafficability, operational complexity, operational safety, operational efficiency, and ATCO workload, which has the characteristics of integration. Meanwhile, all indicators are settled by senior ATC experts, so that this index system can be applied to any ATC system with the characteristics of universality.

On the other hand, the GABP neural network method has higher accuracy and more stable generalization ability, which can be better suitable to operational performance assessment of ATC sector than the traditional BP neural network in terms of comparison of testing errors.

Another important feature is that, taking ATCOPMS in Chengdu for validation, since all the 17 indicators of the multivariate detection index system can be accessed in the operation site, the GABP neural network method could be realized by importing the best-trained neural network for respective ATC sector into application system for the frontline, which will improve the operation management efficiency by presenting real-time assessment result and essential alarm to air traffic service provider, avoiding fatigue and poor timeliness caused by traditional experience-based assessment.

To sum up, the proposed GABP method for operational performance assessment of ATC sector is improved in comprehensiveness by establishing an integrated and quantifiable detection index system, accuracy and reliability by optimizing BP neural network method with Genetic Algorithm, and practicability by updating related software module in application system for the frontline. Last but not least, it should be noted that the assessment results based on the proposed method are relevant only to the sectors from which the initial data have been gathered because of the limitation of the generalization ability of the neural network, which indicates the guidance of future studies.

ACKNOWLEDGEMENTS

The authors would like to thank three anonymous referees and the Editor for their valuable comments and suggestions on our draft paper. The research was partly financially supported by the National Science Foundation jointly with the Civil Aviation Administration of China (No. U1333202) and the Fundamental Research Funds for the Central Universities (No. NJ20140018).

张建平，段力伟，郭静，刘卫东，杨晓嘉，张瑞平

基于GABP神经网络的空中交通管制扇区运行性能评估方法

摘要：为评估空中交通管制扇区运行性能，构建了涵盖运行通行性、运行复杂性、运行安全性、运行经济性以及管制员工作负荷等5类一级指标、17类二级指标的扇区运行性能检测指标体系。利用遗传算法（GA）优化反向传播（BP）神经网络的权值和阈值设置，提出了基于嵌套遗传算法的反向传播（GABP）神经网络的扇区运行性能综合评价算法。实例采集使用了某扇区400组样本数据，其中350组样本进行网络训练、50组进行网络测试，确定了GABP神经网络的17-5-1三层拓扑结构形式。经采用GABP神经网络和传统BP神经网络2种方法进行测试，揭示GABP神经网络方法在误差均值、误差均方差、误差概率等参数上优于传统BP神经网络方法，验证了GABP神经网络方法具有很好的评估精度和泛化能力。所提扇区运行性能测评方法具有全面、准确、可靠和实用的特点，并能够应用于空中交通管制单位实现扇区运行性能实时测评及响应告警。

关键字：空中交通管制扇区；运行性能；检测指标体系；遗传算法；BP神经网络；综合评价

REFERENCES

[1] Ball M, Barnhart C, Nemhauser G, et al. Air transportation: irregular operations and control. Handbooks in Operations Research and Management Science. 2007;14(1):1-67.

[2] Bennell JA, Mesgarpour M, Potts CN. Airport runway scheduling. 4OR–Quarterly Journal of Operations Research. 2011;4(2):115-138.

[3] Reich PG. Analysis of long-range air traffic systems: separation standards I. Journal of Navigation. 1966;19(1):88-98.

[4] Reich PG. Analysis of long-range air traffic systems: separation standards II. Journal of Navigation. 1966;19 (2):169-186.

[5] Reich PG. Analysis of long-range air traffic systems: separation standards III. Journal of Navigation. 1966;19(3):331-347.

[6] Anderson D, Lin XG. Collision risk model for a crossing track separation methodology. Journal of Navigation. 1996;49(2):337-349.

[7] Brooker P. Lateral collision risk in air traffic systems: a Post-Reich event model. Journal of Navigation. 2003;56(3):399-409.

[8] Brooker P. Longitudinal collision risk for ATC track systems: a hazardous event model. Journal of Navigation. 2006;59(1):55-70.

[9] Xu X, Li D, Li X. Research on safety assessment of flight separation. ACTA Aeronautica et Astronautica Sinica. 2008;29(6):1411-1418.

[10] Li D, Xu X, Li X. Research on safety target levels of air collision. Chinese Journal of Ergonomics. 2008;14(2):41-44.

[11] Brooker P. Aircraft collision risk in the North Atlantic region. Journal of the Operational Research Society. 1984;35(8):695-703.

[12] Cox ME, Harrison D, Moek G, et al. European studies to investigate the feasibility of using 1000 ft vertical separation minima above FL 290, Part I. Journal of Navigation. 1991;44(2):171-183.

[13] Cox M E, Harrison D, Moek G, et al. European studies to investigate the feasibility of using 1000 ft vertical separation minima above FL 290, Part II. Journal of Navigation. 1992;45(1):91-106.

[14] Cox ME, Harrison D, Moek G, et al. European studies to investigate the feasibility of using 1000 ft vertical separation minima above FL 290, Part III. Journal of Navigation. 1993;46(2):245-261.

[15] ICAO. Doc9689-AN/953(1st Edition)-1998. Manual on Airspace Planning Methodology for the Determination of Separation Minima. Montreal, Canada: ICAO; 1998.

[16] Brooker P. P-RNAV, safety targets, blunders and parallel route spacing. Journal of Navigation. 2004;57(3):371-384.

[17] ICAO. Doc9750-AN/963(3rd Edition)-2007. Global Air Navigation Plan. Montreal, Canada: ICAO; 2007.

[18] ICAO. Doc9613-AN/937(3rd Edition)-2008. Performance-based Navigation Manual. Montreal, Canada: ICAO; 2008.

[19] ICAO. Doc9883(1st Edition)-2009. Manual on Global Performance of the Air Navigation System. Montreal, Canada: ICAO; 2009.

[20] Kuchar JK, Yang LC. A review of conflict detection and resolution modeling methods. IEEE Transactions on Intelligent Transportation Systems. 2000;4(1):179-189.

[21] Barnhart C, Fearing D, Odoni A, et al. Demand and capacity management in air transportation. EURO Journal on Transportation and Logistics. 2012;1(1-2):135-155.

[22] Pellegrini P, Rodriguez J. Single European sky and single European railway area: A system level analysis of air and rail transportation. Transportation Research – Part A: Policy and Practice. 2013;57(1):64-86.

[23] Wu D. A new model to assess airspace capacity based on controller workload. Journal of Wuhan University of Technology (Transportation Science & Engineering). 2014;38(4):846-849.

[24] Tian Y, Yang S, Wan L, Yang Y. Research on the method of sector dynamic capacity evaluation. System Engineering-Theory & Practice. 2014;34(8):2163-2169.

[25] FAA, EUROCONTROL. 2012 Comparison of Air Traffic Management-Related Operational Performance: U.S./Europe. Washington D.C., U.S.A.: FAA Communications, 2013-AJR-887; 2013.

[26] Li Y, Hu M, Xie H, Peng Y. Terminal area utilization rate evaluation based on extension multi-level state classification. System Engineering and Electronics. 2013;35(12):2533-2539.

[27] Yan S, Yao L, Zhao Y. Research on the Methods of Air Traffic Congestion Measures. Journal of Transportation Engineering and Information. 2009;7(1):11-16.

[28] Zhao Y, Chen K. Air traffic congestion assessment method based on evidence theory. Proceedings of the 2010 Chinese Control and Decision Conference (CCDC 2010). 2010 May 26-28; Xuzhou, China. Institute of Electrical and Electronics Engineers (IEEE); 2010. p. 426-429.

[29] Air Traffic Management Bureau of CAAC. Statistical methods of civil aviation normal flights. Beijing, China: CAAC; 2012.

[30] Samà M, D'Ariano A, Pacciarelli D. Rolling horizon approach for aircraft scheduling in the terminal control area of busy airports. Transportation Research-Part E: Logistics and Transportation Review. 2013;60(1): 140-155.

[31] Samà M, D'Ariano A, Pacciarelli D. Optimal aircraft scheduling and routing at a terminal control area during disturbances. Transportation Research-Part C: Emerging Technologies. 2014;47(1):61-85.

[32] Radio Technical Commission for Aeronautics (RTCA). Final report of RTCA task force 3: free flight implementation. Washington D.C., U.S.A.: RTCA Inc.; 1995.

[33] Laudman IV, Shelden SG, Branstrom R, et al. Dynamic density: an air traffic management metric. Washington, D.C., U.S.A.: NASA/TM-1998-112226; 1998.

[34] Delahaye D, Puechmorel S, Hansman RJ, et al. Air traffic complexity map based on nonlinear dynamical systems. Air Traffic Control Quarterly. 2004; 12(4):367-388.

[35] Gianazza D, Guittet K. Selection and evaluation of air traffic complexity metrics. Portland, U.S.A.: 25th Digital Avionics Systems Conference; 2006.

[36] Lee K, Feron E, Pritchett A. Describing airspace complexity: airspace response to disturbances. Journal of Guidance, Control and Dynamics. 2009;132(1): 210-222.

[37] Cong W, Hu M, Xie H. Research on refinement method of air traffic complexity metrics system. Journal of Wuhan University of Technology (Transportation Science & Engineering). 2014;38(3):611-618.

[38] ICAO. Doc9426-AN/924 (1st Edition)-1984. Air Transport Service Planning Manual, Part II, Appendix C. Montreal, Canada: ICAO; 1999

[39] Arnab M, Washington YO. The factors affecting air traffic controller workload: a multivariate analysis based upon simulation modeling of controller workload. Transportation Research Record. 2002;1788:58-69.

[40] Zhang J, Hu M, Liu W. Integrated Evaluation for Operation Performance of Air Traffic Control in Terminal Area. Journal of Southwest Jiaotong University. 2012;02:341-347.

[41] Zhang J, Yu H, Zou G. Research on evaluation factors for operation performance of air traffic control in terminal area. Journal of Civil Aviation University of China. 2012;03:18-21.

[42] Zhang J, Hu M, Wu Z, Zhang R. An Improved Integrated Evaluation Method on Operation Performance of Air Traffic Control Based on BP Network. Journal of Southwest Jiaotong University. 2013;03:553-558.

[43] Wang L, Zeng Y, Chen T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. Expert Systems with Applications. 2015;42(2):855-863.

[44] Roudbari A, Saghafi F. Intelligent modeling and identification of aircraft nonlinear flight. Chinese Journal of Aeronautics. 2014;27(4):759-771.

[45] Li Y, Wang R, Xu M. Rescheduling of observing spacecraft using fuzzy neural network and ant colony algorithm. Chinese Journal of Aeronautics. 2014;27(3):678-687.

[46] Wang DM, Wang L, Zhang GM. Short-term wind speed forecast model for wind farms based on genetic BP neural network. Journal of Zhejiang University (Engineering Science). 2012;46(5):837-841.

[47] Demuth H, Beale M. Neural Network Toolbox. User's Guide; 2015.

[48] Yu Y-Y, Chen Y, Li T-Y. Improved genetic algorithm for solving TSP. Control and Decision. 2014;29(8):1483-1488.