JOURNAL OF ADVANCED TRANSPORTATION
J. Adv. Transp. 2016; **50**:315–325
Published online 25 June 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/atr.1321

ATCEM: a synthetic model for evaluating air traffic complexity

Mingming Xiao, Jun Zhang, Kaiquan Cai* and Xianbin Cao

School of Electronics and Information Engineering, Beihang University, Beijing, China

SUMMARY

Air traffic complexity, which measures the disorder of air traffic distribution, has become the critical indicator to reflect air traffic controller workload in air traffic management (ATM) system. However, it is hard to assess the system accurately because there are too many correlated factors, which make the air traffic complexity nonlinear. This paper presents an air traffic complexity evaluation model with integrated classification using computational intelligence (ATCEM). To avoid redundant factors, critical factors contributing to complexity are analyzed and selected from numerous factors in the ATCEM. Subsequently, to construct the mapping relationship between selected factors and air traffic complexity, an integrated classifier is built in ATCEM. With efficient training and learning based on aviation domain knowledge, the integrated classifier can effectively and stably reflect the mapping relationship between selected factors and the category of air traffic complexity to ensure the precision of the evaluation. Empirical studies using real data of the southwest airspace of China show that the ATCEM outperforms a number of state-of-the-art models. Moreover, using the critical complexity factors selected in ATCEM, the air traffic complexity could be effectively estimated. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS: air traffic complexity; complexity factors; integrated classification

1. INTRODUCTION

Air traffic complexity is the quantitative description of the disorder of air traffic distribution [1] and acts as an indicator of air traffic controller (ATCo) workload [2, 3]. It relates to the characteristics of both traffic flow pattern and airspace structure. Nowadays, air traffic situation is becoming increasingly complex because dynamic optimizations of traffic flow and airspace configuration are employed to safely accommodate the rapid development of air transportation. As a result, in air traffic management (ATM) system, more ATCo resources are needed to keep aircraft flying safely, efficiently, and expeditiously. To reasonably allocate the limited ATCo resources over airspace, for example, to reconfigure sectors and assign ATCos to sectors, it is important to accurately measure the air traffic complexity.

To assess the traffic complexity in real air traffic operation, two challenges need to be tackled. First and foremost, the complexity factors are hard to select because too many correlated factors affect air traffic complexity [4, 5]. The factors include static factors and dynamic factors. The static factors are fixed and given by the spatial and physical characteristics of airspace (or a sector), such as terrain, number of airways, air route structure, crossing waypoints, and navigation aids. The dynamic factors vary as a function of time and depend on the air traffic pattern, for example, the number/density of aircraft, separation between aircraft, mix of aircraft, proximity of aircraft, divergence/convergence of aircraft, aircraft sensitivity, and aircraft speeds. Second, the process of driving air traffic complexity is a very complicated mechanism, because of not only many influencing factors but also severe nonlinearities [1, 6].

Therefore, it is very hard to develop a precise mathematical model for air traffic complexity. In literature, many existing studies focus on creating more relevant factors to explain the air traffic

*Correspondence to: Kaiquan Cai, School of Electronics and Information Engineering, Beihang University, Mailbox 8, No. 37 Xueyuan Road, Haidian District, Beijing 100191, China. E-mail: logcai@163.com

complexity, such as traffic density, intrinsic attribute [1], and probabilistic factor [7]. But, a single factor can only reflect air traffic complexity in a certain aspect.

In order to research air traffic complexity from a holistic point of view, some studies have proposed a "dynamic density (DD)" model [8–10] to measure control-related workload as a weighted sum of selected complexity factors, and the weights of these factors were determined through regression analysis of the complexity data. Because the functional relationship between complexity factors and workload is nonlinear, complicated, and largely unknown, a "black-box" approach is employed. For example, Gianazza *et al.* [11, 12] applied a back-propagation neural network (BPNN) to model the relationship between a small set of six complexity factors and three levels of air traffic complexity. The six factors are selected from 28 studied factors, and the three levels of air traffic complexity are used to guide sector reconfiguration. Neural network is a nonlinear modeling tool that can extract and learn the relationship between input and output data. However, its instability and lack of robustness [13] reduce its accuracy in high-speed air transportation system.

In ATM domain, the main motivation of evaluating air traffic complexity is to reflect ATCo workload [2–4]. The higher the air traffic complexity is, the higher the workload will be. The information about whether airspace is overloaded or not is critical to provide guidelines to ATCo resource allocation and traffic flow management to ensure a safe operation environment for multiple aircraft. Based on analysis of the problem, air traffic complexity evaluation can be accurately characterized as a classification problem [12]. It can be defined as a process of classifying the levels of air traffic complexity in airspace (or sector) according to complexity factors. Only three levels of air traffic complexity, that is, low, medium, and high, need to be considered. Low complexity denotes that the required workload ensuring aircraft safety exceeds what ATCo can provide, while the opposite is true for high complexity. Medium complexity means this is a balance between the required workload and the provided workload.

Based on the analysis in the preceding text, this paper presents an effective method, namely air traffic complexity evaluation model with integrated classification (ATCEM), to classify air traffic complexity. First, critical complexity factors are analyzed and selected from numerous factors in the ATCEM to avoid redundant factors and scale up the evaluation model. Thereafter, to construct the mapping relationship between selected factors and air traffic complexity, an integrated classifier is built in ATCEM. In this classifier, a number of weak classifiers have been adaptively integrated as a strong classifier inspired by [13, 14] using evolutionary learning based on aviation domain knowledge. With efficient training and learning, the integrated classification method in ATCEM can effectively and stably reflect the mapping relationship between selected factors and air traffic complexity. Empirical studies using real data from the southwest airspace of China show that the ATCEM outperforms a number of state-of-the-art models. Moreover, the critical complexity factors selected in ATCEM can be used for evaluating air traffic complexity effectively.

The rest of this paper is organized as follows: Section 2 reviews and analyzes the air traffic complexity factors. The ATCEM is proposed in Section 3. The specific genetic algorithm (GA)-based method for selecting complexity factors and integrated classification method are described in detail. Section 4 presents the experimental studies, which include critical factor selection and empirical comparison between ATCEM and state-of-the-art models. Lastly, conclusions and future work are presented in Section 5.

2. AIR TRAFFIC COMPLEXITY FACTORS REVIEW AND ANALYSIS

An "air traffic complexity factor" is a parameter or attribute that influences the level of air traffic complexity. Till now, research on complexity factors has drawn much attention. Kopardekar *et al.* [8–10] reviewed some complexity factors, proposed since 1963, and identified that aircraft count, sector geometry, traffic flows, separation standard, aircraft performance characteristics, and weather are the most common factors that affect air traffic complexity. Delahaye and Puechmorel [1] presented some novel intrinsic factors, that is, aircraft density, divergence, convergence, and insensitivity to measure air traffic complexity. Lee *et al.* [15] defined air traffic complexity by the sum of all aircraft's heading changes in response to an intrusive aircraft within a sector. Moreover, Prandini [7] presented a probabilistic factor to measure midterm traffic complexity based on aircraft's intent information and current state. Crespo and Weigang [16] developed an agent evaluation function to airspace DD.

DOI: 10.1002/atr

A list of 28 complexity factors that have been consistently found to be relevant to air traffic complexity has been presented in Table I. For a more thorough review of the listed factors, readers can refer to the cited source literature.

The listed factors in Table I are defined from different perspectives. It is hard to describe the air traffic complexity accurately using a single factor. For example, there are four aircraft in both sectors A and B in Figure 1. The sectors are of the same size, but the spatial attributes (mainly the horizontal

Table I. Notations and definitions of the list of 28 air traffic complexity factors [11].

Notation	Definition
Nb	Number of aircraft within a sector [8, 9, 21, 23]
Nb^2	Squared number of aircraft within a sector [8, 9]
N_{ds}	Number of descending aircraft within a sector [9, 21, 23]
N_{cl}	Number of climbing aircraft within a sector [9, 21, 23]
F_5, F_{15}, F_{30}, F_{60}	Incoming flow of the controlled sector in future 5, 15, 30, and 60 minutes, respectively [3]
inter_hori	Number of potential crossings (irrespective of the aircraft direction on their trajectories) with angle greater than 20° [11]
inter_vert	The ratio of flight phase (stable/climbing/descending) [11]
creed_ok	Conflict perception indicator of pairs of aircraft in which vertical separation occurs prior to separation [11, 24]
creed_pb	Conflict perception indicator of pairs of aircraft in which vertical separation does not occur prior to separation [11, 24]
σ_{gs}^2	Variance of ground speed of aircraft within a sector [9, 21]
σ_{gs}/gs	Ratio of standard deviation of speed to average speed [9, 21]
avg_vs	Average vertical speed of aircraft within a sector [11]
vpro_1	Vertical proximity measure 1, inverse of the mean weighted horizontal separation distance between aircraft pairs [8, 9, 21]
vpro_2	Vertical proximity measure 2, inverse of the average minimum horizontal separation distance between aircraft pairs [8, 9, 21]
hpro_1	Horizontal proximity measure, inverse of the mean weighted vertical separation distance between aircraft pairs [8, 9, 21]
V	Geometric volume of a sector [11]
Dens	Density indicator (mean) [1, 3, 11]
track_disoder	Variability in headings (mean) [1, 3, 11]
speed_disoder	Variability in speed (mean) [1, 3, 11]
Div	Divergence between pairs of aircraft in the controlled sector (mean) [1, 3, 11]
Conv	Convergence between pairs of aircraft in the controlled sector (mean) [1, 3, 11]
sensi_d	Sensitivity indicator that measures the difficulty in solving potential conflicts in case of divergence in the controlled sector (mean); a situation with a "high sensitivity" is easier to resolve for the air traffic controller than one with a "low sensitivity" [1, 3, 11]
insen_d	Insensitivity indicator in case of divergence in the controlled sector (mean) [1, 3, 11]
sensi_c	Sensitivity indicator that measures the difficulty in solving potential conflicts in case of convergence in the controlled sector (mean); a situation with a "high sensitivity" is easier to resolve for the air traffic controller than one with a "low sensitivity" [1, 3, 11]
insen_c	Insensitivity indicator in case of convergence in the controlled sector (mean) [1, 3, 11]

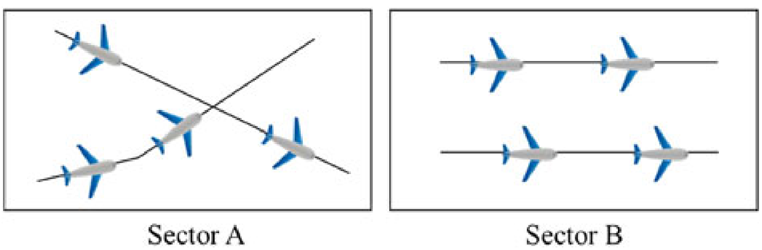

Figure 1. An example of comparison of air traffic complexity between two sectors.

DOI: 10.1002/atr

proximity, i.e., *hpro*_1) within them are different. Potential conflict exists in sector A. It is seemingly more difficult to keep the traffic safe in section A, and the air traffic complexity of sector A is significantly higher than sector B. Hence, the air traffic complexity is the consequence of joint contributions of a group of complexity factors.

In addition, too many input factors and redundancy among these factors lead to over-fitting when constructing a mapping relationship between the factors and air traffic complexity. For example, there are at least three factors (e.g., *inter_hori*, *creed_ok*, and *creed_ok*) to describe the influence of potential conflict in traffic complexity.

Therefore, it is important to obtain a useful set of factors to evaluate air traffic complexity. In literature, different models (i.e., DD and BPNN model) have their own sets of complexity factors. Even for the same DD model, different studies (e.g., FAA William J. Hughes Technical Center, NASA Ames Research Center, and Metron Aviation) have used different sets of complexity factors. Their factors are mainly selected either depending on domain-specific expertise or by principal component analysis and sequential forward selection (PCA&SFS) [11], a feature selection method without considering the nonlinear correlation among features. Till now, no comprehensive and generally accepted set of complexity factors has been defined [3].

To effectively evaluate the air traffic complexity in any traffic operation scenario, it is important to select an effective and comprehensive set of complexity factors. In this paper, the listed 28 factors, which have been consistently found to be relevant to air traffic complexity, constitute a "factor pool". Subsequently, adequate complexity factors will be selected from this "factor pool" to effectively evaluate the air traffic complexity.

3. AIR TRAFFIC COMPLEXITY EVALUATION MODEL WITH INTEGRATED CLASSIFICATION

As stated in Section 1, air traffic complexity evaluation can be defined as the process of classifying traffic complexity into different levels according to a set of complexity factors. The framework of ATCEM is shown in Figure 2. ATCEM is a synthetic model, consisting of input variables, output variables, and a mapping relationship between inputs and outputs. The mapping relationship consists of factor selection and classification.

The input variables are the factors that contribute to air traffic complexity. Here, "factor pool" (defined in Section 2), which consists of a number of important air traffic complexity factors, is an ideal input set.

The output parameters are the levels of air traffic complexity in airspace (or sector). Here, air traffic complexity is classified into three levels, that is, low, medium, and high level. Our ATCEM can also accommodate more detailed classification of air traffic complexity. Low complexity means traffic pattern is simple (e.g., low-density traffic operation and no conflict) and the workload needed to ensure air traffic safely is below the workload ATCo can provide. Adjacent low-complexity sectors can be merged to save ATCo resources. On the contrary, high complexity means traffic is hard to control (e.g., high-density air traffic and high number of potential crossings and potential conflicts) and the workload needed for keeping air traffic safe exceeds the workload ATCo can provide. Medium complexity means a good balance on the sector, where the workload needed to ensure aircraft safety is approximately equal to the workload ATCo can provide.

Because the functional relationship between complexity factors and level of air traffic complexity is nonlinear, complicated, and largely unknown, an integrated classifier that is adaptively integrated

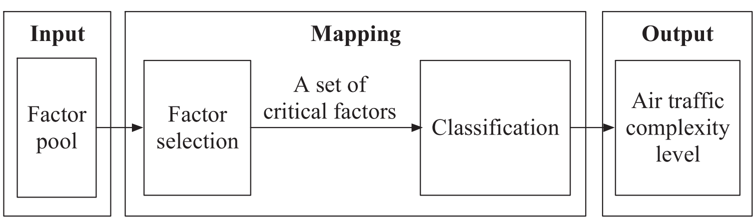

Figure 2. The framework of air traffic complexity evaluation model with integrated classification.

DOI: 10.1002/atr

based on a number of nonlinear weak classifiers to extract and learn the relationship between input and output data can be a good candidate approach in evaluating the air traffic complexity. However, too many input variables and redundancy among these variables would lead to over-fitting while constructing the mapping relationship between input and output data [11]. Hence, before training a classifier for this mapping, selecting critical complexity factors as classifier's input is important and necessary in ATCEM.

Consequently, the ideas for constructing our ATCEM are as follows: First, an effective and comprehensive set of complexity factors (without redundant factor) is selected as input set for evaluating air traffic complexity level. Then, based on aviation domain knowledge, an integrated classification method to construct the mapping relationship between selected factors and air traffic complexity is proposed.

In the following subsections, details of factor selection and integrated classification employed in the ATCEM will be introduced.

3.1. *Complexity factor selection with genetic algorithm (GA)*

From the analysis of complexity factors in Section 2, it is not simple to select critical complexity factors. On one hand, absence of any critical factors in ATCEM would result in incorrect classification of air traffic complexity level. On the other hand, existence of redundant factors would lead to over-fitting during the construction of mapping relationship between input and output data in ATCEM. Various techniques have been proposed to select an optimum subset of features from a larger set of possible features [17]. Based on the advantages of intelligent algorithm, that is, ensuring the chosen feature set is optimum and without considering the relationship among features, a specific GA-based method has been designed to select an effective and comprehensive set of complexity factors from the "factor pool" in ATCEM.

The flowchart of the GA-based selection method is shown in Figure 3. In this method, a nonlinear classifier is applied to classify the levels of air traffic complexity, and fitness is calculated with classification accuracy. Here, an evolutionary BPNN [18, 19] is generated and trained as the classifier. Air traffic complexity is classified into $Y=\{1,2,\ldots,K\}$ levels, where $K=3$ and labels 1, 2, and 3 represent low, medium, and high levels of air traffic complexity, respectively.

The sample space for training the nonlinear classifier is obtained from n_S sectors and n_T time periods and denoted as $S=\{(f_1,y_1),(f_2,y_2),\ldots,(f_i,y_i),\ldots,(f_N,y_N)\}$. Here, $N=n_S\cdot n_T$ is the number of sample data, and f_i and $y_i\in Y$ are the vector of the 28 complexity factors and the levels of air traffic complexity,

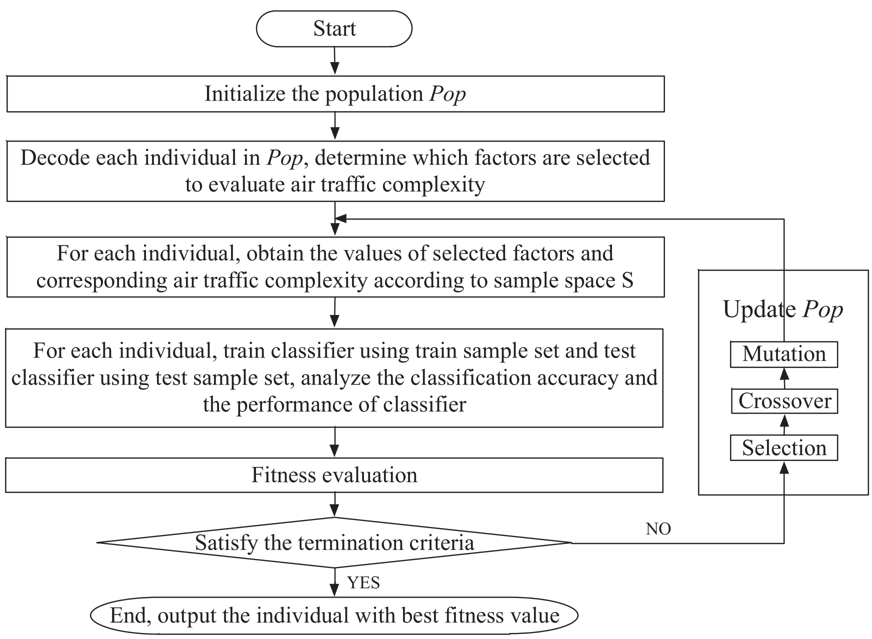

Figure 3. The flowchart of genetic algorithm-based selection method.

respectively, in a given sector and time period. All the samples were divided randomly into two parts with 60% and 40% samples. They were used as train set S_{train} and test set S_{test}, respectively.

Let *Max_Gen* be the maximum number of generations in the whole optimization, *NP* be the population size, and α and β be the crossover rate and mutation rate, respectively. The details of GA-based selection method are as follows:

(1) Encoding: Each individual is encoded as a binary vector $[x_1, x_2, \ldots, x_{28}]$, in which, if $x_i = 1$, the i-th factor in the "factor pool" is chosen as an input parameter of ATCEM; otherwise, the i-th factor is not chosen.
(2) Initialization: First, the population *Pop* is initialized, that is, all genes in *NP* individuals in *Pop* are generated randomly as either 0 or 1. Then, the classifier (a single BPNN) is initialized for classifying air traffic complexity. The classifier consists of an input layer of 28 complexity factors and an output layer of a probability vector, denoted as $[O_L, O_M, O_H]$. O_L, O_M, and O_H are the corresponding probabilities of classifying the air traffic into low complexity, medium complexity, and high complexity based on input data, respectively. The target of $[O_L, O_M, O_H]$ is [1, 0, 0], or [0, 1, 0], or [0, 0, 1].
(3) Fitness evaluation: In order to avoid the influence on fitness value caused by the randomness of initial weights and threshold values of BPNN, the GA [19] is applied to evolve the weights and threshold values of BPNN. The fitness of each individual in population *Pop* is evaluated according to Equation (1).

$$\text{Fitness} = \left|\overline{BIC}^2_{\text{train}} - \overline{BIC}^2_{\text{test}}\right| \cdot [8 - (P^G_{\text{train}} + P^L_{\text{train}} + P^M_{\text{train}} + P^H_{\text{train}} + P^G_{test} + P^L_{\text{test}} + P^M_{\text{test}} + P^H_{\text{test}})] \quad (1)$$

Here, the subscripts "train" and "test" denote the results of training set and testing set, respectively. P^G is the global proportion of correct classification of input vectors, and P^L, P^M, and P^H are the percentages of correct classifications for the low, medium, and high levels of air traffic complexity, respectively. $\overline{BIC}$ represents the performance of BPNN in terms of goodness of fit and is calculated by Equation (2) [11]. In Equation (2), λ is the number of unadjusted parameters of the BPNN (i.e., the number of weights and threshold values of BPNN), K is the total levels of air traffic complexity (we have $K = 3$), N is the size of sample data (i.e., the size of training set or test set), ln*(N)* is the log-likelihood, and $t^{(n)}$ and $y^{(n)}$ are the target and output vectors, respectively. The lower fitness is, the better BPNN has been trained and the better set of complexity factors obtained.

$$\overline{BIC} = 2 \cdot \lambda \cdot \ln(N) - 2 \cdot \sum_{n=1}^{N} \sum_{k=1}^{K} t_k^{(n)} \cdot \ln\left(y_k^{(n)}\right) \quad (2)$$

(4) Genetic operators: Roulette wheel selection operator, two-point crossover operator, and the uniform mutation operator are employed for evolving the population *Pop*.

3.2. Integrated classification in ATCEM

Having selected a set of critical factors, how these factors affect the air traffic complexity is still unknown. Here, an integrated classification method has been employed to construct the mapping relationship between selected factors and air traffic complexity.

In this classification method, a number of weak classifiers with strong nonlinear mapping ability between input and output data, such as BPNN [14, 18, 19], are generated and trained, and then all these classifiers are adaptively integrated together in an ensemble based on aviation domain knowledge to form a strong classifier to evaluate air traffic complexity. Let T be the number of weak classifiers. The flowchart of integrated classification in ATCEM is provided in Figure 4.

In our ATCEM, T evolutionary BPNNs [18, 19], which can better reflect nonlinear relationship between input and output data by evolving their weights and threshold values with evolutionary algorithm, are generated and trained as weak classifiers. To improve the accuracy and stability of weak classifiers, an adaptive boosting algorithm [20] is utilized to integrate these evolutionary BPNNs. Adaptive boosting is a meta-algorithm that can be used in conjunction with many other learning

DOI: 10.1002/atr

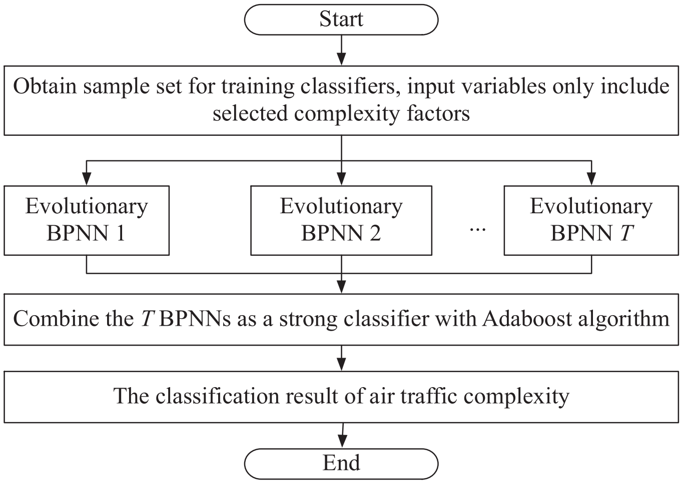

Figure 4. The flowchart of integrated classification in air traffic complexity evaluation model with integrated classification. BPNN, back-propagation neural network; AdaBoost, adaptive boosting.

algorithms to improve their performance. Hence, the outputs of all weak classifiers are combined using a weighted sum to represent the final output of our ATCEM.

The sample space here, denoted as $S=\{(f_{o1}, y_1), (f_{o2}, y_2), \ldots, (f_{oi}, y_i), \ldots, (f_{oN}, y_N)\}$, is the same as that stated in Section 3.1 except for the vector of complexity factor. $f_{oi} = \left[f_{oi}^1, f_{oi}^2, \ldots, f_{oi}^l\right]$ is the vector of the l selected complexity factors (i.e., the optimum set of factors selected from the "factor pool" with GA-based selection method, which will be detailed in Section 4.1). The vectors of complexity factors are used as input of the integrated classifier, while different levels of air traffic complexity are used as the output of ATCEM (denoted as $[p_L, p_M, p_H]$), whose targets are [1, 0, 0], [0, 1, 0], and [0, 0, 1].

4. EXPERIMENTAL STUDIES

In this section, experiments are carried out to identify a set of critical complexity factors in our ATCEM, and analyze the performance of ATCEM by comparing with existing models for measuring air traffic complexity.

All experiments were carried out on data from southwest airspace of China, which consists of seven sectors (Figure 5). The database used was extracted between 12:00 AM and 04:00 PM on 28 July 2010, provided by the Air Traffic Control Center of China, including radar data of aircraft, the characteristics

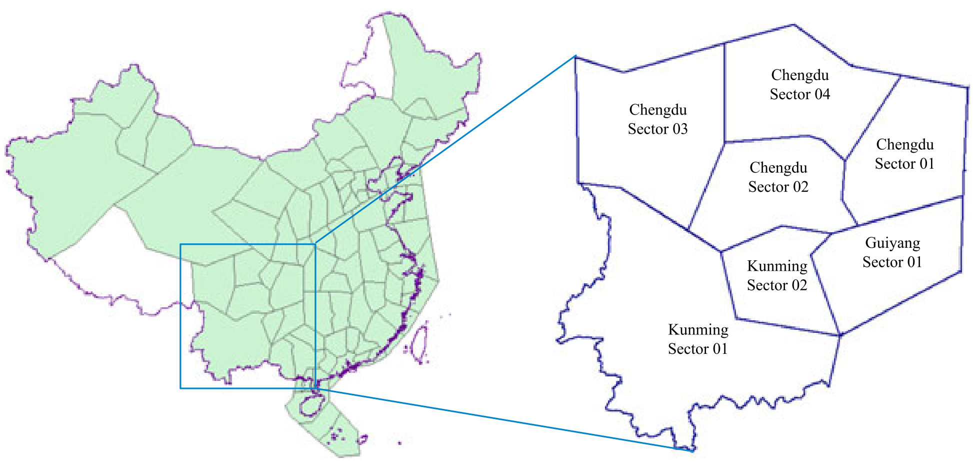

Figure 5. The southwest airspace of China and the corresponding seven sectors.

DOI: 10.1002/atr

of sectors, and air traffic complexity of sectors. The radar data of aircraft were recorded every 4 seconds, and the degrees of sector complexity (low, medium, or high air traffic complexity) were recorded every minute of the day. All complexity factors in the “factor pool” were computed every minute for each sector according to the radar data of aircraft. A total of 6720 samples were obtained and divided randomly into two parts with 60% and 40% samples, respectively, which were used as train set S_{train} and test set S_{test} for ATCEM. For each model throughout our experiments, the results were obtained on the basis of 25 independent runs to estimate the performance of ATCEM model under the stochastic behavior that may be caused by the randomness in GA.

4.1. Air traffic complexity factor selection

The first experiment is aimed to select an effective and comprehensive set of complexity factors, which will be used as input variables of our ATCEM, with GA-based selection method. The parameters of GA-based selection method are presented in Table II.

Seven air traffic complexity factors are selected from the “factor pool”. Details of these selected factors are presented as follows:

(1) Volume of sector (*V*): It is a geometric attribute of a controlled sector. In general, a sector with larger size can accommodate more flights under separation standard, which leads to less workload. In Reference [11], sector’s volume has been proved to greatly improve the prediction of sector status, that is, the level (low, medium, or high level) of air traffic complexity.
(2) Number of aircraft within a sector (*Nb*): For a sector, more aircraft operating will increase the traffic complexity and induce more workload. The volume of a sector and the number of aircraft within a sector can be used together for roughly comparing complexity between the sectors. But usually, the traffic complexity cannot be measured based on these two variables only (see the example in Figure 1 in Section 2).
(3) Incoming flow of the controlled sector in future 5 minutes (F_5): This variable predicts the number of aircraft that will enter into the controlled sector. The more the number of aircraft, the more will be the complexity.
(4) Ratio of standard deviation of speed to average speed (σ_{gs}/gs): This variable can depict the disorder of aircraft operation. σ_{gs} is the standard deviation of aircraft speed within controlled sector. Low standard deviation indicates less performance variation (i.e., speed disorder) between aircraft. All aircraft flying at the same lower constant speed will be easy to control. Hence, with same average speed, higher σ_{gs}/gs will increase air traffic complexity.
(5) Number of potential crossings (*inter_hori*): This variable counts the potential crossings (irrespective of the aircraft direction on their trajectories) with angles greater than 20° [11]. It has a significant impact on increasing sector complexity because potential conflicts may occur during controlling aircraft that have potential crossings.
(6) Vertical proximity (*vpro_2*): It is defined as the inverse of average minimum vertical separation between aircraft pairs [21]. Higher vertical proximity would increase the air traffic complexity and cause controller to focus the attention on this pair of aircraft because of the possibility of separation violation.
(7) Sensitivity indicator (*sensi_c*): It measures how fast aircraft are moving towards each other and represents difficulty solving in potential conflicts in case of convergence in a controlled sector. A situation with a “high sensitivity” is easier to resolve for the ATCo than one with a “low sensitivity”.

Table II. The parameters of GA-based selection method for selecting critical complexity factors.

Parameters	Description	Value
NP	Population size	20
n	Number of complexity factors	28
Max_Gen	Number of generations in overall optimization	100
α	Crossover rate	0.7
β	Mutation rate	0.05

GA, genetic algorithm.

DOI: 10.1002/atr

With the aforementioned seven factors, our GA-based selection method searched for the best fitness value, which related to percentage of correct classification and performance of classifier (Equation (1)). It is probably because the combination of these seven factors can cover almost all critical characteristics of air traffic in a sector. For example, the incoming flow of a controlled sector in future 15, 30, and 60 minutes (F_{15}, F_{30}, F_{60}) can be roughly predicted based on the information of F_5 and σ_{gs}/gs; the intrinsic attributes [11] of air traffic (from the 20th factor to 28th factor in Table I) can be represented by a combination of σ_{gs}/gs, *inter_hori*, *vpro*_2, and *sensi_c*. Further advantages of the selected set of factors will be presented in Section 4.2.

4.2. *Comparing ATCEM to existing model*

This experiment is aimed to evaluate the efficacy of ATCEM by comparing it with existing models. As stated in Section 1, neural network can model the nonlinear complexity relationship between input and output data and has been applied for air traffic complexity evaluation. Hence, the evaluation model proposed by Gianazza *et al.* [11] is applied for comparison, including the structure of BPNN and the complexity factor selection method, that is, combined PCA and SFS (PCA&SFS) method. We denote this compared model as BPNN_PCA in this paper.

With PCA&SFS factor selection method, six complexity factors are selected. They are sector volume (*V*), number of aircraft within a sector (*Nb*), incoming flow of the controlled sector in future 60 minutes (F_{60}), aircraft density indicator (*Dens*), sensitivity indicator that measures the difficulty to solve potential conflicts in case of divergence in the controlled sector (*sensi_d*), and variability in headings (*track_disorder*). Two of these factors, *V* and *Nb*, are also selected with our GA-based algorithm.

In order to evaluate the efficacy of each evaluation model, both $\overline{BIC}$ value and percentages of correct classification of sector complexity (i.e., $\{P^G, P^L, P^M, P^H\}$) are taken into account. The lower $\overline{BIC}$ is, the better is the model obtained in terms of goodness of fit and model complexity. The higher the percentage of correct classification is, the better is the air traffic complexity evaluation model obtained. The $\overline{BIC}$ value and classification accuracies calculated from both training phase and testing phase were analyzed statistically in terms of their average values and standard deviations over 25 independent runs for each model, and the results are presented in Table III.

Because both factor selection method and classification method are different between ATCEM and BPNN_PCA model, it is unclear whether superiority exists in selecting complexity factors with GA-based method. Hence, we have further compared the ATCEM with IC_PCA model to investigate the contribution of selecting complexity factors with GA-based method. IC_PCA is the same as ATCEM, but with selection method replaced by PCA&SFS. The results obtained by IC_PCA are also shown in Table III.

The output of all three models is the probability vector $[O_L, O_M, O_H]$. O_L, O_M, and O_H are the corresponding probabilities of classifying the air traffic into low complexity, medium complexity, and

Table III. Comparisons among ATCEM, BPNN_PCA, and IC_PCA for evaluating air traffic complexity.

Evaluating model	Results				
ATCEM	$\overline{BIC}_{\text{train}}$	P^G_{train}	P^L_{train}	P^M_{train}	P^H_{train}
	1.8253 (0.0255)	**78.6%** (**0.0053**)	76.5% (**0.0111**)	**81.6%** (**0.0157**)	**76.4%** (**0.0177**)
	$\overline{BIC}_{\text{test}}$	P^G_{test}	P^L_{test}	P^M_{test}	P^H_{test}
	2.2376 (**0.0360**)	**76.7%** (**0.0054**)	75.7% (0.0279)	**79.4%** (**0.0193**)	**73.8%** (0.0192)
BPNN_PCA	$\overline{BIC}_{\text{train}}$	P^G_{train}	P^L_{train}	P^M_{train}	P^H_{train}
	2.7406 (3.7562)	73.9% (0.0798)	75.5% (0.1582)	74.4% (0.1557)	72.1% (0.0601)
	$\overline{BIC}_{\text{test}}$	P^G_{test}	P^L_{test}	P^M_{test}	P^H_{test}
	3.1359 (3.8558)	71.9% (0.0789)	75.2% (0.1585)	71.9% (0.1507)	69.9% (0.0667)
IC_PCA	$\overline{BIC}_{\text{train}}$	P^G_{train}	P^L_{train}	P^M_{train}	P^H_{train}
	1.8693 (**0.0241**)	76.0% (0.0067)	**79.3%** (0.0124)	77.3% (0.0167)	72.2% (0.0186)
	$\overline{BIC}_{\text{test}}$	P^G_{test}	P^L_{test}	P^M_{test}	P^H_{test}
	2.2549 (0.0367)	74.0% (0.0080)	**77.9%** (**0.0225**)	75.2% (0.0226)	70.0% (**0.0126**)

ATCEM, air traffic complexity evaluation model with integrated classification; PCA, principal component analysis; BPNN, back-propagation neural network.

high complexity based on input data, respectively. The target of $[O_L, O_M, O_H]$ is [1, 0, 0], or [0, 1, 0], or [0, 0, 1]. These models will classify the air traffic complexity to a level that has the highest probability. For each level of air traffic complexity, percentage of correct classification (P^L, P^M, or P^H) is calculated based on real data samples and is analyzed statistically in terms of their average values and standard deviations over 25 independent runs for each model.

In each row of the table, the best value is highlighted in boldface. It is clear that ATCEM has significantly improved $\overline{BIC}$ and almost all the classification accuracies for sector complexity. P^G is the global proportion of correct classification for all input vectors, and P^L, P^M, and P^H are the percentages of correct classifications for the low, medium, and high levels of air traffic complexity, respectively. The probabilities P^G and P^H have been increased up to 5% and 4%, respectively, which are very critical in ATM system. Besides, the standard deviation of the $\overline{BIC}$ value and all classification accuracies obtained by ATCEM are much lower than those obtained by BPNN_PCA, which means ATCEM is much more stable on air traffic complexity evaluation. We also note that the ATCEM, whose input complexity factors are obtained by GA-based selection method, performs better than IC_PCA, especially in terms of P^G, P^M, and P^H.

Furthermore, the Wilcoxon rank-sum test [22] was carried out on the results obtained by 25 runs of the three compared models, and the one that is significantly better than the others (with the significance level of 5%) is underlined. It can be observed that ATCEM significantly outperformed the BPNN_PCA model.

Benefiting from our GA-based factor selection and integrated classification, it is not surprising to find the superiority of ATCEM for air traffic complexity classification. GA-based selection method can ensure the chosen factors set is optimum without considering the relationship among factors, and moreover, the seven selected factors $\{V, Nb, F_5, \sigma_{gs}/gs, inter\_hori, vpro\_2, sensi\_c\}$ in our ATCEM cover almost all important characteristics of air traffic in a sector. The superiority of our selected factors set can be noted from the comparison between ATCEM and IC_PCA. Furthermore, from the comparison between IC_PCA and BPNN_PCA, it can be seen that integrated classifier performs better than single classifier, that is, the BPNN. The integrated classification method, which decreases the $\overline{BIC}$ value and the standard deviation of $\overline{BIC}$ value and improves the classification accuracy, is more stable and effective to reflect the nonlinear complex relationship between selected factors and air traffic complexity.

5. CONCLUSION AND FUTURE WORK

In ATM system, it is critical to effectively measure the air traffic complexity. In this paper, a novel ATCEM is presented to measure the air traffic complexity. ATCEM is composed of input variables, output variables, and a classifier that constructs mapping relationship between input and output data. A specific GA-based selection method is presented to select an effective and comprehensive set of complexity factors as input variables. Then, an integrated classification is built in ATCEM to classify the level of the air traffic complexity according to selected factors. With efficient training and learning, ATCEM has selected seven efficient and critical factors, and with these selected factors, ATCEM can effectively and stably evaluate the air traffic complexity. Empirical studies on the real-world data of the southwest airspace of China clearly demonstrated the efficacy and stability of ATCEM.

Although preliminary results are promising, we firmly believe that the ATCEM presented here can be further improved to increase the classification performance of air traffic complexity. A more effective integration method to form a stronger classifier with weak classifiers is under investigation to enhance the performance of our ATCEM.

6. LIST OF ABBREVIATIONS

ATCEM	Air traffic complexity evaluation model with integrated classification
ATCo	Air traffic controller
ATM	Air traffic management
BPNN	Back-propagation neural network
BPNN_PCA	An air traffic complexity evaluation model with BPNN and PCA&SFS
DD	Dynamic density

DOI: 10.1002/atr

FAA	Federal aviation administration
GA	Genetic algorithm
IC_PCA	The same air traffic complexity evaluation model as ATCEM, but factor selection method replaced by PCA&SFS
NASA	National aeronautics and space administration
PCA&SFS	principal component analysis and sequential forward selection

ACKNOWLEDGEMENTS

This work is partially supported by the National Basic Research Program of China (grant no.2011CB707004), Key Project in the National Science & Technology Pillar Program (grant no.2012BAG04B01), National Science Fund for Distinguished Young Scholars (grant no. 61425014), and Foundation for Innovative Research Groups of the National Natural Science Foundation of China (grant no. 61221061).

REFERENCES

1. Delahaye D, Puechmorel S. Air traffic complexity: towards intrinsic metrics. *The 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy* 2000.
2. Prandini M, Piroddi L, Puechmorel S, Brazdilova SL. iFly deliverable D3.1, complexity metrics applicable to autonomous aircraft. Project no. TREN/07/FP6AE/S07.71574/037180 IFLY, 2009.
3. Djokic J, Lorenz B, Fricke H. Air traffic control complexity as workload driver. *Transportation Research Part C-Emerging Technologies* 2010; **18**(6): 930–936.
4. Mogford RH, Guttman JA, Morrow SL, Kopardekar P. The complexity construct in air traffic control: a review and synthesis of the literature. FAA Technical Note, DOT/FAA/-CT TN95/22, 1995.
5. Hilburn B. Cognitive complexity in air traffic control: a literature review. Eurocontrol Technical Report, EEC Note No. 04/04, 2004.
6. Prandini M, Piroddi L, Puechmorel S, Brazdilova SL. Toward air traffic complexity assessment in new generation air traffic management systems. *IEEE Transactions on Intelligent Transportation Systems* 2011; **12**(3): 809–818.
7. Prandini M. A probabilistic measure of air traffic complexity in 3-D airspace. *International Journal of Adaptive Control and Signal Processing* 2010; **24**(10): 813–829.
8. Kopardekar P. Dynamic density: a review of proposed variables. FAA WJHTC internal document, overall conclusions and recommendations, Federal Aviation Administration, 2000.
9. Kopardekar P, Magyarits S. Measurement and prediction of dynamic density. *The 5th USA/Europe ATM R&D Seminar*, 2003.
10. Kopardekar P, Schwartz A, Magyarits S, Rhodes J. Airspace complexity measurement an air traffic control simulation analysis. *International Journal of Industrial Engineering-Theory Applications and Practice* 2009; **16**(1): 61–70.
11. Gianazza D, Guittet K. Selection and evaluation of air traffic complexity metrics. *The 25th Digital Avionics Systems Conference* 2006; 1-12.
12. Gianazza D. Forecasting workload and airspace configuration with neural networks and tree search methods. *Artificial Intelligent* 2010; **174**(7-8): 530–549.
13. Hansen L, Salamon P. Neural network ensembles. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 1990; **12**(10): 993–1001.
14. Hippert H, Pedreira C, Souza R. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Transaction on Power Systems* 2001; **16**(1): 44–55.
15. Lee K, Feron E, Pritchett A. Air traffic complexity: an input-output approach. *Proceedings of the 2007 American Control Conference*, New York, United states, 2007.
16. Crespo AF, Weigang L. Agent evaluation function considering airspace Dynamic Density. *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2010.
17. Guyon I, Elisseeff A. An introduction to variable and feature selection. *The Journal of Machine Learning Research* 2003; **3**: 1157–1182.
18. Bishop CM. *Neural Networks for Pattern Recognition*. *Oxford University Press*, Inc.: New York, 1995.
19. White D, Ligomenides P. GANNet: a genetic algorithm for optimizing topology and weights in neural network design. *New Trends in Neural Computation, Lecture Notes in Computer Science* 1993; **686**: 322–327.
20. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* 1997; **55**(1): 119–139.
21. Chatterji GB, Sridhar B. Measures for air traffic controller workload prediction. *Proceedings of the First AIAA Aircraft Technology, Integration and Operations Forum*, Los Angeles, CA, 2001.
22. Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin* 1945; **1**(6): 80–83.
23. Sridhar B, Sheth KS, Grabbe S. Airspace complexity and its application in air traffic management. *Proceedings of the 2nd USA/Europe Air Traffic Management 1998 R&D Seminar*, Orlando, Florida, 1998.
24. Averty P. Conflict perception by ATCS admits doubt but not inconsistency. *Proceedings of the 3rd USA/Europe Air Traffic Management R&D Seminar*, 2000.

DOI: 10.1002/atr