Research
Smart Process Manufacturing: Deep Integration of AI and Process Manufacturing—Article

A Data and Knowledge Collaboration Strategy for Decision-Making on the Amount of Aluminum Fluoride Addition Based on Augmented Fuzzy Cognitive Maps

Weichao Yue, Weihua Gui, Xiaofang Chen *, Zhaohui Zeng, Yongfang Xie

School of Information Science and Engineering, Central South University, Changsha 410083, China

A B S T R A C T

In the aluminum reduction process, aluminum fluoride (AlF_3) is added to lower the liquidus temperature of the electrolyte and increase the electrolytic efficiency. Making the decision on the amount of AlF_3 addition (referred to in this work as MDAAA) is a complex and knowledge-based task that must take into consideration a variety of interrelated functions; in practice, this decision-making step is performed manually. Due to technician subjectivity and the complexity of the aluminum reduction cell, it is difficult to guarantee the accuracy of MDAAA based on knowledge-driven or data-driven methods alone. Existing strategies for MDAAA have difficulty covering these complex causalities. In this work, a data and knowledge collaboration strategy for MDAAA based on augmented fuzzy cognitive maps (FCMs) is proposed. In the proposed strategy, the fuzzy rules are extracted by extended fuzzy k-means (EFKM) and fuzzy decision trees, which are used to amend the initial structure provided by experts. The state transition algorithm (STA) is introduced to detect weight matrices that lead the FCMs to desired steady states. This study then experimentally compares the proposed strategy with some existing research. The results of the comparison show that the speed of FCMs convergence into a stable region based on the STA using the proposed strategy is faster than when using the differential Hebbian learning (DHL), particle swarm optimization (PSO), or genetic algorithm (GA) strategies. In addition, the accuracy of MDAAA based on the proposed method is better than those based on other methods. Accordingly, this paper provides a feasible and effective strategy for MDAAA.

1. Introduction

The aluminum reduction cell, hereafter referred to as "the cell," is a complex multi-variable system, which is characterized by energy balance and mass balance coupling. The electrolyte temperature (ET) can be reduced by decreasing the liquidus temperature based on aluminum fluoride (AlF_3) addition, and thus reducing the loss of molten aluminum [1,2]. A well-shaped hearth can be obtained with a precise AlF_3 feeding amount (AFA) to a certain degree [3]. Some research indicates that a well-shaped cell hearth will result in high current efficiency [4,5]. However, an inaccurate AFA may cause a large fluctuation of the side ledge (SL), which will prevent the ideal energy equilibrium from being achieved. As a result of the inherent complexity of the reduction process, making the decision on the amount of AlF_3 addition (MDAAA) mainly relies on technicians and experts. However, it is difficult for inexperienced technicians to perform this task. Because experienced experts may not always be available, circumstances of excess or insufficient AFA frequently occur. Therefore, it is desirable for an accurate AlF_3 addition to be determined using a scientific strategy.

These problems have attracted the attention of researchers. There are three types of research on MDAAA, all of which mainly focus on controlling the AlF_3 concentration (C_{AlF_3}). The first type of research takes an empirical approach that depends on understanding the dynamic of AlF_3. C_{AlF_3} is monitored by analyzing electrolyte samples, which is done very sporadically. This method has revealed a very strong correlation between C_{AlF_3} and temperature [6]. Temperature and electrolyte sample analysis with a time lag (TL) are used in C_{AlF_3} adjusting strategies in the control feedback loop; building a logic rule base is the core method for these

---

strategies [7–9]. The second type of research considers AFA as a function of deviation from a target C_{AIF_3} and/or temperature. In practice, C_{AIF_3} was found to change with the SL thickness, and some linear regression models for MDAAA were proposed [10–12]. In the third type of research, strategies are proposed based on the AIF_3 mass balance and/or energy balance. MDAAA models have been built by analyzing AIF_3 evolution from cells, and C_{AIF_3} control strategies were developed based on estimation and decoupling techniques with detailed process and plant knowledge [13–17]. The methods in the first type of research always rely on human experience, and it is easy for human subjectivity to influence knowledge model construction. Because of the complexity of making a decision about the amount of AIF_3, it is difficult for methods of the second type to capture all of the complex features of AIF_3 addition. Due to the detection of dead zones in the aluminum reduction cell, it is difficult to implement refined AIF_3 addition using methods of the third type, which are based on AIF_3 mass balance and/or energy balance.

Existing research on MDAAA mainly focuses on data-driven or knowledge-driven methods alone. However, data-driven methods may fail to cover the complex characteristics of the cell, and knowledge-driven methods may be overly subjective. Therefore, it is desirable to develop a model that combines historical data with the experience of experts. To address this challenge, modeling with fuzzy cognitive maps (FCMs) seems practical, as it is characterized by intuition and the simplicity of causal representations [18]. FCMs have been widely used in decision analysis, control, modeling, and prediction [19–22]. Here, we present only a few examples. In Ref. [23], in order to track the maximum power point of a photovoltaic array, FCMs were used for the fuzzy controller. In Ref. [24], a fuzzy multiple attribute decision-making model was built in combination with the technique for order of preference by similarity to ideal solution (TOPSIS) and FCMs. In Ref. [25], FCMs were used to assess the performance of a tunnel-boring machine, and experiential knowledge was captured and utilized.

FCMs consist of concepts and edges; the former introduce the qualitative analysis, while the latter quantitatively indicate the causality [26]. Each concept represents a numeric state. The causal relationship between concepts, which is provided by domain experts in most current achievements, denotes the influencing degree to which a concept changes other concepts [27]. The utilization of experience to identify concepts and edge strength is the core of these methods [28]. However, subjectivity may cause inaccuracy of FCMs modeling [29]. At present, the existing contributions of FCM learning algorithms are mainly divided into three classes: Hebbian-based [29–31], evolutionary-based [32,33], and hybrid-based [34,35], where the latter combines Hebbian-based and evolutionary-based learning algorithms. Although these approaches are widely used for training FCMs, it is sometimes difficult to guarantee avoidance of the local optimization and to identify the global optima.

In this study, a data and knowledge collaboration strategy for MDAAA is proposed, combined with experiential knowledge from experts and data from the aluminum reduction process. The available data is used to extract meaningful fuzzy rules based on fuzzy decision trees and the clustering method, and is also used to detect the edge strength using the state transition algorithm (STA). The initial structure of MDAAA provided by experts is amended using the above fuzzy rules. The problem of having to rely on authoritative experts for FCMs modeling can then be alleviated. The accuracy of MDAAA modeling based on FCMs is sensitive to the edge strength [29], which can be relaxed by detecting strength using the STA. Based on the augmented FCMs, the AFA can be obtained by removing the normalization of the concepts. To the best of our knowledge, this is the first time that a collaboration model that simultaneously integrates expert knowledge with production data

is used for MDAAA based on augmented FCMs. In this study, the validity of the proposed strategy is verified.

The outline of this paper is as follows. Section 2 analyzes the role and evolution of AIF_3, and describes the difficulties of and solutions to MDAAA. Section 3 provides the details of fuzzy decision trees and extended fuzzy k-means (EFKM), which are used to extract fuzzy rules. The STA is then introduced to detect strength. Section 4 describes the initial structure design and the learning problem. Section 5 models the MDAAA based on augmented FCMs, verifies the effectiveness of the proposed strategy, and provides the discussion. The last section gives the conclusions.

2. Analyses of AlF₃ addition

2.1. Role analysis of AlF₃

Research on the aluminum reduction process has been going on for over a century; due to its complexity and high degree of nonlinearity, optimal cell operations are still significant challenges around the world. The superheat degree is the D-value between the ET and the liquidus temperature. During the process, the Al_2O_3 dissolves in the electrolyte at a suitable superheat, and metal aluminum is produced on the cathode [4]. Ref. [2] shows that the current efficiency will decrease by 1.2%–1.5% when the ET increases by 10 °C. AIF_3 addition makes it possible to maintain a suitable superheat with a low ET; this is the most important factor in lowering the liquidus temperature because of the loss of AIF_3, which must be added as required in the reduction process. However, with unreasonable AIF_3 addition, the anode effect may occur frequently as the solubility of Al_2O_3 decreasing, resulting in the destruction of the energy balance and mass balance. When the superheat increases with excessive AIF_3 addition, SL melting occurs, which is damaging to the cell life. Therefore, MDAAA is crucial for the aluminum reduction process. The schematic of a cell is shown in Fig. 1.

2.2. Evolution analysis of AlF₃

AIF_3 evolution is divided into two types, based on its characteristics. The first type involves the neutralization reaction of AIF_3. Alumina contains certain impurities—mainly Na_2O and CaO—which neutralize AIF_3. The second type of evolution involves the emission and recycling of AIF_3; during this process, a mass of particles together with hydrogen fluoride volatilize from the cell at a high temperature, including particulate AIF_3, $NaAlF_4$, CaF_2, Na_3AlF_6, and Al_2O_3. In addition, an aspirator is used for exhaust collection. The unstable $NaAlF_4$ resolves into $Na_5Al_3F_{14}$ and AIF_3 as the temperature decreases. The exhaust is then purified with Al_2O_3, and the AIF_3 from the purification is added to the cell [14].

2.3. Difficulties of MDAAA

Many factors must be considered in the MDAAA process, increasing the complexity of the decision-making problem. Like most industry processes, the aluminum reduction process interacts with external disturbance and internal environment, as illustrated by Fig. 2. However, it is also different from most industrial processes. The ET is very high, at above 960 °C, and complex electrochemical reactions occur in the strong magnetic and corrosive environment.

Three aspects of the external disturbance and internal environmental have an influence on the aluminum reduction process:

(1) External disturbances such as an anode change (AC), aluminum tapping (AT), or beam raising (BR), will disturb the energy equilibrium, causing a variation in C_{AIF_3}.
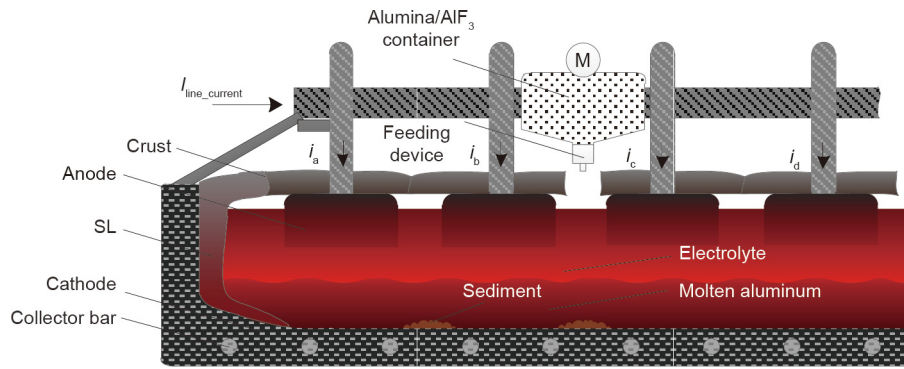
Fig. 1. Schematic of an aluminum reduction cell. $I_{line\_current}$, i_a, i_b, i_c, i_d are series current of cell, and anode current of each anodic bar, respectively; and M is metering installation for feeding.
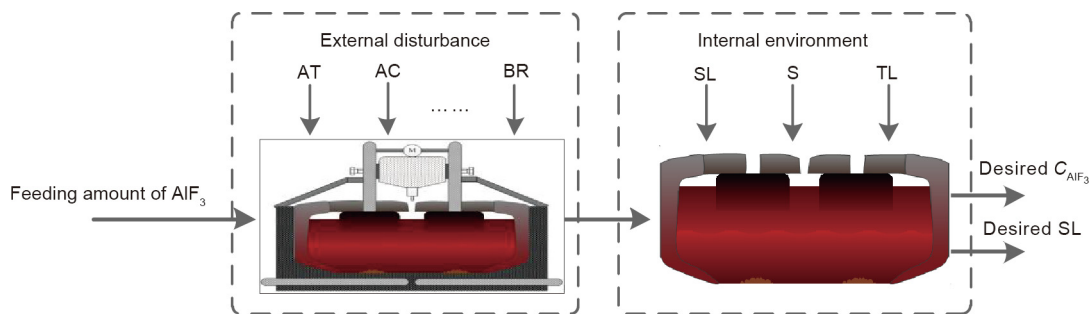

Fig. 2. The external disturbance and internal environment of the aluminum reduction process. AT: aluminum tapping; AC: anode change; S: sludge.

(2) The SL and sludge will melt or solidify with variation of the superheat, which also affects C_{AlF_3}. However, as a result of the high temperature, the degree of influence is imponderable.

(3) Because of the TL of the cell, the response of the AlF₃ addition will be delayed. Therefore, the temperature will not decrease immediately. Over time, the SL will become thinner with an increase in the superheat, resulting in a heat loss. Because of the thin SL, the C_{AlF_3} will decrease with more solidified electrolyte melting into the cell. After that, the ET will slightly recover.

Correspondingly, the difficulties of AlF₃ addition increase in practice, including the following:

(1) Disturbances: External operations and the environment will have an imponderable influence on C_{AlF_3}.

(2) Difficulties of mechanism modeling: Due to erosion and the high temperature of the electrolyte, there is no mature real-time detection service to analyze the electrolyte composition. C_{AlF_3} is detected by means of time-consuming sampling and chemical analyzing. The variation in C_{AlF_3} will be unavailable in real time. In addition, the C_{AlF_3} evolutionary process is complex, and it will be difficult to derive a mechanism model to describe the process theoretically.

(3) Insufficient/excessive addition: Due to the TL of the cell, the cell is out of sync between the influence of the AlF₃ addition and the variation of the superheat; this makes it easy to carry out insufficient or excessive addition.

2.4. Solutions to MDAAA

Based on the above analysis, a new solution for AlF₃ addition is proposed, as shown in Fig. 3. The solution contains two main aspects: the data-driven method and the knowledge-driven method. The concepts selection of FCMs and value range of the

concepts and weights depend on experts. Moreover, the initial structure for MDAAA modeling will be provided by experts. EFKM is proposed to design the membership functions with noisy production data, which is the data-driven method. Next, the fuzzy membership degrees are obtained for the input of the fuzzy decision trees. The fuzzy decision trees are used to extract the fuzzy rules, which are used to amend the initial structure in order to obtain a desired structure. The STA is introduced to detect the causality degree in this study. At last, the MDAAA model is obtained—that is, the augmented FCMs model. In general, knowledge guides the model building, and data is used for amendment.

3. Background of the proposed strategy

3.1. FCMs theory

As a simple intuitive graphical representation and efficient inference mechanism for complex systems, FCMs are a combination of fuzzy logic and neural networks, and have widespread applications [19]. FCMs can be described by a set of concepts (i.e., variables of a system) and relationships (i.e., causality between variables) [20]. FCMs are usually constructed by domain experts who have intimate knowledge of systems. This experiential knowledge is transformed into concepts collection and relation strength in some way. Fig. 4 illustrates a simple example of FCMs where C_i is a concept with the numerical value A_i. The value A_i refers to the active degree of a concept, which varies in a normalized range of [0, 1] or [−1, 1] [19,36]. The strength w_{ij} denotes the causality degree between the cause variable C_i and the effect variable C_j, which ranges from −1 to 1 or a trivalent logic within {−1, 0, 1}. However, the trivalent logic is not fit for describing and
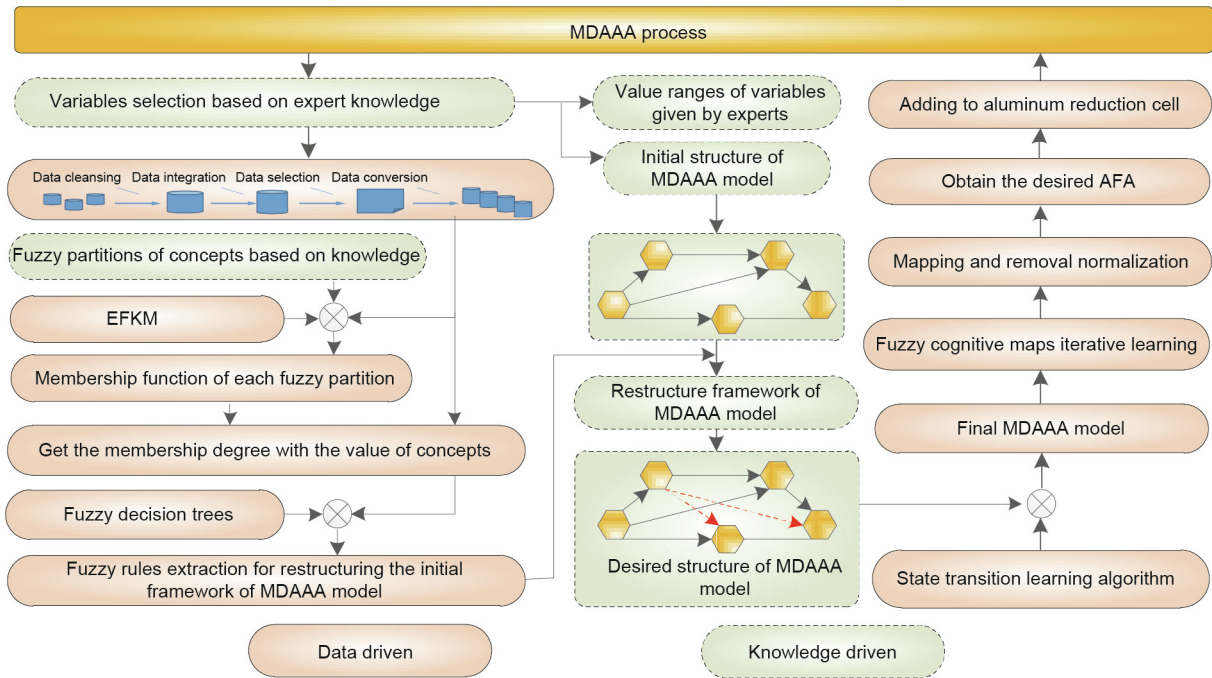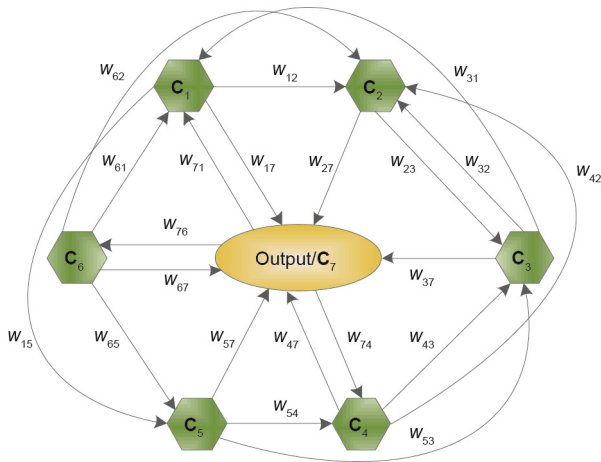
Fig. 3. Solving the MDAAA.

Fig. 4. A simple example of FCMs.

modeling the real world [37]. The adjacency matrix of FCMs corresponding to Fig. 4 is defined as **W**.

$$\mathbf{W} = \begin{bmatrix} 0 & w_{12} & 0 & 0 & w_{12} & 0 & w_{17} \\ 0 & 0 & w_{23} & 0 & 0 & 0 & w_{27} \\ w_{31} & w_{32} & 0 & 0 & 0 & 0 & w_{37} \\ 0 & w_{42} & w_{43} & 0 & 0 & 0 & w_{47} \\ 0 & 0 & w_{53} & w_{54} & 0 & 0 & w_{57} \\ w_{61} & w_{62} & 0 & 0 & w_{65} & 0 & w_{67} \\ w_{71} & 0 & 0 & w_{74} & 0 & w_{76} & 0 \end{bmatrix} \quad (1)$$

Given the initial state $\mathbf{A}(t) = [A_1(t); A_2(t); \dots; A_n(t)]$ (where n is the number of concepts) and adjacency matrix **W**, then the new state of each concept is calculated iteratively in the following way.

$$A_i(t+1) = f\left(A_i(t) + \sum_{j=1, j\neq i}^{n} A_j(t)w_{ji}\right) \quad (2)$$

where $A_i(t+1)$ is the value of concept $\mathbf{C}_i$ at time $(t+1)$; $A_i(t)$ is the value of concept $\mathbf{C}_i$ at time t; w_{ji} is an element of **W**; and f is a threshold function and makes the results map to an interval [0, 1] or [−1, 1] [35].

3.2. Principle of the STA

Learning algorithms play an important role in improving the accuracy of computing results based on FCMs [29]. Thus far, Hebbian-based, evolutionary-based, and hybrid-based algorithms have been applied in existing contributions [38,39]. In those works, the researchers mainly focused on obtaining a desired adjacency matrix **W** using experiential knowledge and/or historical data.

In this study, the STA, which is adopted to deal with non-convex optimization problems, is a global optimization algorithm [40–43]. Each solution is considered as a state of the problem for the STA, and the updating of the current solution is regarded as a state transition. The generation of the candidate solutions in the STA can be described as follows:

$$\begin{cases} \mathbf{X}_{t+1} = A_t\mathbf{X}_t + B_t\mathbf{u}_t \\ y_{t+1} = f(\mathbf{X}_{t+1}) \end{cases} \quad (3)$$

where $\mathbf{X}_t \in \mathbb{R}^n$ stands for a candidate solution; A_t and B_t represent transformation operators; $\mathbf{u}_t$ is a function of **X** and the historical state; and $f(\cdot)$ is the objective function.

The candidate solutions are generated based on the following four special state transformation operators.

(1) Rotation transformation

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \varepsilon_a \frac{1}{w\| \mathbf{X}_t \|_2} \mathbf{R}_r\mathbf{X}_t \quad (4)$$

where ε_a is the rotation factor; $\mathbf{R}_r \in \mathbb{R}^{w \times w}$ is a random matrix, and each element of $\mathbf{R}_r$ ranges from −1 to 1; and $\| \cdot \|_2$ is the 2-norm of a vector. With a given radius a, the candidate solution can be generated by the rotation transformation in a domain of a hypersphere; the rotation transformation is a local search operator.

(2) Translation transformation

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \varepsilon_b \mathbf{R}_t \frac{\mathbf{X}_t - \mathbf{X}_{t-1}}{\| \mathbf{X}_t - \mathbf{X}_{t-1} \|_2} \qquad (5)$$

where ε_b is the translation factor, and each element of $\mathbf{R}_t \in \mathbb{R}^{w \times w}$ is a random variable ranging from 0 to 1. The translation transformation can generate a line search, which is only performed when a better solution can be found by other transformation operators.

(3) Expansion transformation

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \varepsilon_c \mathbf{R}_e \mathbf{X}_t \qquad (6)$$

where ε_c is the expansion factor, $\mathbf{R}_e \in \mathbb{R}^{w \times w}$ is a random diagonal matrix, and the elements of ε_c obey the Gaussian distribution such that the mean value and standard deviation are 0 and 1, respectively. The expansion transformation is a global search operator that is able to search in the whole space.

(4) Axesion transformation

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \varepsilon_d \mathbf{R}_d \mathbf{X}_t \qquad (7)$$

where ε_d is the axesion factor, $\mathbf{R}_d \in \mathbb{R}^{w \times w}$ is a random diagonal matrix, and the elements of ε_d obey the Gaussian distribution such that the mean value and standard deviation are 0 and 1, respectively, and only one random position has a nonzero value. The axesion transformation is designed for a single-dimensional search as well as the global search operator.

For a given solution, many different candidate solutions can be generated by the aforementioned state transition operators. In this study, the number of candidates generated by a certain operator is set to 30, and in every iteration, the transformation operators are alternately and independently applied. The FCMs can learn with the objective function based on the STA to find the optimal $\mathbf{W}$.

3.3. Knowledge extraction and membership functions design

3.3.1. Knowledge extraction based on fuzzy decision trees

In this study, as a representation of knowledge, fuzzy rules are used to amend the initial structure of the MDAAA model. Many computational intelligence knowledge-extraction algorithms are available from the historical database, including neural networks [44], support vector machines [45], and so on. Although these methods are helpful for building a knowledge base, they always suffer from inadequately or improperly expressing and dealing with the vagueness and ambiguity associated with experts' thinking, reasoning, cognition, and consciousness [46]. For aluminum reduction, fuzzy variables, being from cognition, are always used to describe the cell information [1]. In addition, most technicians do not have enough specific knowledge to understand the results generated by the above algorithms. Therefore, it is expected to use the algorithms whose results are comprehensible and evaluative, rather than using black box approaches. Simple linguistically explicable rules are a good fit for technicians. A fuzzy decision tree induction method is thus introduced to extract knowledge.

The fuzzy decision tree induction method can explicitly express, measure, and incorporate the cognitive uncertainties of technicians into the knowledge induction process. In the induction process, this method can reduce the ambiguity of the classification with fuzzy evidence to construct fuzzy decision trees [46]. Accordingly, the fuzzy decision tree induction method is suitable for dealing with the problem of cognitive uncertainties in the aluminum reduction process [46].

3.3.2. Membership functions design based on EFKM

The membership degrees used to extract fuzzy rules are usually provided by experts. However, the subjectivity of the experts may affect the results. Therefore, a memberships generating method is proposed. Because of the severe environment of the cell, the measured data contain a great deal of noise, and isolated points even appear. EFKM is proposed in order to deal with the problem of the traditional fuzzy k-means (FKM) being sensitive to isolated points [47]. Based on EFKM, we obtain k clustering centers, c_1, $c_2, \ldots, c_k$, where k is equal to the number of fuzzy partitions in practice, which is decided by the technicians.

The distance between the sample points x_i and the cluster centers c_k is defined as follows:

$$d_{ik} = \alpha_k \frac{|x_i - c_k|}{|x_i| + |c_k|} + \frac{1 - \alpha_k}{N} \sum_{x_j \in U} \frac{|x_j - c_k|}{|x_j| + |c_k|} \qquad (8)$$

where x_i and x_j are the ith and jth samples, respectively; c_k is the kth clustering center; N is the number of adjacent sample points, decided by the clustering centers numbers; U represents a collection of neighborhood sample points; and α_k is defined as the smoothing parameter, which is self-adjustable based on spatial neighborhood information, where the more homogeneous the sample space is, the smaller α_k will be. When target sample points are in the regional margin, α_k will be bigger, and the memberships of the sample points will be slightly affected. α_k is shown by Eqs. (9) and (10).

$$\tilde{\alpha}_k = \sqrt{\frac{1}{N} \sum_{x_j \in U} \left[d(x_k, x_j) - \frac{1}{N} \sum_{x_j \in U} d(x_k, x_j) \right]^2} \Bigg/ \sqrt{\frac{1}{N} \sum_{x_j \in U} \left[x_j - \frac{1}{N} \sum_{x_j \in U} \left(x_j - \frac{1}{N} \sum_{x_j \in U} x_j \right) \right]^2} \qquad (9)$$

$$\alpha_k = \tilde{\alpha}_k / \max\left(\tilde{\alpha}_k \right) \qquad (10)$$

The membership of sample x_i belonging to category k is represented by μ_{ik}.

$$\mu_{ik} = \sum_{j=1}^{k} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{-2}{m-1}} \qquad (11)$$

where m is the weighted index number, and the default value of m is 2.

The clustering centers are shown by the following equation.

$$c_k = \sum_{k=1}^{n} \left(\mu_{ik} \right)^m x_k \Bigg/ \sum_{k=1}^{n} \left(\mu_{ik} \right) \qquad (12)$$

Then, a method of membership functions constructed with the clustering centers for each fuzzy division of every concept is proposed, as illustrated by Eqs. (13)–(15).

When $i = 1$, the membership functions μ_{e_1} for the first fuzzy division are given by the following:

$$\mu_{e_1} = \begin{cases} 1 - (x - c_1)/(x_{\min} - c_1) * 0.5, & \text{if } x_{\min} \leq x < c_1 \\ 1 - (x - c_1)/(c_2 - c_1), & \text{if } c_1 \leq x < c_2 \\ 0. & \text{otherwise} \end{cases} \qquad (13)$$

When $1 < i < k$, the membership functions μ_{e_i} for the ith fuzzy division are given by the following:

$$\mu_{e_i} = \begin{cases} 1 - (x - c_1)/(c_{i-1} - c_i), & \text{if } c_{i-1} \leq x < c_i \\ 1 - (x - c_1)/(c_2 - c_1), & \text{if } c_i \leq x < c_{i+1} \\ 0. & \text{otherwise} \end{cases} \qquad (14)$$

When $i = k$, the membership functions μ_{e_k} for the kth fuzzy division are illustrated by the following:

$$\mu_{e_k} = \begin{cases} 1 - (x - c_k)/(c_{k-1} - c_k), & \text{if } c_{i-1} \leq x < c_i \\ 1 - (x - c_k)/(x_{\max} - c_k) * 0.5, & \text{if } c_k \leq x < x_{\max} \\ 0. & \text{otherwise} \end{cases} \qquad (15)$$

where μ_{e_i} denotes the membership degree of sample x attaching to the ith fuzzy division, c_i is the ith clustering center, and $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the universe, respectively.

In this section, we obtain k clustering centers based on EFKM and the membership functions of the fuzzy divisions based on Eqs. (13)–(15). Based on these methods, we can obtain the membership degrees of the production data that are used to extract the fuzzy rules.

4. New decision models constructed for the AFA

The AFA is usually computed by means of a simple formula or decided by technicians. However, due to the cell's complexities, it is difficult for formulas to involve the characteristics, and technicians' experiential knowledge is not always correct. Therefore, it guarantees the accuracy of the AFA based on data-driven or knowledge-driven methods alone. In this section, a new MDAAA model is constructed based on data and knowledge collaboration, combing with FCMs, knowledge acquisition techniques and data processing methods. The main steps of MDAAA modeling are as follows:

Step 1: The concepts and value ranges and the causal directions between concepts are determined by an expert. This makes it possible to obtain the initial structure of the MDAAA. This step is knowledge-driven.

Step 2: The production data is preprocessed, including selection and conversion. Next, the clustering centers based on EFKM are obtained, which are used to generate the membership function of each fuzzy division. The membership degrees are used to extract the fuzzy rules. This step is data-driven.

Step 3: The fuzzy rules are used to amend the initial structure and to obtain the desired structure of MDAAA. This step is knowledge-driven.

Step 4: The STA is introduced to detect the causality degree (weights). The final MDAAA model is obtained, and the desired AFA is obtained. This step is data-driven.

Step 5: The target AFA is added to the cell by means of feeding devices.

Based on the above steps, the details of the MDAAA model building process are as follows.

4.1. Initial structure design of the MDAAA

Thirteen concepts are selected by an expert. Reducing the number of concepts to 13 makes it possible to not only cover the characteristics of MDAAA, but also reduce the computing complexity. Table 1 lists the 13 concepts with their descriptions and numbers of fuzzy division.

Definition 1. One computing cycle is defined as the difference between today and yesterday:

Table 1
Description and fuzzy partitions of concepts.

Concepts	Descriptions	Abbreviations	Number of fuzzy divisions
C_1	Molecular ratio of electrolyte	MR	5
C_2	Electrolyte temperature	ET	5
C_3	Mean voltage	MV	3
C_4	Aluminum level	AL	3
C_5	Electrolyte level	EL	3
C_6	Cumulative changes of C_1	$\sum \Delta MR$	3
C_7	Cumulative changes of C_2	$\sum \Delta ET$	3
C_8	Cumulative changes of C_3	$\sum \Delta MV$	3
C_9	Cumulative changes of C_4	$\sum \Delta AL$	3
C_{10}	Cumulative changes of C_5	$\sum \Delta EL$	3
C_{11}	AlF$_3$ feeding amount	AFA	5
C_{12}	Setting voltage	SV	3
C_{13}	Tapping amount	TA	5

$$\sum \Delta MR(t - i) = MR(t - i) - MR(t - i - 1) \tag{16}$$

where MR is the molecular ratio of electrolyte, $MR(t)$ represents the current value of the molecular ratio, and $MR(t - i)$ denotes the value on the $(t - i)$th day. The definitions of the other concepts are the same as those for the MR.

Definition 2. Due to the particularity of aluminum reduction, the change in MR is not obvious. It is difficult for fewer computing cycles to reflect the recent changes in the MR. However, there is no need to consider more computing cycles. In practice, the cumulative values of four computing cycles for the MR are reasonable, which are defined as follows:

$$\sum \Delta MR = \sum_{t=0}^{3} \Delta MR(t - i) \tag{17}$$

Based on experiential knowledge, the MR is the current value ratio of NaF to AlF$_3$ [4]. The lower the MR, the lower the ET [1]. The thickness of the SL changes with variation in the ET; furthermore, changes in the MR may be caused by electrolyte solidification or SL melting in the past few days, where the higher the ET, the higher the MR [1]. Thus, there are connections between nodes C_1 and C_2. In addition, the electrolyte resistance varies with changes in the MR, where the higher the mean voltage (MV), the higher the ET [1]. Therefore the MR, ET, and MV should be considered for the voltage setting and the AFA, and connections exist from the MR, ET, and MV to the AFA. These cumulative changes indicate the variation trend of the concepts; clearly, there are connections existing from $\sum \Delta MR$, $\sum \Delta ET$, and $\sum \Delta MV$ to AFA. The heat dissipation of the aluminum level (AL) and the thermal insulation of the electrolyte level (EL) are usually used to adjust the cell energy balance. The higher the AL, the lower the ET, and the higher the EL, the higher the ET. The MV is used to adjust the energy import. Connections from the MR, ET, MV, and AL to the setting voltage (SV) should exist, due to influences on the energy variation. Accordingly, these should be considered in the MDAAA modeling, along with the cumulative changes. In practice, the tapping amount (TA), AFA, and SV are used to adjust the energy balance, and influence each other.

Based on the above analysis, the initial structure of MDAAA is obtained, as illustrated in Fig. 5. The initial structure will be restructured using the methods proposed in Sections 3.3.1 and 3.3.2. However, the casual relationships are still unknown. The STA is introduced to detect weights in the following section.

4.2. FCMs learning with the STA

Experts are subjective in determining the casual relationships for simple FCMs. However, as shown in Fig. 5, it is difficult for an expert to build complicated FCMs. Accordingly, there is an urgent need to develop a learning algorithm. In this study, the STA is introduced as a learning algorithm to eliminate the need for expert intervention. Although the STA has been widely used in many fields [40–43], this is the first time it is used for the FCMs learning process. When the objective function reaches a minimum value, a set of the desired weights of the augmented FCMs is obtained, and the concepts' values are in the predefined interval. The main goal is to detect an adjacency matrix that brings the FCMs to a steady state in the predefined interval of each weight.

The objective function must quantitatively measure the suitability of a given candidate solution; it iteratively calculates the difference between the estimated values and the real values of the concepts. In our proposal, the objective function with the predefined interval of the concepts and weights is shown in Eq. (18). In the aluminum reduction process, the values of the concepts vary within a certain range. Therefore, the values of the concepts are
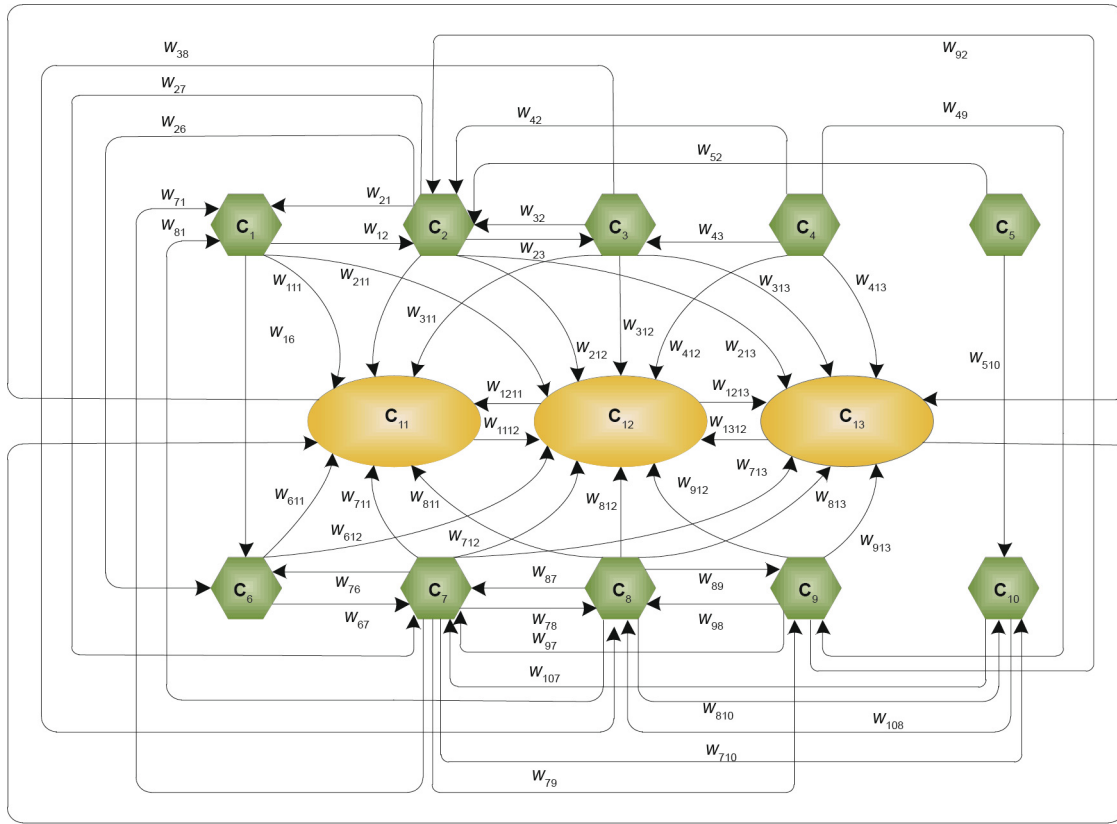
Fig. 5. The initial structure of MDAAA.

restricted within an interval, $A_{l,\text{lb}} \leq A_l \leq A_{l,\text{ub}}$. In practice, the causality degree is either positive or negative. For example, the higher the AL, the greater the energy loss. The causality degree between the AL and the ET is negative. However, the higher the EL, the lower the energy loss. The causality degree between the ET and the EL is positive. Accordingly, based on the real-life situation, the values of w_{ijn} must be restricted within an interval, $w_{ijn,\text{lb}} \leq w_{ijn} \leq w_{ijn,\text{ub}}$. In this study, w_{ijn} is restricted in $[-1, 0]$ or $[0, 1]$.

$$\min \quad J = \frac{1}{K-1} \cdot \frac{1}{M} \sum_{k=1}^{K-1} \sum_{m=1}^{M-1} A_i(t) - \hat{A}_i(t)$$

$$\text{s.t.} \quad A_{l,\text{lb}} \leq A_l \leq A_{l,\text{ub}} \quad l = 1, 2, \ldots, M \tag{18}$$

$$w_{ijn,\text{lb}} \leq w_{ijn} \leq w_{ijn,\text{ub}} \quad n = 1, 2, \ldots, N$$

where K and M are the number of weights and concepts, respectively; K is the number of learning records (historical data); $A_{l,\text{lb}}$ and $A_{l,\text{ub}}$ are the lower bound and upper bound of the concepts, respectively; $A_i(t)$ is calculated by Eq. (2); $\hat{A}_i(t)$ is the desired response; and $w_{ijn,\text{lb}}$ and $w_{ijn,\text{ub}}$ are lower bound and upper bound of each weight.

5. Results and discussion

In order to verify the feasibility and ability of the proposed strategy, the results of an experimental study of AlF$_3$ addition in the aluminum reduction process are discussed in this section. The problem is investigated in a certain aluminum reduction plant in Binzhou City, Shandong Province, China. The industrial data for June and July 2016 were collected. In the aluminum reduction process, the AL, MV, EL, and ET are measured every day. The MR is measured every two days, with the averages of the closest two days used for the missing values. The corresponding cumulative changes of the above variables are accessed every day. The MR, ET, MV, AL, and EL and the corresponding cumulative changes are shown in Figs. 6 and 7, respectively.

5.1. The initial structure of the MDAAA model amendment

In order to alleviate the need for expert intervention for MDAAA modeling, a knowledge-extraction method combined with EFKM and the membership function generating method is proposed.

First, clustering centers are generated by EFKM. The statistical results are shown in Table 2, including the minimum value, maximum value, and clustering centers. The number of clustering centers being equal to the number of fuzzy partitions—that is, the value of k—is based on experiential knowledge. The membership functions of the MR, $\sum \Delta$MR, TA, SV, and AFA are shown in Fig. 8. Other membership functions of the concepts are similar to those shown in Fig. 8, and are omitted here to save space.

Second, as a result of the complexity of the MDAAA model, it is difficult for experts to depict the causalities. The fuzzy rule is an expression of the association between concepts, and is also a means of tacit knowledge representation, as illustrated by the following:

- If $\mathbf{C}_i$ is high and $\mathbf{C}_j$ is low, then $\mathbf{C}_k$ is normal $\iff \mathbf{C}_i \xrightarrow{w_{ij}} \mathbf{C}_j \xrightarrow{w_{jk}} \mathbf{C}_k$

Accordingly, the connections can be mined based on knowledge-extraction methods. Fuzzy decision trees are implemented to extract fuzzy rules, whose variables include the MR, ET, MV, AL, EL, $\sum \Delta$MR, $\sum \Delta$ET, $\sum \Delta$MV, $\sum \Delta$EL, $\sum \Delta$AL, AFA, SV,
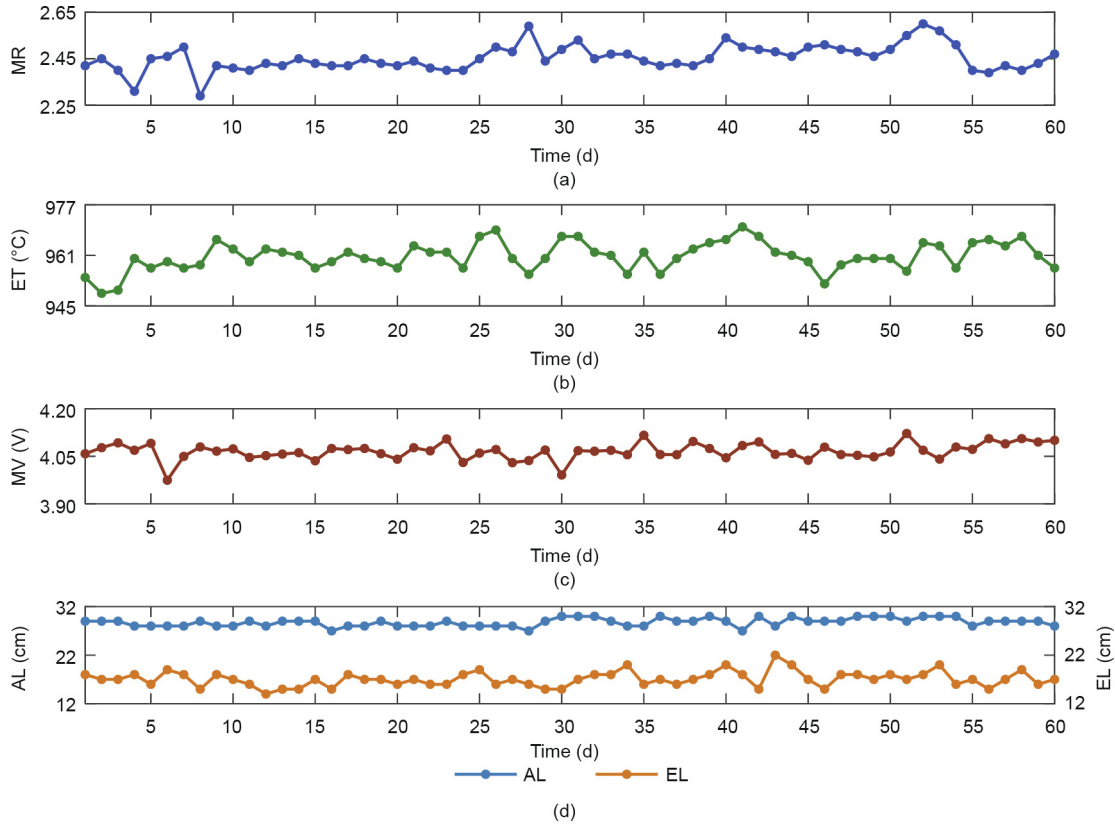
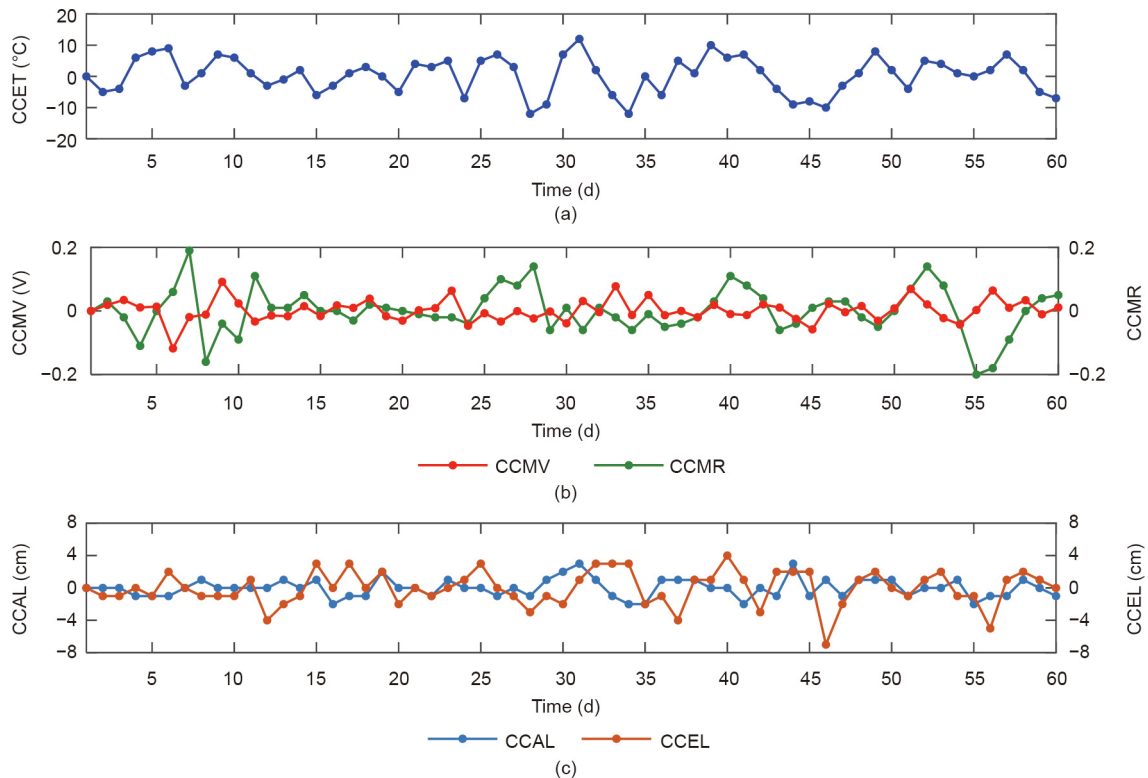**Fig. 6.** The values of (a) MR, (b) ET, (c) MV, and (d) AL and EL.

Fig. 7. The cumulative changes in (a) ET, (b) MV and MR, and (c) AL and EL. CCET: cumulative changes of ET; CCMV: cumulative changes ofMV; CCMR: cumulative changes of MR; CCAL: cumulative changes of AL; CCEL: cumulative changes of EL.

Table 2
Clustering centers c_i based on the EFKM algorithm.

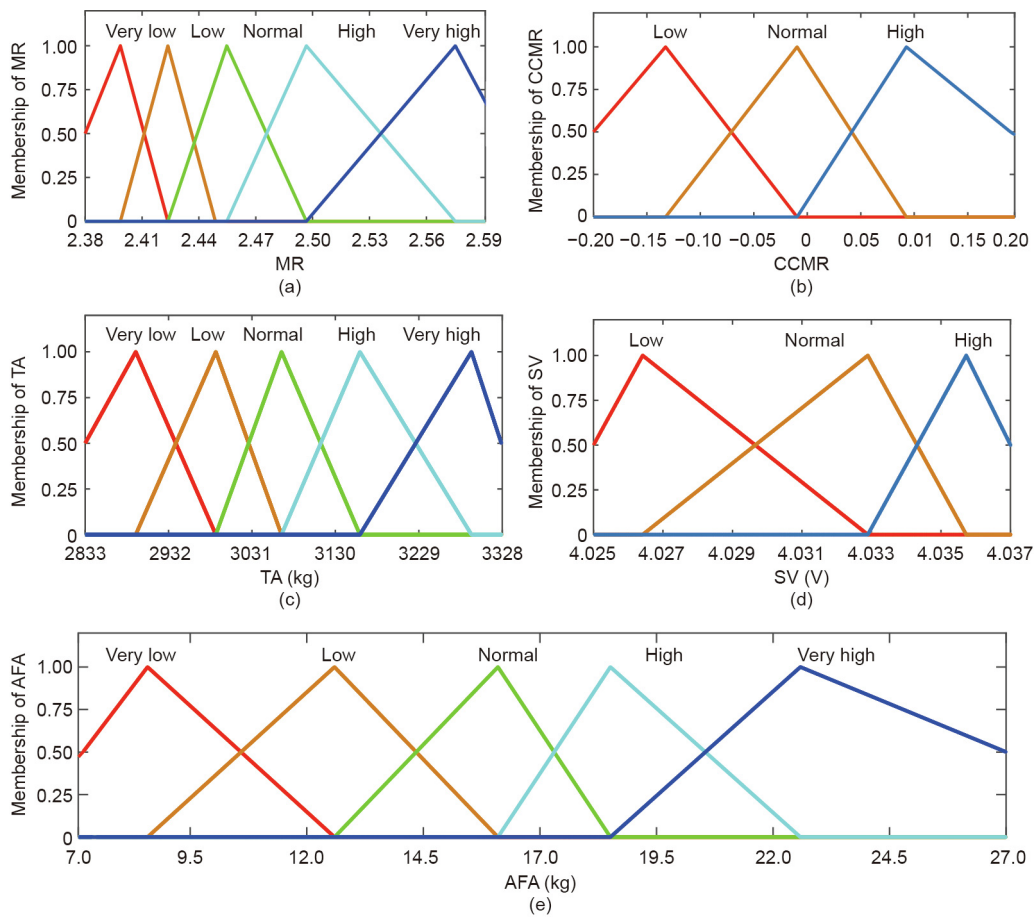| Concepts	Names	Minimum	Maximum	Clustering centers c_i
C_1	MR	2.36	2.60	[2.3985, 2.4236, 2.4547, 2.4967, 2.5752]
C_2	ET (°C)	949	970	[950.33, 956.51, 959.79, 962.59, 966.81]
C_3	MV (V)	3.9746	4.1219	[4.0357, 4.0665, 4.0994]
C_4	AL (cm)	27	30	[27.9219, 28.9931, 29.9964]
C_5	EL (cm)	14	22	[15.4547, 17.4399, 19.9062]
C_6	$\sum \Delta$MR	−0.20	0.19	[−0.132527, −0.009562, 0.092622]
C_7	$\sum \Delta$ET (°C)	−12	12	[−6.7340332, 1.151452, 6.927109]
C_8	$\sum \Delta$MV (V)	−0.118	0.092	[−0.0274, 0.0104, 0.0650]
C_9	$\sum \Delta$AL (cm)	−2	3	[−1.1714, 0.0113, 1.1817]
C_{10}	$\sum \Delta$EL (cm)	−7	4	[−4.1944, −0.7519, 2.0486]
C_{11}	AFA (kg)	7.2	27	[8.585, 12.595, 16.101, 18.512, 22.587]
C_{12}	SV (V)	4.025	4.037	[4.0264, 4.0329, 4.0357]
C_{13}	TA (kg)	2833	3327	[2893.1, 2988.1, 3066.1, 3159.2, 3291.5]

Fig. 8. Fuzzy membership functions generated based on EFKM and Eqs. (13)–(15) for: (a) MR, (b) CCMR, (c) TA, (d) SV, and (e) AFA.

and TA. The membership degrees of these variables, being the input of the fuzzy decision trees, are generated based on Eqs. (13)–(15). The pseudocodes of the fuzzy decision trees induction method are shown in Algorithms 1–6, where G and $GG_{ambiguity}$ are the classification ambiguity and ambiguity function, respectively. D_{Ai} and D_{Ci} are the ith dataset of attributes D_A and classification D_C, respectively. S and A_{ij} are the fuzzy subsethood and the jth partition of ith attribute, respectively. In addition, D_{Aij}, D_{Cmk} are the data of the jth partition of ith attribute and the data of

the jth partition of ith classification, respectively. ms and G_CE are the number of partition and the ambiguity of each parent node, respectively. AMBIGUITY is the ambiguity function. C_mu and $G\_A_i\_A_{jm}$ are the membership of classification and the class ambiguity for each fuzzy partition of each parent node and child node, respectively. ClassAmbiguityWithP is the class ambiguity function with parent node. The muE is the membership of fuzzy evidence. muF is the membership of each fuzzy partition. muC is the membership of each classification.

Algorithm 1. Main function.

Input: Dataset of attributes and classification D_A, D_C, parameters α, β
Output: Fuzzy value
1: **for** each data D_A **do**
2: **if** $D_A(i, j) < \alpha$ **then**
3:   $D_A(i, j) = 0$;
4: **end if**
5: **end for**
6: **for** each attribute A_i **do**
7: $G\_A_i = GG_{ambiguity}(D_{Ai}; D_{Ci})$;
8: value$\_A_i = \min(G\_A_i)$;
9: fprintf('*The rootnode is A_i*');
10: **for** each fuzzy partition of root node **do**
11: $S\_A_{ij}\_C(i) = \text{subsethood}A\_B(D_{Aij}, D_{Cmk}, ms)$;
12:   **if** $\max(S\_A_{ij}\_C) \geq \beta$ **then**
13:    fprintf('The leaf node of each fuzzy partition of root node is');
14:   **end if**
15: **end for**
16: **end for**
17: **for** each parent node **do**
18: $G\_CE = \text{AMBIGUITY}(FuzzyEvidence\_C\_mu, D_{Ai}, D_{Ci})$;
19:   **for** each fuzzy partition of each parent node and child node **do**
20:    $G\_A_i\_A_{jm} = \text{ClassAmbiguityWithP}(muE, muF, muC)$;
21:   **end for**
22:    $[childnode\_A_{jm}\_\min\ index\_\min] = \min(G\_A_i\_A_{jm})$;
23: **if** $childnode\_A_{jm}\_\min < G\_CE\_A_{jm}$ **then**
24:    fprintf('*The child node of A_{jm} is %s*');
25: **end if**
26: **end for**

Algorithm 2. Classification ambiguity function.

Input: Fuzzy memberships fuzzy events D and C
Output: The classification ambiguity
1: **function** $G\_A_i = GG_{ambiguity}(D_A, D_C)$
2: **for** dataset $D_A D_C$ **do**
3:   $D_{Ai}\_column = \text{sum}(D_{Ai})$
4:   $subsethood\_A_i(i, :) = \text{subsethood}A\_B(D_{Ai}(:, i), D_{Ci})$;
5: **end for**
6: $weights\_A_i = D_{Ai}\_column/sum\_A_i$;
7: $G\_A_i = weights\_A_i * (ambiguity(subsethood\_A_i(i, :) = \max(subsethood\_A_i); ns_i, ns_j))$;
8: **end function**

Algorithm 3. Classifying the possibility function.

Input: Fuzzy memberships of evidence, categories and numbers of categories
Output: Possibility of Classifying an object to class C_i
1: **function** FuzzyEvidence_CE = FuzzyEvidence_C_mu(muE, muC, C)
2: **for** each D_A **do**
3:   subsethood_temp(i) = subsethoodA_B(muE, muC(:, i), size(muC, 1));
4: **end for**
5: FuzzyEvidence_CE = subsethood_temp/max(subsethood_temp);
6: **end function**

Algorithm 4. Subsethood function.

Input: Fuzzy memberships of A, B, and number of instances, N
Output: The subsethood between A and B
1: **function** subsethood_AB = subsethoodA_B(muA, muB, N)
2: **for** each D_A **do**
3:   $val(:, j) = \min(muA, muB(:, j))$;
4: **end for**
5: subsethood_AB = sum(val)/sum(muA_{ij});
6: **end function**

Algorithm 5. Ambiguity function.

Input: Fuzzy memberships of A, number of categories of A and B
Output: Ambiguity of B
1: **function** E_A = ambiguity(muA, nA, nB)
2: The sorting result $p_i\_A_i = \text{sort}(muA, 2, 'descend')$;
3: $p_i\_A_i (:, nB + 1) = 0$;
4: **for** each sort results **do**
5:   $E\_a\_A_j(:, i) = (p_i\_A_i (:, i) - p_i\_A_i(:, i + 1)) * \log(i)$;
6: **end for**
7: E_A = sum($E\_a\_A_j$; 2);
8: **end function**

Algorithm 6. Classification ambiguity with fuzzy partitioning.

Input: Fuzzy memberships of E, P, numbers of partitioning and instances
Output: Classification ambiguity with fuzzy partitioning
1: **function** G_P_F = ClassAmbiguityWithP(muE, muF, muC, K, N)
2: **for** each D_A **do**
3:   $M\_E_i F(:, k) = \min(muE(:, k) , muF)$;
4:   $w\_E\_F(k) = \text{sum}(M\_E_i\_F) = \text{sum}(M\_E_i\_F\_sum, 2)$;
5:   $G\_E_i\_F(k) = \text{subsethood}A\_B(M\_E_i\_F(:, k), muC, \text{size}(muC, 2))$;
6: **end for**
7: G_P_F = $w\_E\_F * G\_E_i\_F$;
8: **end function**

Due to the complexity of the MDAAA and the limitations of technician cognition, some rules are ignored. The specificity to the aluminum reduction process, C_i, not only affects C_j; C_j also influences C_i. Because of the large number of fuzzy rules produced by the fuzzy decision trees, we present only the specific rules that differ from the rules suggested by the experts. These rules are used to amend the MDAAA structure, presented as follows:

- If C_1 is high and C_2 is high and C_4 is normal and C_7 is low then C_{11} is very high;
- If C_1 is very high and C_2 normal and C_3 is high and C_4 is low and C_9 is high then C_{11} is very high;
- If C_1 is high and C_2 is very low and C_5 is high and C_{10} is low then C_{13} is very high;
- If C_3 is low and C_1 is normal and C_6 is high and C_7 is high and C_8 is high and C_9 is high then C_{11} is very low;
- If C_3 is low and C_5 is high then C_{13} is high;
- If C_4 is high and C_1 is normal and C_6 is high and C_7 is high then C_{12} is low;
- If C_5 is normal and C_1 is normal and C_2 is normal and C_3 is high and C_4 is low then C_{11} is high;
- If C_8 is high and C_{10} is low then C_{13} is low;
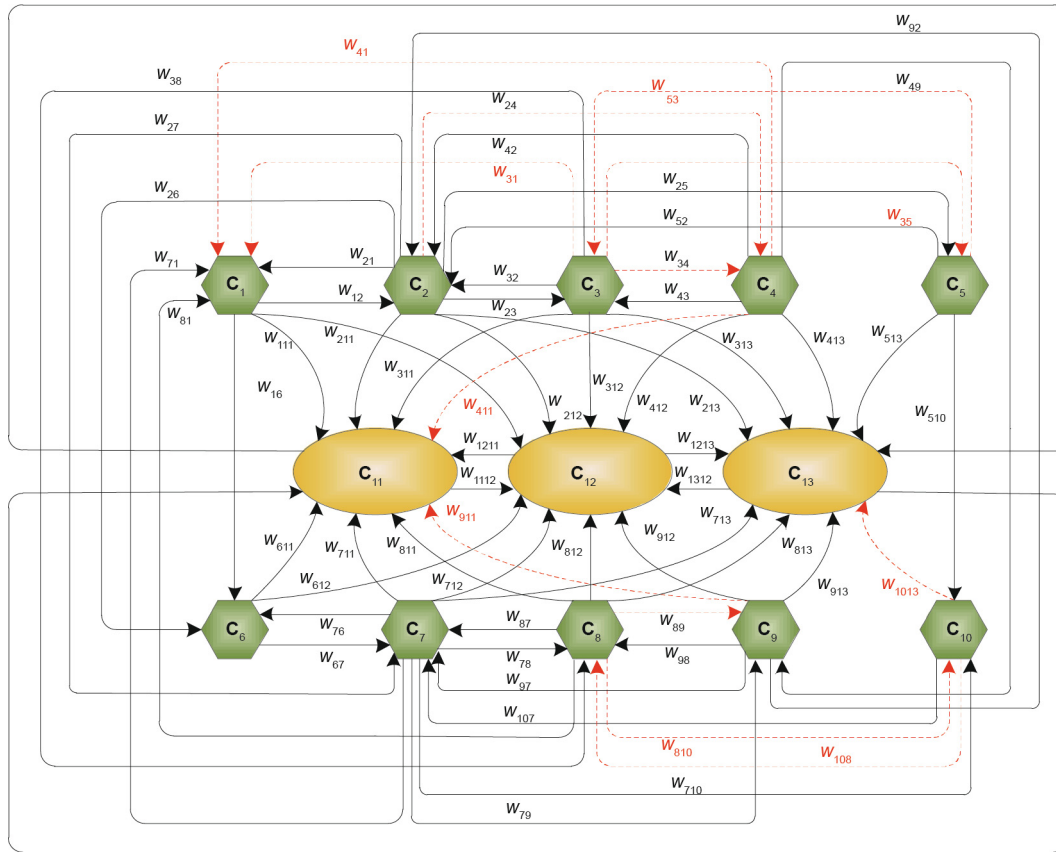- If C_{10} is normal and C_8 is low and C_9 is high then C_{12} is high.

Fig. 9. The reconstructed structure of the MDAAA model.

Third, based on the above fuzzy rules, the potential connections among these concepts are found. The initial structure is reconstructed using the newfound connections among the concepts; the new structure of MDAAA is shown in Fig. 9.

The process of knowledge acquisition is based on the data-driven method. The special rules above are used to reconstruct the initial structure, and obtain an augmented structure. The process of reconstruction involves data and knowledge fusion, and is a data and knowledge collaboration process.

5.2. The MDAAA model obtained with FCMs learning

At present, we obtain the desired structure of the MDAAA model without weights. In the following parts, we will detect the weights to build augmented FCMs based on the STA. The vector $A_{initial}$ represents the initial states of the concepts at a given time in the aluminum reduction process. Because A_i is in $[-1, 1]$, the actual production data should be normalized by the following equation:

$$A = \frac{\tilde{A} - \tilde{A}_{min}}{\tilde{A}_{max} - \tilde{A}_{min}} \times \left(U_{range} - D_{range}\right) + D_{range} \tag{19}$$

where $\tilde{A}$ is the real concept value; $\tilde{A}_{max}$ and $\tilde{A}_{min}$ are the maximal and minimal real data selected from the history database, respectively; and U_{range} and D_{range} are the maximal and minimal normalized data, respectively. In cases where the real concept values are only positive, $U_{range} = 1$ and $D_{range} = 0$, while in cases where the real concept values are able to be both negative and positive, $U_{range} = 1$ and $D_{range} = -1$ in this study.

At last, because we want the truth value of the concepts/variables for the real system, the normalization is removed for A_{final} by the following equation:

$$\tilde{A}_{real} = \frac{A_{final} - D_{range}}{U_{range} - D_{range}} \times \left(\tilde{A}_{max} - \tilde{A}_{min}\right) + \tilde{A}_{min} \tag{20}$$

where A_{final} is the final value after the concept iterations, and $\tilde{A}_{real}$ is the truth value of a variable for a real system.

The $A_{initial}$ is selected as follows:

$$A_{initial} = [\ 0.1618, 0.6190, 0.5588, 0.6667, 0.1250, 0.0769,$$
$$-0.0833, -0.0324, 0.2000, -0.0910, 0.6455, 0.8330, 0.1478]$$

The threshold function is given by the following:

$$f(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

The vector A_{final} is the final state of reachability, produced in the convergence region. The $A_{final,STA}$ is obtained based on the STA.

$$A_{final,STA} = [0.5663, 0.8121, 0.8710, 0.1220, 0.2892, 0.0950,$$
$$-0.0817, 0.2631, 0.6934, -0.1298, 0.6000, 0.7883, 0.5879]$$

The numerical weights among the concepts are shown in Table 3.

Based on the STA, the learning results of the FCMs are shown in Fig. 10. For a comparison with the STA, the differential Hebbian learning (DHL) algorithm is selected, as proposed by Dickerson and Kosko [48]. The results are illustrated by Fig. 11, which present the values of the concepts with 50 iterations. The results indicate that after the 32nd iteration, the FCMs reach an equilibrium region,

Table 3
Numerical weights among the concepts of Fig. 9.

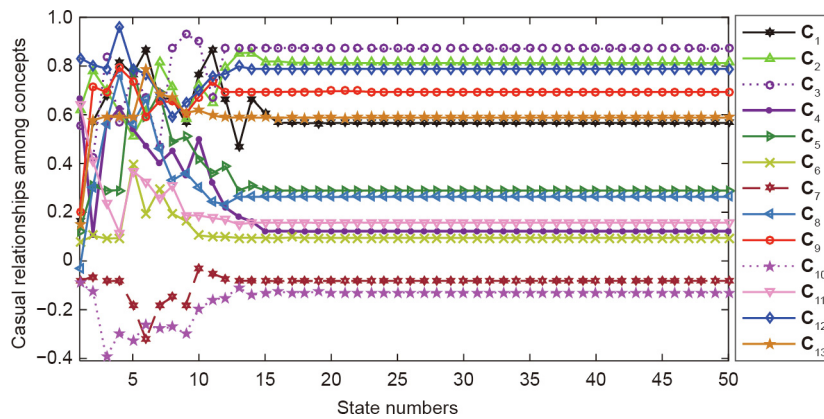| Nodes	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}
C_1	0	0.3852	0	0	0	0.1932	0	0	0	0	0.7515	0.0005	0
C_2	0.1587	0	0.0157	0.0045	0.2789	0.0956	0.3707	0	0	0	0.5664	0.0082	−0.8986
C_3	0.7577	0.2198	0	−0.9733	0.0016	0	0	0.6807	0	0	0.0591	−0.5183	−0.0759
C_4	−0.0883	0.1985	0.9655	0	0	0	0	0	0.9332	0	−0.4101	0.3067	0.4516
C_5	0	0.6874	0.9975	0	0	0	0	0	0	0.0905	0	0	−0.0069
C_6	0	0	0	0	0	0	0	0.11	0	0	0.7233	0.0299	0
C_7	0.0299	0	0	0	0	0.9175	0	0.9771	0.0012	0.8652	0.1644	−0.0444	−0.1181
C_8	0.0018	0	0	0	0	0	0.064	0	−0.8566	−0.8986	0.0938	−0.0085	−0.0683
C_9	0	0.515	0	0	0	0	−0.7632	0.1832	0	0	−0.6158	0.2322	0.8993
C_{10}	0	0	0	0	0	0	0.8998	0.3719	0	0	0	0	−0.9210
C_{11}	0	0	0	0	0	0	0	0	0	0	0	0.2277	−0.0042
C_{12}	0	0	0	0	0	0	0	0	0	0	−0.2555	0	0.0638
C_{13}	0	0	0	0	0	0	0	0	0	0	−0.3092	0.9766	0

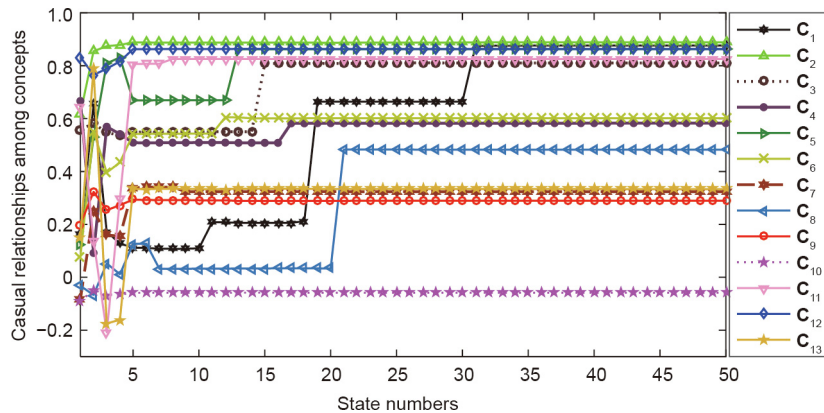**Fig. 10.** FCMs learning results based on the STA.

Fig. 11. FCMs learning results based on the DHL algorithm.

while 16 iterations are required based on the STA. With the above $A_{initial}$, the $A_{final,DHL}$ is as follows:

$$A_{final,DHL} = [0.8322, 0.8784, 0.7993, 0.5847, 0.8567, 0.6000, \\ 0.3547, 0.4437, 0.3046, −0.0673, 0.8157, 0.8555, 0.3689]$$

Fig. 12 indicates that after the 37th iteration, the FCMs converge into a stable region based on particle swarm optimization (PSO). With the same $A_{initial}$, the final calculated values of the concepts $A_{final,PSO}$ based on PSO are as follows:

$$A_{final,PSO} = [0.3481, 0.9460, 0.8624, 0.0477, 0.6316, 0.6609, \\ 0.4790, 0.5618, 0.1174, −0.1409, 0.4041, 0.1020, 0.3381]$$

Fig. 13 presents 100 iterations; after the 74th iteration, the FCMs converge into a stable region with a slow convergence speed based on a genetic algorithm (GA). With the above $A_{initial}$, the final calculated values $A_{fianl,GA}$ of the concepts based on GA are as follows:

$$A_{final,GA} = [0.8537, 0.75, 0.934, 0.5828, 0.7644, 0.5460, 0.5237, \\ 0.5670, 0.4524, 0.0295, 0.4049, 0.0818, 0.4187]$$

Figs. 10–13 reveal that the convergence speed based on the STA is faster than those based on the DHL, PSO, and GA. In order to validate the proposed strategy for the AFA, the values of C_{11} that were computed based on the STA, DHL, PSO, and GA are converted to real values using Eq. (20).
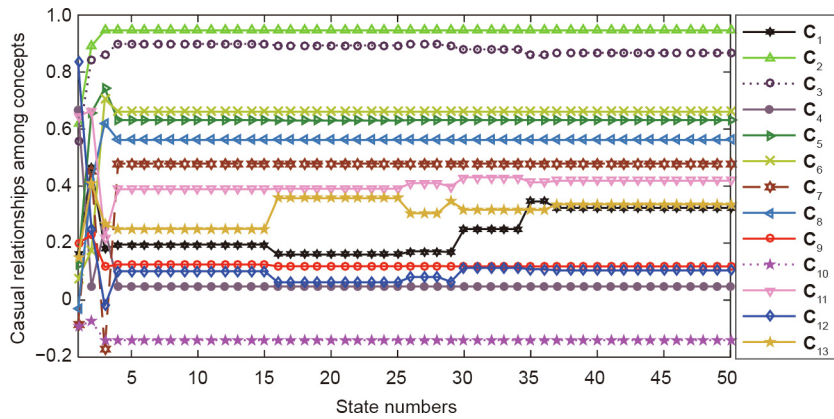
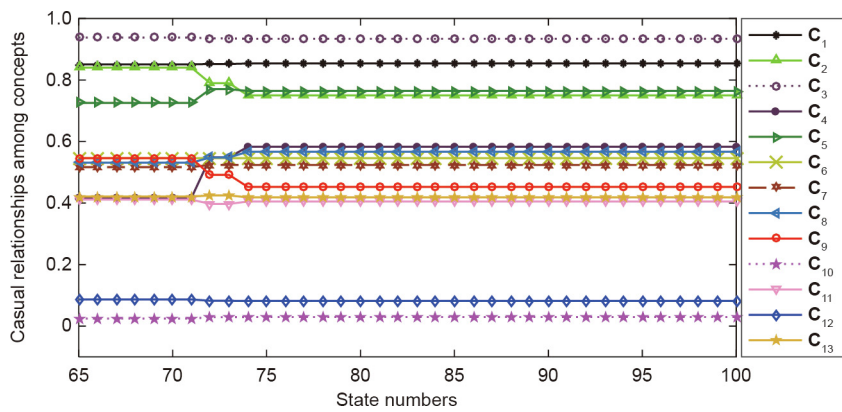**Fig. 12.** FCMs learning results based on the PSO.

Fig. 13. FCMs learning results based on a GA.

First, the lower vibration (VI), waving (WA), and MV are symbols of better operations of AlF$_3$ addition than may be obtained by consulting technicians or experts. Corresponding to lower VI, WA, and MV, the truth values of C_1–C_{13} are selected for validation. Verification in this study was carried out over about two months. The data for the VI, WA, and MV are shown in Fig. 14.

Second, the MDAAA model is used to compute the AFA based on the STA, DHL, PSO, and GA. Existing strategies for the AFA are used for comparison; these are the linear programming model [12] and the fuzzy control method [49]. In practice, as a result of the inherent feeding pattern of the cell, the total AFA is divided into batches, with 1.6 kg being added once to the cell by one feeder. The total feeding times of all AlF$_3$ feeders based on the amended and initial structure are shown in Figs. 15 and 16, respectively. In addition, the feeding times are shown in Fig. 17 based on the initial structure using the FKM. Table 4 provides the analysis results of the MDAAA based on the amended structure, and of the initial structure with EFKM based on the initial
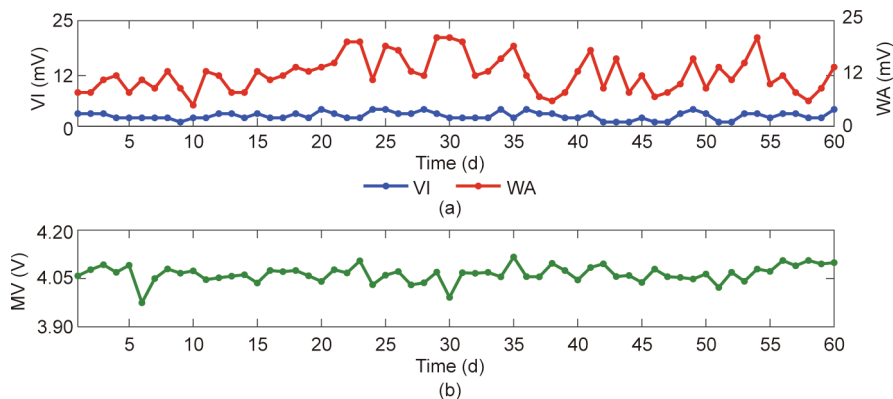

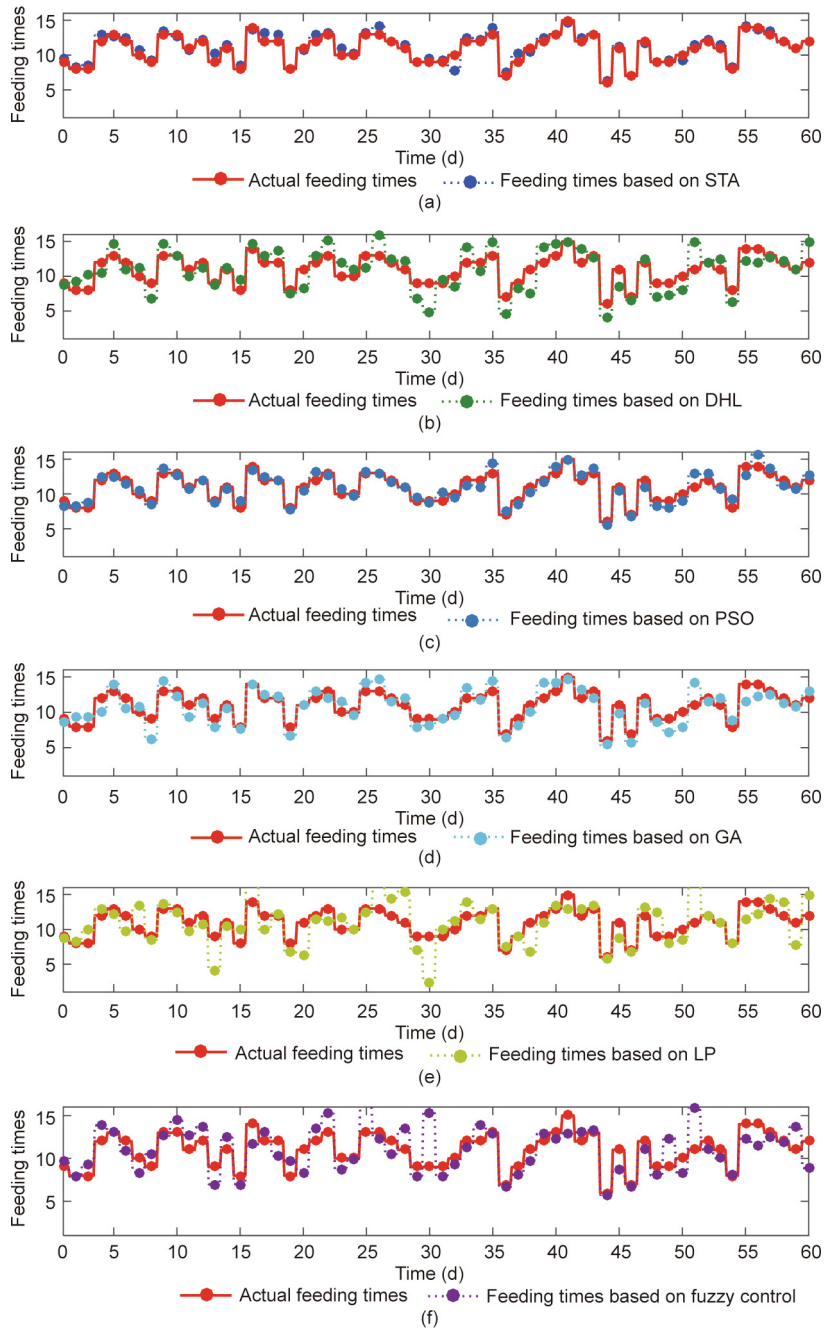**Fig. 14.** The values of (a) VI and WA, and (b) MV.

Fig. 15. Feeding times based on each algorithm under the amended structure: (a) based on STA; (b) based on DHL; (c) based on PSO; (d) based on GA; (e) based on linear programming (LP); (f) based on fuzzy control.

structure with FKM, and highlights the minimum value of the mean absolute percentage error (MAPE).

Third, Table 4 and Figs. 16 and 17 show that the MAPEs of the AFA are 10.1028 and 10.0766 based on FKM and EFKM, respectively, under the initial structure, using the STA. This means that we can obtain a higher accuracy for the AFA based on EFKM and the STA under the initial structure. In comparison with the above result, according to Table 4 and Fig. 15, when the MAPE of AFA is 6.2521, based on the STA and EFKM under the amended structure, it is lower than the above result. This indicates that the accuracy of the AFA based on the amended structure is greater than that based on the initial structure. Moreover, in contrast to the existing research achieve-

ments for MDAAA, Fig. 15(e) and Fig. 15(f) reveal that the proposed strategy for MDAAA is more efficient, being based on EFKM, fuzzy decision trees, and the STA to obtain augmented FCMs.

5.3. Summary

The MDAAA model based on augmented FCMs was designed by data and knowledge collaboration; data was selected from the aluminum production database, and knowledge was extracted from aluminum reduction experts and production data. Knowledge-extraction methods were used to enrich the structure of the MDAAA model in order to weaken the dependence on experts.
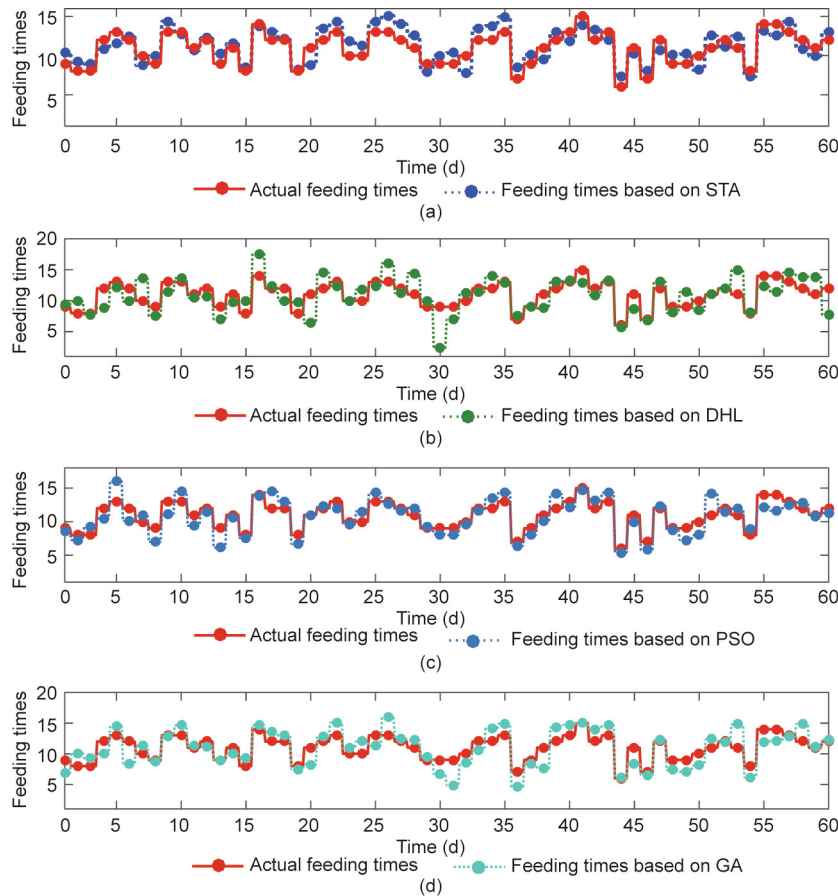
Fig. 16. Feeding times based on each algorithm with an initial structure using EFKM: (a) based on STA; (b) based on DHL; (c) based on PSO; (d) based on GA.

The data-driven methods were used to eliminate the subjectivity of experts, while the knowledge-driven methods were used to guide the MDAAA model construction. Due to the complexities of the aluminum reduction process, it is very difficult for technicians to provide all the fuzzy rules. EFKM and fuzzy decision trees were used for fuzzy rules extraction in order to reconstruct the initial structure of MDAAA. Due to the large number of weights, it was difficult for these to be provided by experts. In this study, the issue was solved by introducing the STA. In essence, the main task of this study was to propose a new strategy for MDAAA using FCMs based on data-driven and knowledge-driven methods. In practice, this is the first time data and knowledge collaboration based on FCMs has been introduced to MDAAA.

The data used in this study were selected to correspond to low VI, low WA, and low MV per day. First, the STA had a faster convergence rate than other learning methods, according to Figs. 10–13, and the MAPE was lower than that based on other learning algorithms. Second, the MAPE based on EFKM was lower than that based on FKM, and compared with the initial structure, the MAPE based on the STA was lower under the amended structure. In addition, having a maximum absolute error equal to 0.4498 meant that only an average of 0.8094 kg existed between the real feeding amount and the calculated value, which was acceptable and reasonable for the complicated aluminum reduction process. Third, in comparison with the existing MDAAA strategies, the proposed method had a better performance.

6. Conclusions

In this study, a data and knowledge collaboration strategy for MDAAA was proposed based on augmented FCMs. In practice, the AFA is decided by technicians, who combine the data report with experiential knowledge. FCMs are an efficient tool for capturing kinds of knowledge, dealing with complicated models, and processing knowledge problems, as they represent domain knowledge both visually and descriptively. Therefore, FCMs are more comprehensible than other strategies for aluminum reduction technicians. Knowledge-driven techniques were used for concepts selection and initial structure construction, and data-driven methods were used for knowledge extraction in order to reconstruct the initial structure, which enriched the knowledge-based system. The learning algorithm STA was introduced to detect the weights among the concepts of the FCMs; the results were compared with those of other learning algorithms. In addition, the proposed strategy for MDAAA was compared with those in existing research. The results show that the proposed strategy is valid and more effective than other strategies for MDAAA. The proposed strategy has potential for application in automatic decision-making for AlF$_3$ addition. The results show the prospective performance of the MDAAA model based on augmented FCMs. These results encourage us to continue work on another challenging and relevant problem—namely, feeding interval optimization in the aluminum reduction process.
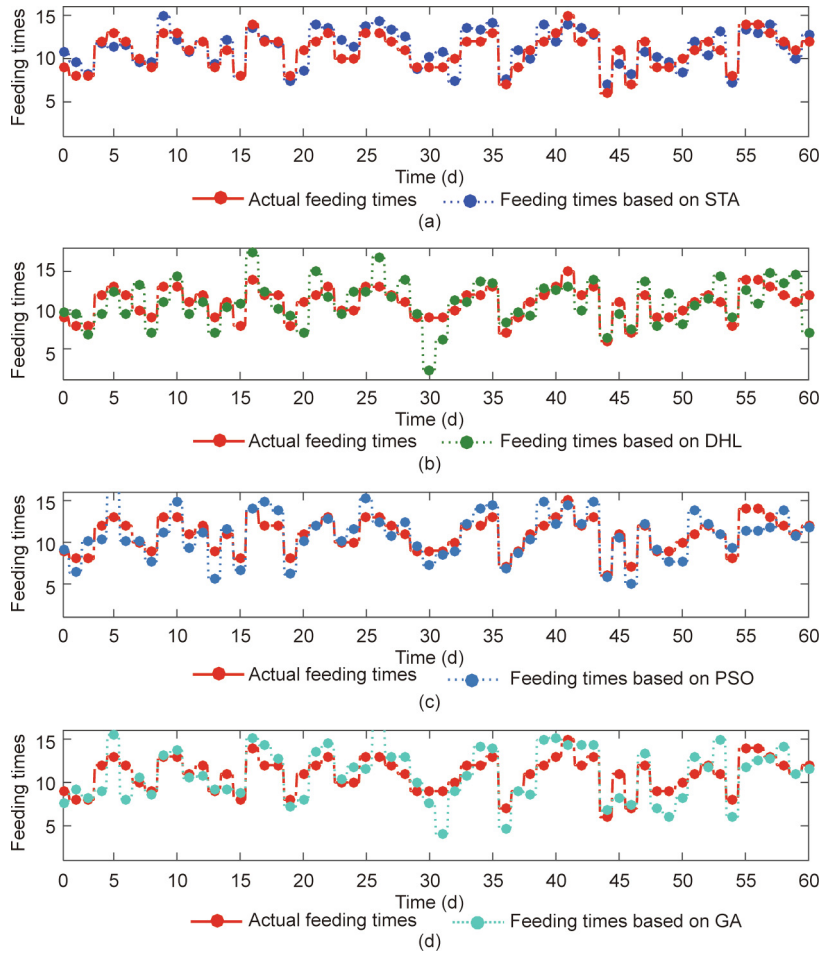
Fig. 17. Feeding times based on each algorithm with an initial structure using FKM: (a) based on STA; (b) based on DHL; (c) based on PSO; (d) based on GA.

Table 4
MAPE of each learning algorithm based on the amended and initial structure.

MAPE	Learning algorithm				Existing strategies	
	STA	DHL	PSO	GA	LP	Fuzzy control
Initial structure using FKM	10.1028	16.7194	11.2917	14.3077	—	—
Initial structure using EFKM	10.0766	15.0187	10.1657	14.0593	—	—
Amended structure	6.2521	13.9768	7.6852	9.6341	15.1288	13.4102

Acknowledgements

Compliance with ethics guidelines

Weichao Yue, Weihua Gui, Xiaofang Chen, Zhaohui Zeng, and Yongfang Xie declare that they have no conflicts of interest or financial conflicts to disclose.

References

[1] Yue W, Chen X, Gui W, Xie Y, Zhang H. A knowledge reasoning Fuzzy-Bayesian network for root cause analysis of abnormal aluminum electrolysis cell condition. Front Chem Sci Eng 2017;11(3):414–28.

[2] Kvande H. Current efficiency of alumina reduction cells. In: Campbell PG, editor. Light metals 1989: Proceedings of the Technical Sessions by the TMS Light Metals Committee at the 118 TMS Annual Meeting; 1989 Feb 27–Mar 3; Las Vegas, NV, USA. Pittsburgh: The Minerals, Metals & Materials Society; 1989. p. 261–8.

[3] Zeng S, Li J, Peng Q. Model predictive control of superheat for prebake aluminum production cells. In: DeYoung DH, editor. Light Metals 2008: Proceedings of the Technical Sessions Presented by the TMS Aluminum Committee at the TMS 2008 Annual Meeting & Exhibition; 2008 Mar 9–13; New Orleans, LA, USA. Pittsburgh: The Minerals, Metals & Materials Society; 2009. p. 347–50.

[4] Desclaux P. AlF$_3$ additions based on bath temperature measurements. In: Zabreznik R, editor. Light Metals 1987: Proceedings of the Technical Sessions Sponsored by the TMS Light Metal Committee at the 116th Annual Meeting; 1987 Feb 24–26; Denver, CO, USA. Pittsburgh: The Minerals, Metals & Materials Society; 1987. p. 309–13.

[5] Guo J, Gui W, Wen X. Multi-objective optimization for aluminum electrolysis production process. J Cent South Univ 2012;43(2):548–53.

[6] Kolås S. Defining and verifying the "correlation line" in aluminum electrolysis. JOM 2007;59(5):55–60.

[7] Kloetstra K, Benninghoff S, Stam M, Toebes B. Optimisation of aluminium fluoride control at Aluminum Delfzijl. In: Proceedings of the 7th Australasian Aluminum Smelting Technology Conference and Workshop; 2001 Nov 11–16;

Melbourne, VIC, Australia Sydney: Centre for Electrochemical and Minerals Processing. p. 506–14.

[8] Dupuis M. Excess AlF_3 concentration in bath control logic. In: Proceedings of National Conference on Advancements in Aluminum Electrolysis; 2006 Feb 11; Angul, India; 2006 p. 309–13.

[9] Salt DJ. Bath chemistry control system. In: Bearne G, Dupuis M, Tarcy G, editors. Essential readings in light metals. Cham: Springer; 2016. p. 798–803.

[10] Meghlaoui A, Al Farsi YA, Aljabri NH. Analytical and experimental study of fluoride evolution. In: Schneider WA, editor. Light Metals 2002: Proceedings of the Technical Sessions Presented by the TMS Aluminum Committee at the 131st TMS Annual Meeting; 2002 Feb 17–21; Seattle, WA, USA. Pittsburgh: The Minerals, Metals & Materials Society; 2002. p. 283–8.

[11] Entner PM, Gudmundsson GA. Further development of the temperature model. In: Hale W, editor. Light Metals 1996: Proceedings of the Technical Sessions Presented by the TMS Light Metals Committee at the 125th TMS Annual Meeting; 1996 Feb 4–8; Anaheim, CA, USA. Pittsburgh: The Minerals, Metals & Materials Society; 1996. p. 445–9.

[12] Zeng S, Cui F. Dynamic decision model for amount of AlF3 addition in industrial aluminum electrolysis. In: Proceedings of the 3rd International Conference on Mechatronics, Robotics and Automation; 2015 Apr 20–21; Shenzhen, China. Paris: Atlantis Press; 2015. p. 307–18.

[13] Hyland MM, Patterson EC, Stevens-Mcfadden F, Welch BJ. Aluminium fluoride consumption and control in smelting cells. Scand J Metall 2001;30 (2):404–14.

[14] Kolås S, Støre T. Bath temperature and AlF$_3$ control of an aluminum electrolysis cell. Control Eng Pract 2009;17(9):1035–43.

[15] Huang Y, Qu X, Zhou JM. Coupled heat/mass-balance model for analyzing correlation between excess AlF$_3$ concentration and aluminum electrolyte temperature. Trans Nonferrous Met Soc China 2009;19(3):724–9.

[16] Haupin W, Kvande H. Mathematical model of fluoride evolution from hall-héroult cells. In: Bearne G, Dupuis M, Tarcy G, editors. Essential readings in light metals. Cham: Springer; 2016. p. 903–9.

[17] Drengstig T, Ljungquist D, Foss BA. On the AlF$_3$ and temperature control of an aluminum electrolysis cell. IEEE Trans Contr Syst Technol 1998;6(2):157–71.

[18] Wee YY, Cheah WP, Tan SC, Wee KK. A method for root cause analysis with a Bayesian belief network and fuzzy cognitive map. Expert Syst Appl 2015;42 (1):468–87.

[19] Papageorgiou EI, Salmeron JL. A review of fuzzy cognitive maps research during the last decade. IEEE Trans Fuzzy Syst 2013;21(1):66–79.

[20] Mls K, Cimler R, Vaščák J, Puheim M. Interactive evolutionary optimization of fuzzy cognitive maps. Neurocomputing 2017;232:58–68.

[21] Nápoles G, Concepción L, Falcon R, Bello R, Vanhoof K. On the accuracy-convergence tradeoff in sigmoid fuzzy cognitive maps. IEEE Trans Fuzzy Syst 2018;26(4):2479–84.

[22] Zdanowicz P, Petrovic D. New mechanisms for reasoning and impacts accumulation for rule-based fuzzy cognitive maps. IEEE Trans Fuzzy Syst 2018;26(2):543–55.

[23] Kottas TL, Boutalis YS, Karlis AD. New maximum power point tracker for PV arrays using fuzzy controller in close cooperation with fuzzy cognitive networks. IEEE Trans Energ Convers 2006;21(3):793–803.

[24] Baykasoğlu A, Gölcük İ. Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS. Inf Sci 2015;301:75–98.

[25] Zhang L, Chettupuzha AJA, Chen H, Wu X, AbouRizk SM. Fuzzy cognitive maps enabled root cause analysis in complex projects. Appl Soft Comput 2017;57:235–49.

[26] Rezaee MJ, Yousefi S, Babaei M. Multi-stage cognitive map for failures assessment of production processes: an extension in structure and algorithm. Neurocomputing 2017;232:69–82.

[27] Froelich W, Pedrycz W. Fuzzy cognitive maps in the modeling of granular time series. Knowl Base Syst 2017;115:110–22.

[28] Obiedat M, Samarasinghe S. A novel semi-quantitative fuzzy cognitive map model for complex systems for addressing challenging participatory real life problems. Appl Soft Comput 2016;48:91–110.

[29] Papageorgiou EI. Learning algorithms for fuzzy cognitive maps: a review study. IEEE Trans Syst Man Cybern C 2012;42(2):150–63.

[30] Stach W, Kurgan L, Pedrycz W. Data-driven nonlinear Hebbian learning method for fuzzy cognitive maps. In: Proceedings of 2008 IEEE World Congress on Computational Intelligence; 2008 Jun 1–6; Hong Kong, China. New York: Institute of Electrical and Electronics Engineers; 2008. p. 1975–81.

[31] Konar A, Chakraborty UK. Reasoning and unsupervised learning in a fuzzy cognitive map. Inf Sci 2005;170(2–4):419–41.

[32] Froelich W, Salmeron JL. Evolutionary learning of fuzzy grey cognitive maps for the forecasting of multivariate, interval-valued time series. Int J Approx Reason 2014;55(6):1319–35.

[33] Salmeron JL, Rahimi SA, Navali AM, Sadeghpour A. Medical diagnosis of rheumatoid arthritis using data driven PSO-FCM with scarce datasets. Neurocomputing 2017;232:104–12.

[34] Luo X, Wei X, Zhang J. Guided game-based learning using fuzzy cognitive maps. IEEE Trans Learn Technol 2010;3(4):344–57.

[35] Baykasoglu A, Durmusoglu ZDU, Kaplanoglu V. Training fuzzy cognitive maps via extended great deluge algorithm with applications. Comput Ind 2011;62 (2):187–95.

[36] Christoforou AS, Andreou AS. A framework for static and dynamic analysis of multi-layer fuzzy cognitive maps. Neurocomputing 2017;232:133–45.

[37] Mateou NH, Moiseos M, Andreou AS. Multi-objective evolutionary fuzzy cognitive maps for decision support. In: Proceedings of the IEEE Congress on Evolutionary Computation; 2005 Sep 2–4; Edinburgh, UK. New York: Institute of Electrical and Electronics Engineers; 2005. p. 824–30.

[38] Stach W, Kurgan L, Pedrycz W, Reformat M. Genetic learning of fuzzy cognitive maps. Fuzzy Sets Syst 2005;153(3):371–401.

[39] Papageorgiou EI, Stylios CD, Groumpos PP. Active Hebbian learning algorithm to train fuzzy cognitive maps. Int J Approx Reason 2004;37(3):219–49.

[40] Zhou X, Yang C, Gui W. State transition algorithm. J Ind Manage Optim 2012;8 (4):1039–56.

[41] Zhang F, Yang C, Zhou X, Gui W. Fractional-order PID controller tuning using continuous state transition algorithm. Neural Comput Appl 2018;29 (10):795–804.

[42] Han J, Yang C, Zhou X, Gui W. A new multi-threshold image segmentation approach using state transition algorithm. Appl Math Model 2017;44:588–601.

[43] Han J, Yang C, Zhou X, Gui W. Dynamic multi-objective optimization arising in iron precipitation of zinc hydrometallurgy. Hydrometallurgy 2017;173: 134–48.

[44] Kulluk S, Özbakır L, Baykasoğlu A. Fuzzy DIFACONN-miner: a novel approach for fuzzy rule extraction from neural networks. Expert Syst Appl 2013;40 (3):938–46.

[45] Barakat N, Bradley AP. Rule extraction from support vector machines: a review. Neurocomputing 2010;74(1–3):178–90.

[46] Yuan Y, Shaw MJ. Induction of fuzzy decision trees. Fuzzy Sets Syst 1995;69 (2):125–39.

[47] Jain AK. Data clustering: 50 years beyond k-means. Pattern Recognit Lett 2010;31(8):651–66.

[48] Dickerson JA, Kosko B. Virtual worlds as fuzzy cognitive maps. Presence (Camb) 1994;3(2):173–89.

[49] Zeng SP, Li JH, Ren BJ. Fuzzy determination of AlF$_3$ addition and aluminum tapping volume in aluminum electrolyzing process. Metall Ind Auto 2008;32 (1):18–21.