Inference and Multi-level Learning in a Belief Rule-Based Expert System to Predict Flooding

Raihan Ul Islam
Pervasive and Mobile Computing Laboratory
Luleå University of Technology, Sweden
Email: raihan.ul.islam@ltu.se

Mohammad Shahadat Hossain
Department of Computer Science and Engineering
University of Chittagong, Bangladesh
Email: hossain_ms@cu.ac.bd

Karl Andersson
Pervasive and Mobile Computing Laboratory
Luleå University of Technology, Sweden
Email: karl.andersson@ltu.se

***Abstract*—Floods are one of the most dangerous catastrophic events. By the year 2050 flooding due to rise of ocean level may cost one trillion USD to coastal cities. Since flooding involves multi-dimensional elements, its accurate prediction is difficult. In addition, the elements cannot be measured with 100% accuracy. Belief rule-based expert systems (BRBESs) can be considered as an appropriate approach to handle this type of problem because they are capable of addressing uncertainty. However, BRBESs need to be equipped with the capacity to handle multi-level learning and inference to improve its accuracy of flood prediction. Therefore, this paper proposes a new learning and inference mechanism, named joint optimization using belief rule-based adaptive differential evolution (BRBaDE) for multi-level BRBES, which has the capability to handle multi-level learning and inference. Various machine learning methods, including Artificial Neural Networks (ANN), Support Vector Machine (SVM), Linear Regression and Long Short Term Memory have been compared with BRBaDE. The result exhibits that our proposed learning mechanism performs betters than learning techniques as mentioned above in terms of accuracy in flood prediction.**

***Contribution*—A new learning and inference mechanism for multi-level BRBES has been proposed in this paper, which has the capability to handle multi-level learning and inference.**

***Index Terms*—optimisation, learning, flood prediction, belief-rule-based expert systems**

I. Introduction

Climate change is a big phenomenon for the world, which might cause different types of natural devastation, like sea level rise, storms, intense drought, heavy raining, and flooding [1]. An estimated one trillion USD may be incurred due to flooding caused from sea level rise for the cities near coastal areas until 2050 [2]. Flooding is responsible for the highest number of fatalities as well as socio-economic damages among various natural disasters [1]. Therefore, to save human lives, infrastructure and economy, the prediction of flooding with the highest accuracy is necessary. This would play an important role to develop an evacuation plan as well as to take flood prevention measures.

Normally, flood contributing factors in an area including meteorological, topographical, geological, river characteristics, and human activities are considered [3]. They can be measured using both quantitative and qualitative data. Hence, an integrated framework is necessary, which should be able to process both types of data, enabling the accurate prediction of flooding. By conducting interviews with the flood-affected people, looking at the historical records and literature as well as by deploying various sensors, this data can be acquired [3]. Sensor data may contain duplicate, inconsistent, noisy, incorrect, and erroneous data, which can lead to different types of uncertainty [4]. Furthermore, data gathered from surveying people can be inaccurate or uncertain because people can express some answers in vague linguistic terms, such as 'Big', 'Medium', and 'Small'. Moreover, data acquired from human are usually not accurate because of insufficient knowledge and misperception, resulting in ignorance and incompleteness.

Researchers have tried to predict flooding using different algorithms including Artificial Neural Networks (ANNs), Support Vector Machine (SVM), and Linear Regression [1], [5], [6]. Seal et al. [5] and Shegal et al. [6] considered the use of Linear regression oriented models to predict flooding. Granata et al. [7] used SVM to predict flood water level. Both the linear regression and SVM lack the procedures to handle uncertainty. Furthermore, Solaiman et al. [8] also used ANN to predict flood water level. However, the reasoning by ANN cannot be explained as its intermediate layers are not visible. Hence ANN is considered as a black box [9]. Moreover, ANN endures from growing dimensionality [10]. Thus, to deal with uncertainties, causing from both qualitative and quantitative data, an integrated framework is required necessary to predict flooding with higher accuracy.

By taking into account Belief Rule Base (BRB) as a knowledge representation paradigm as well as Evidential Reasoning (ER) as the inference engine, a Belief rule-based expert system (BRBES) is developed. BRB allows acquiring uncertain knowledge while ER enables handling uncertainties resulting from qualitative and quantitative data. Andersson et al. [11] proposed a system to predict flood using BRBES. However, this BRBES has been transformed into a Web-BRBES to handle sensor data in real time [12]. BRBES's framework is illustrated in Fig. 1, developed by taking into account the multi-level factors, causing flooding in a zone [11][12]. The major causes of flooding (X7), including Meteorologi-

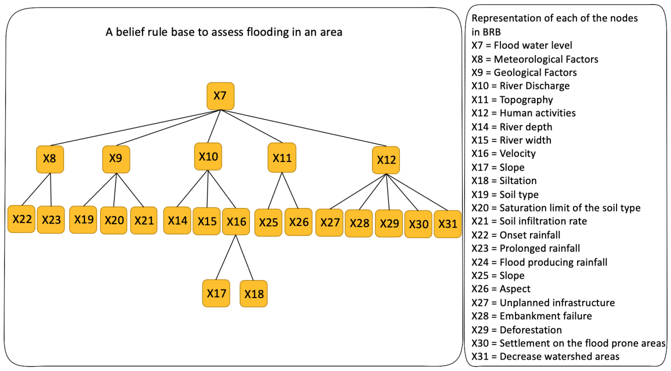

Fig. 1. A multi-level BRB framework for predicting flood water level.

cal Factors (X8), Geological Factors (X9), River Discharge (X10), Topography (X11), and Human Activities (X12) are considered at the intermediate level nodes of this BRBES framework. As mentioned previously, the data of the leaf nodes are collected either by sensors or by interviewing people and they are fed into the Web-BRBES to predict the flooding in a zone. Fig. 1 illustrates the seven BRBs where the lower level BRB is X16, mid-level BRBs are X8, X9, X10, X11, and X12 and the top-level BRB is X7. By following a bottom-up approach information propagates.

The learning parameters of BRBES include rule weight, attribute weight and belief degrees. However, in case of multi-level BRBES the calculation of the value of these learning parameters is difficult. These learning parameters play an important role in measuring the accuracy of the BRBES's output. Hence, it is necessary to develop a procedure allowing the assignment of optimal values to the BRBES's parameters. This can be considered as learning and inferencing for the system. Learning and inference for BRBES was proposed in [13]. Chang et al. [14] came up with a learning algorithm allowing the discovery of the optimal number of referential values known as structure optimization for BRBES. Additionally, Chang et al. [15] proposed a joint optimisation technique which combines the parameter and structure optimisation. All of the above research focuses on the framework of uni-level BRBES. As Fig. 1 illustrates a multi-level BRBES, the development of learning and inference algorithms is necessary. Eventually, flood prediction can be achieved with higher accuracy. Hence, to improve the prediction accuracy of flooding, our proposed solution demonstrates an algorithm, allowing the training of the learning parameters and structure of a multi-level BRBES.

The people of a flood affected zone of Bangladesh were interviewed to acquire data and used to validate both the BRBES and the novel learning algorithm. Moreover, the output of BRBES and the proposed multi-level learning algorithm were compared with the machine learning algorithms including SVM based regression, linear regression, ANN and LSTM. The mentioned machine learning algorithms were outperformed by the proposed learning algorithm. The BRBES and the proposed learning algorithm are capable of handling various uncertainties related to the elements as represented in Fig. 1.

The remaining of the paper is structured as follows. Section 2 presents the recent research works on the prediction of flooding, while Section 3 elaborates the BRBES. The joint optimisation using BRBaDE for multi-level BRBES is elaborated in Section 4. Section 5 discusses the architecture of the proposed learning algorithm, while Section 6 discusses the results and analysis. The conclusion of the paper is included in Section 7.

II. Related Work

Flooding is a complex phenomenon which is influenced by several elements like metrological, geological, unplanned infrastructure, and deforestation. Therefore, it is challenging to predict flood accurately. Researchers have used different machine learning techniques for predicting a flood of an area using several different contributing factors [5][7][16]. Sulaiman et al. [8] used ANN for predicting flood in Pahang, Malaysia, based on rainfall in that area. A comparison of the performance between ANN model and Auto Regression Integrated Moving Average (ARIMA) was carried out. The results demonstrate that the ANN model is dependable in anticipating the risk level of precipitation. The research mentioned above did not consider other flood influential elements related to topography, geological, and human activities, which are presented in Fig. 1. Furthermore, these influential factors are qualitative and quantitative in nature. Hence, BRBES, including various influential factors, has better potential to predict flood accurately than ANN since BRBES has the capability of handling qualitative and quantitative data.

Ganata et al. [7] compared the SWMM (Storm Water Management Model) and SVM based approach to model rainfall-runoff. The SVM model showed greater accuracy for urban hydrology compared to SWMM. SVM fails to address the uncertainty of data as the input data are not distributed over belief degrees.

Sear et al. [5] presented a flood forecasting system based on rainfall, water discharge, and temperature collected by sensors. They used a linear regression model for forecasting flood water level. Sensors dispatch the measured data to the computational server using intermediate nodes. The water levels of flood are predicted and passed to a monitoring server for users to view. Sensor data usually contains various uncertainty [4]. However, the linear regression model-based forecasting model lacks the procedures to address uncertainty. BRBES based systems have the ability to handle uncertainty of various types, which will allow flood prediction with higher accuracy [11].

Thus, due to its capability of addressing uncertainty BRBES will allow more accurate flood prediction than from ANN, SVM regression, and linear regression. Furthermore, the learning of parameters are considered as important to increase the BRBES's accuracy. Usually, the learning parameters and the number of referential values are selected by experts. It is not always possible for the experts to select optimal values, especially when the dataset is significantly larger [17]. However, the optimal values can be learned from the dataset. Therefore, to measure the closer optimal values of the learning parameters, an efficient and robust learning mechanism is needed, which will improve the BRBES's prediction accuracy.

Yang et al. [17] first proposed an optimisation model for BRBES. The authors suggested optimisation of the parameters using the non-linear constrained solver, named *fmincon* from the MATLAB optimisation toolbox. Sequential quadratic programming algorithm is used in the *fmincon* solver, which is a deterministic algorithm. The proposed method is prone to get stuck in local optima due to its gradient-based mechanism. Additionally, this method does not have any mechanism for addressing the uncertainty of the objective functions.

Xu et al. [18] applied the optimisation model mentioned above for pipeline leak detection and demonstrated the usefulness of incorporation of learning in the BRBES. The proposed model by Xu et al. [18] has been improved by adding utility values as one of the learning parameters by Chen et al. [19], which generates better results. However, the later proposed model inherited the shortfalls of the *fmincon* method. Hossain et al. [20] proposed a single level BRBES, allowing the prediction of power usage effectiveness (PUE) for datacentres based on outdoor and server room temperatures. The accuracy of PUE prediction by BRBES has been validated with real-world data and compared with other algorithms, like ANN and the Genetic algorithm. The comparison exhibited that after including learning, the non trained BRBES performed worse than the trained BRBES. Zhou et al. [21] suggested using an evolutionary algorithm named Clonal Selection Algorithm (CSA) due to its success in achieving the optimal or near-optimal solution for problems with non-linear and continuous search space. Besides, other evolutionary algorithms, like Particle Swarm Optimisation [22] and Differential evolution (DE) [23] were used for the optimisation of BRBES, where DE performed better than the others.

Yang et al. [24] came up with a Principal Component Analysis (PCA) based method to discover closer optimal numbers of referential values, which produce better results and requires less computation time as a reduced number of antecedent attributes are used.

Yang et al. [23] also came up with a joint optimisation algorithm for BRBES. A generalisation error based on the Hoeffding inequality theorem was recommended instead of root mean square error. A heuristic algorithm used for structure optimisation and DE algorithm used for parameter optimisation. This study shows the efficiency of DE for parameter optimisation of BRBES. However, discovering the closer optimal values of the control parameters of DE is not always easy. Usually, a trial and error based approach is used for finding the closer optimal values of the control parameters. Islam et al. [25] proposed a Belief Rule-based Adaptive DE (BRBaDE) based joint optimisation algorithm for single level BRBES, where optimal values of the control parameters were identified using BRBES.

In summary, BRBES presents a better opportunity for predicting flood water level considering its ability to address different types of uncertainty and providing a single framework to handle quantitative and qualitative data. Researchers have proposed different learning mechanisms for incorporating learning in BRBES to improve its accuracy. However, most of the research work discussed above presented learning mechanisms for the single level BRBES. Hence, a new learning algorithm is required to predict flood under uncertainty using complex multi-level BRBES.

III. Methodology of Belief Rule-Based Expert Systems

BRBES is an expert system which is capable of handling various categories of uncertainties, including incompleteness, ignorance, vagueness, and ambiguity, while processing both qualitative and quantitative data [26]. The knowledge base of BRBES is represented using BRB while the inference procedure consists of various procedures such as 1) transformation of input data, 2) activation of rule weight, 3) belief degrees update, and 4) aggregation of rules by deploying evidential reasoning [26].

In the following subsection, a brief description of the BRBES methodology and its learning mechanism is given. Afterwards, a new learning mechanism named joined optimisation using belief rule-based adaptive differential evolution (BRBaDE) for multi-level BRBES is proposed.

A. The BRBES knowledge representation

BRB is the improved version of classical IF-THEN rule base. It comprises antecedent and consequent. According to Fig. 1, a BRB consists of X8 (Meteorological factors), X22 (Prolonged rainfall), and X23 (Onset rainfall), where the antecedent attributes are X22 and X23 and the consequent attribute is X8. Each antecedent attribute has referential values while each referential value of each consequent attribute has a belief degree as shown in Eq. (1). To demonstrate the importance of a rule among the others a rule weight is considered.

$$R_k : \begin{cases} \text{IF } (A_1 \text{ is } V_1^k) \text{ AND / OR } (A_2 \text{ is } V_2^k) \text{ AND / OR } \ldots \\ \text{ AND / OR } (A_{T_k} \text{ is } V_{T_k}^k) \\ \text{THEN } (C_1, \beta_{1k}), (C_2, \beta_{2k}), \ldots, (C_N, \beta_{Nk}) \end{cases} \tag{1}$$

$$\text{where } \beta_{jk} \geq 0, \sum_{j=1}^{N} \beta_{jk} \leq 1 \text{ with rule weight } \theta_k,$$

$$\text{and attribute weights } \delta_{k1}, \delta_{k2}, \ldots \delta_{kTk}, k \in 1, \ldots, L$$

where, $A_1, A_2, \ldots, A_{T_k}$ are the antecedent attributes of the k^{th} rule. $V_i^k (i = 1, \ldots, T_k, k = 1, \ldots, L)$ is the referential value of the i^{th} antecedent attribute. C_j is the j^{th} referential value of the consequent attribute. $\beta_{jk} (j = 1, ..., N, k = 1, ..., L)$ is the degree of belief related to the reference value C_j of consequent believed to be true. If $\sum_{j=1}^{N} \beta_{jk} \leq 1$, then the k^{th} rule is considered to be complete; otherwise, it is incomplete. T_K denotes the number of antecedent attributes that are employed in the k^{th} rule. L is the number of belief rules in a BRB. There are two different types of BRB considered on the logical connector of the antecedent attributes. The AND operator represents the conjunctive BRB while the OR operator represents the disjunctive BRB. Considering the logical operators of the BRB, a BRBES can be termed as conjunctive or disjunctive BRBES. An example of belief rule follows:

$$R_k : \begin{cases} \text{IF X17 (Slope) is High} \wedge \\ \text{X18 (Siltation) is Medium} \\ \text{THEN X16 (Velocity) is} \\ \{\text{(Severe, 0.1), (Moderate, 0.4), (Low, 0.6)}\} \end{cases} \tag{2}$$

In this conjunctive rule, Slope and Siltation are the antecedent attributes, while "High" and "Medium" are their associated referential values. The velocity is the consequent attribute with referential values, such as "Severe", "Moderate", and "Low". Since the summation of belief degrees (0.1+0.4+ 0.5) related to the referential values of the consequent attribute is one, the rule is considered to be complete. On the contrary, this rule is incomplete if the summation of belief degrees is less than one because of incomplete information.

B. BRBES inference procedures

An inference procedure allows BRBES to predict values using BRB based on the input while addressing various types of uncertainty. BRBES uses an inference procedure consisting of transformation of input data, calculation of rule activation weight, update of belief degrees and aggregation of rules using evidential reasoning. These steps are described as follows:

1) Input transformation: Among the referential values of the antecedent attribute the distribution of input data is carried out in this inference procedure. For example, in Fig. 1, 30 degree may be considered as the input data of the leaf node X17. This is the antecedent attribute of the BRB X16. This 30 degree is to be converted into X17 antecedent attribute's referential values which are assumed as "High", "Medium", and "Low".

$$H(V_i) = (V_{ij}, \alpha_{ij}), j = 1, \ldots, j_i, i = 1, \ldots, T_k \tag{3}$$

In Eq. 3, the H function converts the input value of the antecedent attribute into the matching degree of its referential values. V_{ij} is the jth referential value of the input. α_{ij} is the matching degree to the referential value. For example, the input 30 degree of antecedent attribute X17 is converted by employing the utility function [26] into matching degrees of its referential values, which are (High, 0.08), (Medium, 0.92), (Low, 0.00).

2) Rule Activation: The converted values of the input data can be termed as the matching degrees. When the matching degree is allocated the rules are called packet antecedent and hence, they are active. The combination of the matching degrees of the attributes of a rule is necessary. This can be achieved by employing following the multiplicative weighted equation:

$$\alpha_k = \prod_{i=1}^{T_k} (\alpha_i^k)^{\bar{\delta}_{ki}} \tag{4}$$

$$\text{where } \bar{\delta}_{ki} = \frac{\delta_{ki}}{max_{i=1,\ldots,T_k}\{\delta_{ki}\}} \text{ so that } 0 \leq \bar{\delta}_{ki} \leq 1$$

Here, T_k is the total number of antecedent attributes in the k^{th} rule. To understand the complementarity or the integration between the antecedent attributes, the use of multiplicative equation is necessary.

3) Weight Activation: The activation weight of each conjunctive rule is determined using the combined matching degree of each rule, as shown in Eq. (5). The activation weight w_ks for the k^{th} rule can be generated by the following expression:

$$w_k = \frac{\theta_k \alpha_k}{\sum_{i=1}^{L} (\theta_i \alpha_i)} \tag{5}$$

Here, θ_k denotes the rule weight and α_k denotes the combined matching degree of the k^{th} rule.

However, the activation weight w_k for disjunctive assumption of the k^{th} rule can be calculated by the following expression:

$$w_k = \frac{\theta_k \sum_{i=1}^{M} \alpha_i^k}{\sum_{i=1}^{L} (\theta_k \sum_{i=1}^{M} \alpha_i^l)} \tag{6}$$

Here, θ_k is the rule weight and α_k is the matching degree of the k^{th} rule. All matching degrees are summed for this disjunctive assumption.

4) Belief Update: In many scenarios, data could not be found for any of the antecedent attributes of a BRB. As an example, the input data for one of the antecedent attributes X17 (Slope) of the BRB X16 (Velocity) cannot be collected because of ignorance, as shown in Fig. 1. The initial belief degrees of BRB X16 should be updated to address the ignorance by using Eq. (7) [26].

$$\beta_{jk} = \bar{\beta}_{jk} \frac{\sum_{t=1}^{T_k} (\lambda(t,k) \sum_{i=1}^{I_t} (\alpha_{ti}))}{\sum_{t=1}^{T_k} \lambda(t,k)} \tag{7}$$

$$\text{where } \lambda(t,k) = \begin{cases} 1 & \text{if the } t^{th} \text{ attribute is used in} \\ & \text{defining rule } R_k (k = 1, ..., T_k) \\ 0 & \text{otherwise} \end{cases}$$

Here, $\bar{\beta}_{jk}$ denotes the original belief degree, while the updated belief degree is β_{jk} of the k^{th} rule. α_{ti} denotes the degree to which the input value belongs to an attribute.

5) Rule Aggregation: The output of the BRBES is generated employing the evidential reasoning algorithm by aggregating the rules for the respective input data. The evidential reasoning is implemented using two approaches; these are analytical and recursive [18][26]. The analytical approach is preferable due to computationally efficiency [26]. Therefore, analytical evidential reasoning is used in this paper by using Eq. (8).

$$\beta_j = \frac{\mu \times [X - \prod_{k=1}^{L}(1-\omega_k \sum_{j=1}^{N}\beta_{jk})]}{1 - \mu \times [\prod_{k=1}^{L} 1-\omega_k]} \tag{8}$$

where $\mu = \left[\sum_{j=1}^{N}\prod_{k=1}^{L}(\omega_k\beta_{jk} + 1 - \omega_k\sum_{j=1}^{N}\beta_{jk}) - (N-1) \times \prod_{k=1}^{L}(1-\omega_k\sum_{j=1}^{N}\beta_{jk})\right]^{-1}$

$$X = \prod_{k=1}^{L}(\omega_k\beta_{jk} + 1 - \omega_k\sum_{j=1}^{N}\beta_{jk})$$

Here, ω_k represents the activation weight of the k^{th} rule, whereas the belief degree associated with one of the consequent reference values is denoted by β_j.

The output of the rule aggregation step is fuzzy value. Using Eq. (9), the fuzzy value is converted into a crisp value.

$$z_i = \sum_{j=1}^{N} u(O_j)\beta_j \tag{9}$$

where z_i is the expected numerical value and $u(O_j)$ is the utility score of each referential value.

C. Optimal learning procedure for multi-level BRBES

The optimal learning algorithm for multi-level BRBES, enabling the robust prediction of flood water level is presented in this subsection. A multi-level BRBES can be termed as the group of multiple BRBs, structured in hierarchical order. Instance, Fig. 1 depicts a group of BRBs, that are organised following the structure of a tree. Domain experts or randomly usually assigned the value of the learning parameters including rule weights (θ_k), attribute weights (δ_i), and belief degrees (β_k) of a BRB are assigned by experts in the domain or by generating random numbers [26]. Both the attribute and rule weights are used to prioritise the antecedent attribute and rules. The uncertainty of the output is elaborated by the Belief degrees of the consequent attribute. Hence, the learning parameters play an important role in the inference mechanism. The rational and logical output from the lower levels goes to upper levels in a multi-level hierarchical BRBES. Eventually, this produces the final result more rational and dependable. Consequently, optimal learning parameters enable producing of a dependable output for the multi-level BRBES.

Hence, in order to calculate the closer optimal values of the learning parameters a new algorithm is required. According to [17], from data, the closer optimal values of the learning parameters can be obtained by employing training algorithms. Usually, in the light of the liner inequality and equality constraints a single objective function is employed to calculate the closer optimal values of the learning parameters. The gap $\xi(p)$ between BRBES output (z_m) and the real system output ($\bar{z}_m$) is required to be reduced during the optimisation process. Let's consider M the number of cases in a training sample, where the input is u_m, the observed output is $\bar{z}_m$, and the simulated output is z_m ($m = 1, \ldots, M$)). The gap $\xi(p)$ is calculated employing Eq. (10).

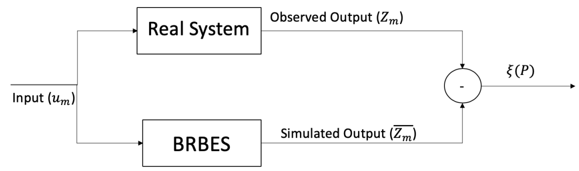

Fig. 2. The learning process of the BRBs.

$$\xi_{(p)} = \frac{1}{M}\sum_{m=1}^{M}(z_m - \bar{z}_m)^2 \tag{10}$$

The training process is conducted on each BRB. To minimise the gap $\xi(p)$, the optimisation of the values of the learning parameters is carried out as elaborated in the following equation:

$$\min_{p} \xi(p) \tag{11}$$

$$P = P(\mu(O_j), \theta_k, \delta_k, \beta_{jk})$$

Eqs. (8) and (9) are employed to build the objective function to support training in the BRB. The learning parameters comprise some constraints. The normalization of the attribute weights, rule weights, utility values of the consequent attributes, and belief degrees is considered between zero to one. The summations of the belief degree for each rule are assumed to be one to ensure the completeness of the rule. Therefore, to reflect the above, the following constraints are taken into account for each of the learning parameters:

- Utility values of the consequent attributes $\mu(O_j)(j = 1, \ldots, n)$:
  $1 \geq \mu(O_j) \geq 0$;
  $\mu(O_i) < \mu(O_j)$; If $i < j$
- Rule weights $\theta_k(k = 1, \ldots, K)$:
  $1 \geq \theta_k \geq 0$;
- Antecedent attribute weights $\delta_k, (k = 1, \ldots, K)$:
  $1 \geq \delta_k \geq 0$;
- Consequent belief degrees for the kth rule $\beta_{jk}, (j = 1, \ldots, n, k = 1, \ldots, L)$:
  $1 \geq \beta_{jk} \geq 0$;
  $\sum_{j=1}^{n}\beta_{jk} = 1$;

IV. Learning in BRBES-based on BRBaDE based joint optimisation

A new joint optimisation technique to improve the accuracy of the BRBES will be proposed in the following sections. The

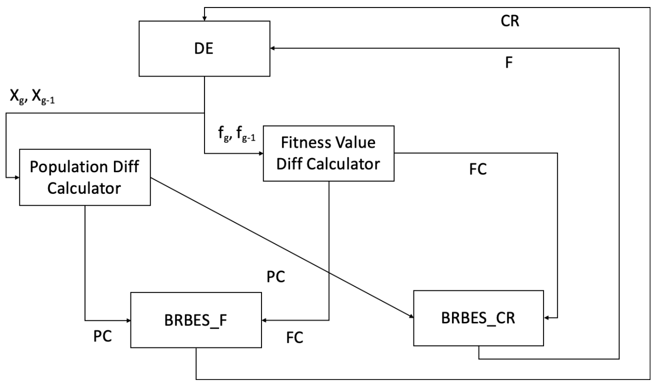

Fig. 3. Belief Rule-Based Adaptive Differential Evolution (BRBaDE).

joint optimisation algorithm has two parts, one is structure, and the other is the parameter optimisation algorithm. This algorithm is described below.

A. Parameter optimisation with BRBaDE

DE is a population-based evolutionary algorithm, which has the capability of avoiding local optima and performs comparatively better in accurately discovering global optima. The mutation (F) and crossover (CR) factors deeply affect the performance of DE. Previous research shows that DE performs well if the F and CR are changed during each iteration . On the contrary, this has not been taken into account by the earlier studies. Hence, BRBaDE, which integrates BRBES and DE [25], has been used for parameter optimization under uncertainty. Fig. 3 illustrates the components of BRBaDE. In BRBaDE, two BRBESs are employed, enabling the prediction of F and CR values by using the changes in population and objective function values in each iteration. Using Eqs. (12), (13), (14), (15), (16), and (17) the changed values of population and objective function under each iteration are calculated. Table II and Fig. 4 present the rule base and the framework of 'BRBES_F', which is used for predicting the value of F during each iteration. Similarly, the rule base and the framework of 'BRBES CR' are presented in Table I and Fig. 4, which is used for predicting the value of CR.

The two BRBESs find the balanced exploration and exploitation in the search space by taking into account the changed population and objective function values under each iteration. Furthermore, a penalty is added to the objective function values if any of the constraints mentioned in subsection III-C are violated.

$$PC = \sqrt{\frac{1}{NP}\sum_{i=1}^{NP}\sum_{j=1}^{D}(x_{j,i}^{g} - x_{j,i}^{(g-1)})^2} \tag{12}$$

$$FC = \sqrt{\frac{1}{NP}\sum_{i=1}^{NP}(f_i^{g} - f_i^{(g-1)})^2} \tag{13}$$

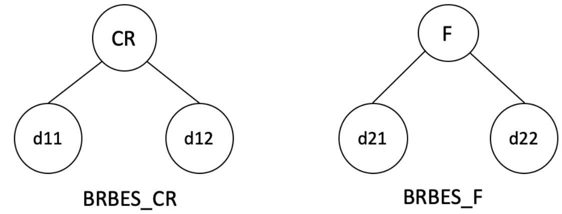

Fig. 4. Two BRBESs used for BRBaDE.

TABLE I
DETAILS OF BRBES_CR

	Antecedent Attributes			Consequent Attribute		
	d_{11}/d_{12}			CR		
Referential Values	Big	Medium	Small	Big	Medium	Small
Utility Values	1	0.5	0	1	0.75	0.1

$$d_{11} = 1 - (1 + PC)e^{-PC} \tag{14}$$

$$d_{12} = 1 - (1 + FC)e^{-FC} \tag{15}$$

$$d_{21} = 2d_{11} \tag{16}$$

$$d_{22} = 2d_{12} \tag{17}$$

In Eq. (12), PC constitutes the changed magnitude of the population vector during the last two iterations, while $x_{j,i}^{g}$ and $x_{j,i}^{g-1}$ are the population vectors on the g^{th} and $(g-1)^{\text{th}}$ iteration respectively. The change in the magnitude of the objective function during the last two iterations is denoted by FC. However, f_i^{g} and $f_i^{(g-1)}$ are the values of the objective function for the i^{th} population in the g^{th} and $(g-1)^{\text{th}}$ iteration. Using Eqs. (14) and (15), the values of PC and FC are normalised between 0 to 1, where d_{11} and d_{12} contains the the altered value of PC and FC respectively. Likewise, the value of PC and FC are altered between 0 to 2 employing Eqs. (16) and (17), which are subsequently used as inputs for BRBES to determine new values of F and CR.

Hence, by integrating BRBES with DE, the BRBaDE enables handling the uncertainty of objective functions. In addition, BRBaDE makes it possible to achieve both optimal exploration and exploitation in the search space. This will eventually enable finding the optimal solution with lower iterations.

B. Structure optimisation

An optimal structure of the multi-level BRB is needed for obtaining better accuracy from BRBES besides parameter

TABLE II
DETAILS OF BRBES_F

	Antecedent Attributes			Consequent Attribute		
	d_{21}/d_{22}			F		
Referential Values	Big	Medium	Small	Big	Medium	Small
Utility Values	2	1	0	2	1	0.1

optimisation. A structure optimisation algorithm is proposed in Algorithm 1 to find the optimal structure of a multilevel BRB. A top-down approach is used for finding the optimal structure for the multi-level BRB. At first, parameter optimisation of the multi-level BRB is performed by the following bottom-up approach. Afterwards, the referential values of the antecedent attributes of the topmost BRB is increased, and parameter optimisation of the multi-level BRB is performed. If the new multi-level BRB performs better than the previous level BRB, then the number of referential values is increased, and parameter optimisation is performed to determine the performance of the new multi-level BRB. If the performance of BRBES is not getting better, then move to the next BRB and perform the similar operation until all the BRBs are traversed. A graphical representation of Algorithm 1 is shown in Fig. 5.

Algorithm 1 Structure optimisation

Let $T = t_{i,j}, \dots t_{n,m}$ denote the multi-level BRB. Here $t_{i,j}$ denotes i^{th} BRB at j^{th}level, where $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$

Input T

Output Optimised multi-level tree X

procedure STRUCTURE_OPTIMISATION(T)
  $T \cup t'_{i,j}$ ▷ Generate new structure for $t_{i,j}$ BRB by increasing the number of referential values. Here $t_{i,j}$ is the top most BRB of multi-level BRB T.
  $previous\_RMSE$:=inf
  $Continue := true$
  while $Continue = true$ **do**
    new_RMSE:=BRBaDE(T) ▷ Parameter optimisation using BRBaDE
    if $new\_RMSE < previous\_RMSE$ **then**
      $X \cup t'_{i,j}$ ▷ replace the previous $t'_{i,j}$ BRB with new $t'_{i,j}$ BRB
      $previous\_RMSE := new\_RMSE$
    else if number of referential value = M **then**
      $i := i + 1$ ▷ move to next BRB at level j
      if all BRBs of the same level is visited **then**
        $j := j + 1$ ▷ move to next level
    $T \cup t'_{i,j}$ ▷ Generate new structure for $t_{i,j}$ BRB by increasing the number of referential values of $t_{i,j}$ BRB till M.
    if $(i = n) AND (j = m)$ **then** ▷ all BRB trees are visited
      $Continue := false$
  return X

In brief, by using BRBaDE the parameter optimisation (PO) is accomplished by taking into account the initial BRB. On the other hand, structure optimisation (SO) is accomplished by employing Algorithm 1. The stop criterion is met when the number of iterations reaches the threshold value and such a scenario is termed as an optimised BRB. Otherwise, the loop continues.

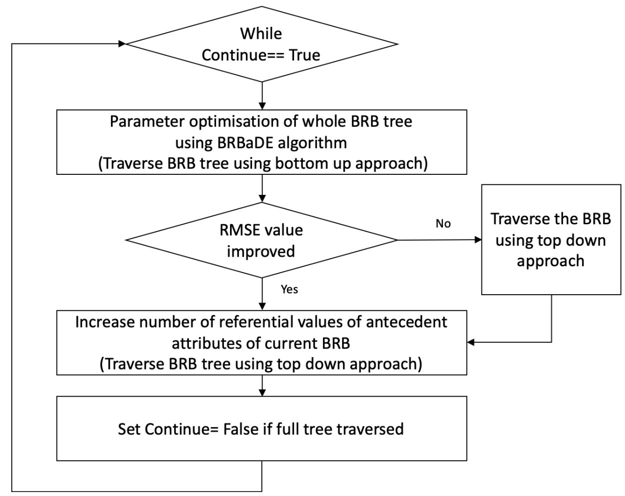

Fig. 5. Flowchart of structure optimisation.

V. SYSTEM ARCHITECTURE OF THE TRAINED BRBES

A system for predicting flood water level using the proposed joint optimisation using BRBaDE is introduced in this section. The system is divided into six modules. These are an Input Module, a BRB Main Module, a Knowledge Base Module, a Knowledge Base Driver Module, a Training Module, a Configuration Module, and a BRB UI Module. This system is based on Web-BRBES [12]. This subsection presents the discussion on each of these modules.

A. Input Module

The BRB receives data from this module. The Input Module uses different data sources which are set in a file of the Configuration Module. It receives data from CSV files and RESTful APIs. This module also receives training data for the learning mechanism.

B. BRB Main Module

The main module of BRB is the core of BRBES. It communicates with other modules and performs flood prediction. This module uses the Training Module to learn about the learning parameters of BRBES utilising the joint optimisation using the BRBaDE algorithm while applying the training data from the Input Module. Afterwards, by taking into account the training learning parameters new rules are created, which are stored in the Knowledge Base Module. After receiving the data from the Input Module, the BRB Main Module distributes the data into the referential values based on the utility function. The inference procedure is carried out based on the ER [26] algorithm outlined in subsection III-B. The matching degree and rule activation weight are calculated by employing Eqs. (4), (5), and (6), which is the part of this inference procedure. Due to the presence of ignorance in data, belief degree update is performed by using Eq. (7). Afterwards, rule aggregation is performed using Eq. (8), which produces the predicted fuzzy flood water level. By using Eq. (9), the

TABLE III
INPUT VARIABLES

No	Input Antecedent	Name of node	Referential Values		
			High	Medium	Low
1	Onset rainfall (mm)	X22	110	56	2
2	Prolonged rainfall (day)	X23	11	5.75	0.5
3	Soil type	X19	1	0.5	0
4	Saturation limit of Soil	X20	20	15	10
5	Soil infiltration rate	X21	0.5	0.25	0
6	Slope (degree)	X17	26	18.50	11
7	Siltation	X18	2	1	0
8	River depth (meter)	X14	10	7	3
9	River width (meter)	X15	105	77	35
10	Slope (degree)	X25	32	16	0
11	Aspect	X26	0.5	0.25	0
12	Unplanned infrastructure	X27	1.1	0.85	0.6
13	Embankment failure	X28	1	0.5	0
14	Deforestation	X29	1.10	0.85	0.6
15	Settlement on the flood prone areas	X30	1.10	0.85	0.6
16	Decrease watershed areas	X31	1.10	0.85	0.6

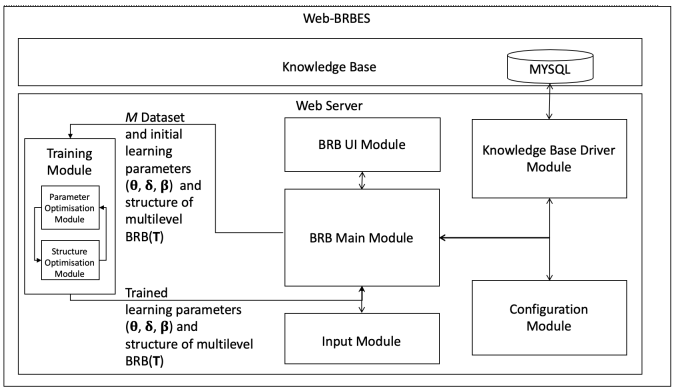

Fig. 6. System Architecture of Web-BRBES.

fuzzy values are converted into crisp values which are the predicted flood water levels.

C. Knowledge Base Module

The BRB is stored in the Knowledge Base Module using MySQL Server. The MySQL Server consists of high-performance query engine, which helps in faster data insertion and querying. This server helps managing large numbers of belief rules generated by the multi-level BRB framework. The knowledge base of the BRBES for flood prediction is constructed by taking into account the framework of BRB as illustrated in Fig. 1, which is later modified by the Training Module. The input antecedent attribute and their corresponding referential values are presented in Table III. The MySQL server also stores the learning parameters' initial values. In this study, the primary BRB was developed, and the values of the learning parameters were assigned by considering the domain experts knowledge.

D. Knowledge Base Driver Module

This module acts as an abstraction layer for the database, which helps to support different relational database management systems (RDBMSs), such as MySQL, PostgreSQL, and SQLite. This module maintains the communication between the BRB Main Module and the Knowledge Base Module.

E. Training Module

The Training Module performs learning for BRBES by taking into account the training dataset using the learning mechanism. It receives the training data from the BRB Main Module, which BRB Main Module receives from the Input Module. The training module gets the initial BRB and learning parameters from the Knowledge Base Module through the BRB Main Module. The optimal values of the learning parameters are found using the joint optimisation procedure mentioned in Section IV. There are two modules to perform the joint optimisation procedure. These are the parameter and structure optimisation modules. These modules are used iteratively based on the joint optimisation procedure to find the optimal structure and values of the learning parameters. The near-optimal values of the learning parameters are stored in the Knowledge Base Module.

F. Configuration Module

This module manages various static parameters (e.g. database source, user credentials, and default values of attribute weights and referential values). A configuration file is used to store these parameters. The configuration module supplies the static parameters upon receiving requests from other modules.

G. BRB UI Module

This component facilitates a web-based user interface (UI) for the BRBES to visualize the predicted flood water level. The UI provides options to predict flood water level for a single data point.

VI. RESULTS AND DISCUSSION

Flood is a devastating natural event which brings havoc to human life, destroys urban and rural infrastructure, and disrupts the economy. The various preventive measure helps to reduce these suffering. Therefore, predicting flood helps in taking the necessary steps to minimise the destructive effects of the flood. Our proposed joint optimisation using BRBaDE will help BRBES to predict flood more accurately. The sadar union of Cox's Bazar upazila of Cox's Bazar, a district of Bangladesh, was considered as the case study area to evaluate the accuracy of flood prediction using trained BRBES [12]. During this survey 307 people located at different places of the Cox's Bazar Sadar union were interviewed. They were asked about the flood of July, 2017 of this union by taking into account of the leaf nodes (X22, X23, X19, X20, X21, X17, X18, X14, X15, X25, X26, X27, X28, X29, X30, and X31) of the BRB framework as shown in Fig.1 as well as illustrated in Table. III. The size of this data is adequate, since the sample size is in between 30 and 500, which is considered suitable for most studies [27]. The experts' opinions on the flood water level have been taken into account as the baseline to compare different methods. We have taken experts' advice as the benchmark for comparing among various learning methods. The dataset was grouped into training and test sets with a ratio

TABLE IV
COMPARISON OF AUC OF TRAINED DISJUNCTIVE AND CONJUNCTIVE BRBES, NON-TRAINED BRBES, LSTM, ANN, SVM, AND LINEAR REGRESSION FOR PREDICTED FLOOD WATER LEVEL (X7). CI: 95% CONFIDENCE INTERVAL

Results	Trained Disj. BRBES	Trained Conj. BRBES	Non-trained BRBES	LSTM	ANN	SVM	Linear Regression
AUC	0.86	0.67	0.58	0.61	0.54	0.44	0.40
CI	0.783-0.936	0.558-0.778	0.462-0.693	0.449-0.726	0.425-0.661	0.311-0.572	0.291-0.526

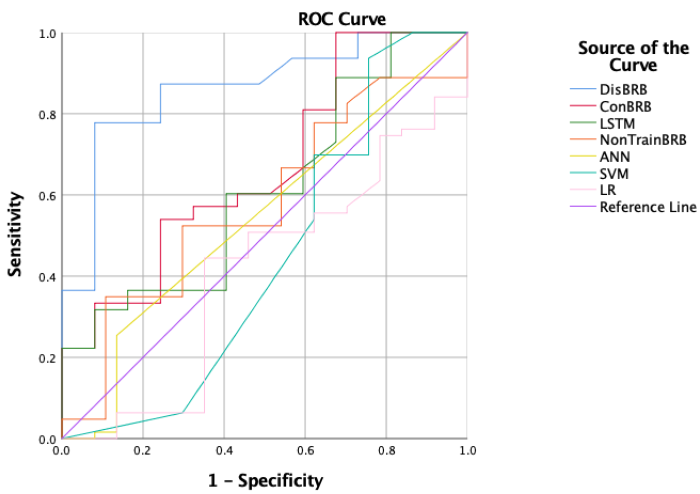

Fig. 7. ROC curves comparison of Non-Trained BRBES, Trained Conjunctive and Disjunctive BRBES, LSTM, ANN, SVM, Linear regression for predicted flood water prediction data.

of 80:20. The cross-validation of five-fold was considered to evaluate the results.

Different evaluation metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Receiver Operator Characteristic (ROC) curves have been used for comparing the performance among different methods. The ROC curves produce a visual representation of the evaluation among various methods. When the Area Under Curve (AUC) achieved by a method is close to one, it is considered to perform better since the highest possible value of AUC is one. Among various machine learning algorithms, LSTM, ANN, SVM based regression, and linear regression can recognize patterns between input and output data. This feature allows them to predict dynamic and non-linear complex natural phenomena, such as flooding, with higher accuracy. Therefore, our proposed learning procedures for multi-level BRBES was compared with the machine learning algorithms, as mentioned earlier for predicting flooding [5][6][16].

Fig. 7 shows the ROC curves for trained Disjunctive and Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression, which are labelled as DisBRB, ConBRB, Non-Trained BRB, ANN, SVM, and LR respectively in the figure. The AUC and the confidence interval (CI) for these algorithms are illustrated in Table IV. Table IV provides the AUCs for trained Disjunctive BRBES, trained Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression as

TABLE V
COMPARISON OF RMSE, MAE OF NON-TRAINED BRBES, TRAINED CONJUNCTIVE AND DISJUNCTIVE BRBES, LSTM, ANN, SVM, LINEAR REGRESSION FOR PREDICTED FLOOD WATER LEVEL (X7).

Results	Trained Disj. BRBES	Trained Conj. BRBES	Non-trained BRBES	LSTM	ANN	SVM	Linear Regression
RMSE	200.01	214.12	215.45	210.47	211.67	220.91	312.54
MAE	100.15	107.07	107.98	105.25	105.67	110.46	157.23

0.86, 0.67, 0.58, 0.61, 0.54, 0.44, and 0.40, respectively. The lower and upper limits of AUC by considering 95% CI, are for trained Disjunctive BRBES, trained Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression 0.783–0.936, 0.558–0.778, 0.462–0.693, 0.449–0.726, 0.425–0.661, 0.311-0.572, and 0.291-0.526, respectively; as shown in Table IV. Therefore, it can be argued that the trained Disjunctive BRBES performs better than trained Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression. The trained Disjunctive BRBES's performance is better not only with respect to AUC but also with respect to both the lower and upper limits having 95% CI. The trained Disjunctive BRBES demonstrates more accuracy than the traditional learning algorithms because it considers various uncertainties. On the other hand, traditional learning algorithms do not address various uncertainty. For this reason, LSTM, ANN, SVM and Linear Regression demonstrate less accuracy, because of their incapability of handling uncertainty.

Furthermore, Table V represents the average RMSE and MAE values for trained Disjunctive and Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression for the test dataset. The RMSE for trained Disjunctive and Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression are 200.01, 214.12, 215.45, 210.47, 211.67, 220.91, and 312.54 respectively. The MAE for trained Disjunctive and Conjunctive BRBES, non-trained BRBES, LSTM, ANN, SVM based regression, and linear regression are 100.15, 107.07,107.98, 105.25, 105.67, 110.46, and 157.23 respectively. From these values, it can be noticed that trained Disjunctive BRBES has predicted the flood water level more accurately than the other techniques. Therefore, it can be concluded that our proposed new learning mechanism helps Disjunctive BRBES to predict flood water levels with higher accuracy than LSTM, ANN, SVM based regression, and linear regression.

VII. CONCLUSION

This research work proposes a new learning mechanism, named joint optimisation using BRBaDE for a multi-level BRB framework. The new learning mechanism allows to learn about the learning parameters and the structure for a multi-level BRB for the training dataset. The proposed joint optimization using BRBaDE has been evaluated and contrasted with other popular algorithms including LSTM, ANN, SVM, and Linear regression. From the results, it can be argued that trained Disjunctive BRBES performs more accurately than the other techniques due to the proposed new learning mechanism, which helps in determining the near-optimal learning parameters and structure of the multi-level BRBES. Henceforth, our proposed learning mechanism will help in more accurate prediction of flooding, which will also help in protecting human life, buildings, roads, and the economy of a country. Nevertheless, the complexity of the learning mechanism increases with a large number of BRBs, which will be considered as future work. Although, the variables for flooding were identified based on our previous research in the future more analytical approach like factor analysis will be used for their identification.

REFERENCES

[1] A. Mosavi, P. Ozturk, and K.-w. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, 2018.

[2] S. Hallegatte, C. Green, R. J. Nicholls, and J. Corfee-Morlot, "Future flood losses in major coastal cities," *Nature climate change*, vol. 3, no. 9, pp. 802–806, 2013.

[3] K. Andersson and M. S. Hossain, "Heterogeneous wireless sensor networks for flood prediction decision support systems," in *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2015, pp. 133–137.

[4] R. U. Islam, M. S. Hossain, and K. Andersson, "A novel anomaly detection algorithm for sensor data under uncertainty," *Soft Computing*, vol. 22, no. 5, pp. 1623–1639, 2018.

[5] V. Seal, A. Raha, S. Maity, S. K. Mitra, A. Mukherjee, and M. K. Naskar, "A simple flood forecasting scheme using wireless sensor networks," *International Journal of Ad hoc, Sensor and Ubiquitous Computing*, vol. 3, no. 1, pp. 45–60, 2012.

[6] V. Sehgal, M. K. Tiwari, and C. Chatterjee, "Wavelet bootstrap multiple linear regression based hybrid modeling for daily river discharge forecasting," *Water resources management*, vol. 28, no. 10, pp. 2793–2811, 2014.

[7] F. Granata, R. Gargano, and G. de Marinis, "Support vector regression for rainfall-runoff modeling in urban drainage: A comparison with the epa's storm water management model," *Water*, vol. 8, no. 3, p. 69, 2016.

[8] J. Sulaiman and S. H. Wahab, "Heavy rainfall forecasting model using artificial neural network for flood prone area," in *IT Convergence and Security 2017*. Springer, 2018, pp. 68–76.

[9] Y. Yang, C. Fu, Y.-W. Chen, D.-L. Xu, and S.-L. Yang, "A belief rule based expert system for predicting consumer preference in new product development," *Knowledge-Based Systems*, vol. 94, pp. 105–113, 2016.

[10] S. Patra and L. Bruzzone, "A batch-mode active learning technique based on multiple uncertainty for SVM classifier," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 497–501, 2012.

[11] K. Andersson and M. S. Hossain, "Smart risk assessment systems using belief-rule-based DSS and WSN technologies," in *2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE)*. IEEE, 2014, pp. 1–5.

[12] R. U. Islam, K. Andersson, and M. S. Hossain, "A web based belief rule based expert system to predict flood," in *17th International Conference on Information Integration and Web-based Applications & Services*. ACM, 2015, pp. 19–26.

[13] J.-B. Yang, J. Liu, J. Wang, G.-P. Liu, and H.-W. Wang, "An optimal learning method for constructing belief rule bases," in *2004 IEEE International Conference on Systems, Man and Cybernetics*, vol. 1. IEEE, 2004, pp. 994–999.

[14] L. Chang, Y. Zhou, J. Jiang, M. Li, and X. Zhang, "Structure learning for belief rule base expert system: A comparative study," *Knowledge-Based Systems*, vol. 39, pp. 159–172, 2013.

[15] L.-L. Chang, Z.-J. Zhou, Y.-W. Chen, T.-J. Liao, Y. Hu, and L.-H. Yang, "Belief rule base structure and parameter joint optimization under disjunctive assumption for nonlinear complex system modeling," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 9, pp. 1542–1554, 2018.

[16] P. Mitra, R. Ray, R. Chatterjee, R. Basu, P. Saha, S. Raha, R. Barman, S. Patra, S. S. Biswas, and S. Saha, "Flood forecasting using internet of things and artificial neural networks," in *2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference*. IEEE, 2016, pp. 1–5.

[17] J.-B. Yang, J. Liu, D.-L. Xu, J. Wang, and H. Wang, "Optimization models for training belief-rule-based systems," *IEEE Transactions on systems, Man, and Cybernetics-part A: Systems and Humans*, vol. 37, no. 4, pp. 569–585, 2007.

[18] D.-L. Xu, J. Liu, J.-B. Yang, G.-P. Liu, J. Wang, I. Jenkinson, and J. Ren, "Inference and learning methodology of belief-rule-based expert system for pipeline leak detection," *Expert Systems with Applications*, vol. 32, no. 1, pp. 103–113, 2007.

[19] L. Chang, Z. Zhou, Y. You, L. Yang, and Z. Zhou, "Belief rule based expert system for classification problems with new rule activation and weight calculation procedures," *Information Sciences*, vol. 336, pp. 75–91, 2016.

[20] M. S. Hossain, S. Rahaman, A.-L. Kor, K. Andersson, and C. Pattinson, "A belief rule based expert system for datacenter PUE prediction under uncertainty," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 2, pp. 140–153, 2017.

[21] Z.-G. Zhou, F. Liu, L.-C. Jiao, Z.-J. Zhou, J.-B. Yang, M.-G. Gong, and X.-P. Zhang, "A bi-level belief rule based decision support system for diagnosis of lymph node metastasis in gastric cancer," *Knowledge-based systems*, vol. 54, pp. 128–136, 2013.

[22] B. Qian, Q.-Q. Wang, R. Hu, Z.-J. Zhou, C.-Q. Yu, and Z.-G. Zhou, "An effective soft computing technology based on belief-rule-base and particle swarm optimization for tipping paper permeability measurement," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 3, pp. 841–850, 2019.

[23] L.-H. Yang, Y.-M. Wang, J. Liu, and L. Martínez, "A joint optimization method on parameter and structure for belief-rule-based systems," *Knowledge-Based Systems*, vol. 142, pp. 220–240, 2018.

[24] J.-B. Yang, Y.-M. Wang, D.-L. Xu, K.-S. Chin, and L. Chatton, "Belief rule-based methodology for mapping consumer preferences and setting product targets," *Expert Systems with Applications*, vol. 39, no. 5, pp. 4749–4759, 2012.

[25] R. U. Islam, X. Ruci, M. S. Hossain, K. Andersson, and A.-L. Kor, "Capacity management of hyperscale data centers using predictive modelling," *Energies*, vol. 12, no. 18, p. 3438, 2019.

[26] M. S. Hossain, S. Rahaman, R. Mustafa, and K. Andersson, "A belief rule-based expert system to assess suspicion of acute coronary syndrome (ACS) under uncertainty," *Soft Computing*, vol. 22, no. 22, pp. 7571–7586, 2018.

[27] N. M. Suki, "Consumer environmental concern and green product purchase in malaysia: structural effects of consumption values," *Journal of Cleaner Production*, vol. 132, pp. 204–214, 2016.