Luleå University of Technology
Department of Computer Science, Electrical and Space Engineering
PERCCOM Master Program

Master's Thesis in
PERvasive Computing & COMmunications
for sustainable development

Md Rafiul Sabbir Hridoy

AN INTELLIGENT FLOOD RISK ASSESSMENT SYSTEM USING BELIEF RULE BASE

2017

Supervisors: Professor Dr. Mohammad Shahadat Hossain - *University of Chittagong*
Assoc. Professor Karl Andersson - *Luleå University of Technology*

Examiners: Professor Eric Rondeau - *University of Lorraine*
Professor Jari Porras - *Lappeeranta University of Technology*
Assoc. Professor Karl Andersson - *Luleå University of Technology*

This thesis is prepared as part of an European Erasmus Mundus program **PERCCOM** - **P**ervasive **C**omputing & **COM**munications for sustainable development.

Co-funded by the Erasmus+ Programme of the European Union

This thesis has been accepted by partner institutions of the consortium (cf. UDL-DAJ, n°1524, 2012 PERCCOM agreement).

Successful defense of this thesis is obligatory for graduation with the following national diplomas:

- Master in Complex Systems Engineering (University of Lorraine);
- Master of Science in Technology (Lappeenranta University of Technology);
- Degree of Master of Science (120 credits) - Major: Computer Science and Engineering; Specialisation: Pervasive Computing and Communications for Sustainable Development (Luleå University of Technology).

ABSTRACT

Luleå University of Technology
Department of Computer Science, Electrical and Space Engineering
PERCCOM Master Program

Md Rafiul Sabbir Hridoy

An Intelligent Flood Risk Assessment System using Belief Rule Base

Master's Thesis - 2017.

95 pages, 20 figures, 20 tables, 1 algorithm and 2 appendices.

Keywords: Belief Rule Base, flood risk assessment, uncertainty, expert systems, optimization, RESTful API

Natural disasters disrupt our daily life and cause many sufferings. Among the various natural disasters, flood is one of the most catastrophic. Assessing flood risk helps to take necessary precautions and can save human lives. The assessment of risk involves various factors which can not be measured with hundred percent certainty. Therefore, the present methods of flood risk assessment can not assess the risk of flooding accurately.

This research rigorously investigates various types of uncertainties associated with the flood risk factors. In addition, a comprehensive study of the present flood risk assessment approaches has been conducted. Belief Rule Base expert systems are widely used to handle various of types of uncertainties. Therefore, this research considers BRBES's approach to develop an expert system to assess the risk of flooding. In addition, to facilitate the learning procedures of BRBES, an optimal learning algorithm has been proposed. The developed BRBES has been applied taking real world case study area, located at Cox's Bazar, Bangladesh. The training data has been collected from the case study area to obtain the trained BRB and to develop the optimal learning model. The BRBES can generate different "What-If" scenarios which enables the analysis of flood risk of an area from various perspectives which makes the system robust and sustainable. This system is said to be intelligent as it has knowledge base, inference engine as well as the learning capability.

ACKNOWLEDGMENTS

Firstly, I would like to convey my gratitude to my supervisors, Professor Dr. Mohammad Shahadat Hossain and Professor Dr. Karl Andersson, for their enormous patient guidance, help, encouragement and inspiration during the tenure of this work. I would also like to thank Raihan Ul Islam, a PhD student in Luleå University of Technology, for his unconditional helps and advices during this work. Along with Professor Shahadat, Professor Karl and Raihan, I engaged in stimulating discussions about the work which has been of significant importance. Their skill, expertise and guidance in the domain, and about research in general has helped me a lot to learn more and become motivated.

I would love to convey my gratitude to the Erasmus Mundus program and PERCCOM committee for giving me the opportunity to participate in this program and complete my study. I would also like to thank all the members and stuffs at University of Lorraine, Lappeenranta University of Technology, ITMO University and Luleå University of Technology. Special thanks to Professor Eric Rondeau, J.P. Georges, Jari Porras and Josef Hallberg for their valuable suggestions and guidance throughout the program.

On personal front, I would love to express my gratitude to my mother and wife for being a constant source of encouragement and support as well as for their belief in whatever I undertake, not only during the thesis but during the two years of my masters program. I would also express my gratitude to all of my friends back home and those whom I met during this program. It has been an wonderful journey with all of you and I am truly lucky to meet these incredibly awesome people.

Skellefteå, May 25, 2017

Md Rafiul Sabbir Hridoy

CONTENTS

List of Figures

List of Tables

List of Algorithms

ABBREVIATIONS AND SYMBOLS

(Alphabetic order)

AI	Artificial Intelligence
API	Application Programming Interface
AUC	Area Under Curve
BDT	Bangladeshi Taka
BFS	Breadth First Search
BPT	Bayesian Probability Theory
BRB	Belief Rule Base
BRBES	Belief Rule Base expert system
CI	Confidence Interval
CRUD	Create, Read, Update and Delete
CSV	Comma Separated Values
DI	Direct Intangible
DST	Dempster-Shafer Theory
DT	Direct Tangible
ER	Evidential Reasoning
ES	Expert Systems
FL	Fuzzy Logic
FLBES	Fuzzy Logic based expert system
FOWM	Federal Office for Water Management
GUI	Graphical User Interface
HTTP	Hypertext Transfer Protocol
HUD	Housing and Urban Development
IAF	Impact Assessment Framework
II	Indirect Intangible
IID	Iterative and Incremental Development
IoT	Internet of Things
IP	Internet Protocol
IT	Indirect Tangible
JSON	JavaScript Object Notation
NFIP	National Flood Insurance Program
RAF	Risk Assessment Framework
REST	Representational State Transfer
ROC	Receiver Operator Curve
SOAP	Simple Object Access Protocol
SQL	Structured Query Language
UDDI	Universal Description, Discovery and Integration
URL	Uniform Resource Locater
USACE	United States Army Corps of Engineers

1 INTRODUCTION

This chapter introduces the research presented in this thesis. It begins with a foreword to give an idea about flooding and briefly presents the motivation and sustainability aspects of this work. The research challenges faced by the author during the thesis work along with aims and objectives are also presented in this chapter. In addition, the scope of the research and the contributions are also discussed. The chapter concludes with an outline of this thesis.

1.1 Foreword

Even in this twenty first century, in the era of significant innovations and technological enhancements, humans are helpless in the hand of natural disasters such as floods, earthquakes, volcanic eruptions etc. Flooding is one of the most devastating catastrophic disasters among all and it causes significant socio-economic losses every year all over the world. It has been noticed that floods are responsible for one-third of all deaths, one-third of all damages and one-third of all injuries from natural disasters [1].

For example, one of the most devastating floods occurred in China during 1941 caused by overflowing of water from Huang He River (Yellow River). It caused the complete inundation of eighty-eight thousand square km of land, four million loss of life and eighty million people were homeless [2]. Since 1990, floods have caused more than 10,000 deaths, and economic losses are greater than US 70 billion in the United States alone [3]. In Europe, during 1971-1995, there were 157 major floods, and the cost of damages in the period of 1991-1995 was estimated as Euro 99 billion by European Environment Agency [4]. In Bangladesh, floods caused serious threat to lives in 1954, 1955, 1974, 1984, 1987, 1988 and 2004 [5]. Recently in 2015, a destructive flood in India, resulted from heavy rainfall, caused the death of 500 people and over 1.8 million people were displaced from the southern part of India with estimated damage ranging from US 3 billion to over US 15 billion.

Since flood is devastating and causes enormous damage, assessing the risk of flooding is very important. It will help to take necessary steps, enabling to save millions of life and hence, losses can be reduced significantly. Several factors such as transportation, agriculture, road network and mental health should be taken into consideration while assessing the risk of flooding. Some of these factors are quantitative in nature while others are qualitative. For example, financial loss can be measured in a quantitative way while health condition should be expressed in a qualitative way. Therefore, various types of uncertainty such as vagueness, imprecision, ambiguity, ignorance and incompleteness can be noticed while measuring these factors. A belief rule based expert system (BRBES) methodology can be employed which has the capability to process heterogeneous as well as data with various types of uncertainties in a single integrated framework to assess the consequences or damages of flooding in an area.

Knowledge base, inference engine and user interface are the most essential components of an expert system. Moreover, different types of uncertainties such as ignorance, vagueness, ambiguity, incompleteness and imprecision can be associated with the dat5 a as pointed out before. There are several reasons for causing these uncertainties such as lack of human knowledge, insufficient data or faulty sensors. Therefore, it is necessary to consider uncertainty along with heterogeneity while designing an expert system. BRBESs methodology use belief rule base to represent uncertain knowledge, while Evidential Reasoning works as the inference engine with the capability of handling both heterogeneous and uncertain data [6].

This research is significant as it helps to build an unified framework considering heterogeneous data for flood risk assessment and their associated uncertainties. Establishing a web based BRB expert system also enables the widespread use of the system for different domains. Eventually, the expert system is also implemented in real case study area and performance analysis is conducted. The procreation of such expert system enables companies and academia to come up with several interesting real time systems such as disease diagnosis, disaster recovery system like assessment of risk of flooding, which is the use case for this research.

The purpose of this research work falls directly under the scope of ICT for sustainability. The expert system helps to make possible applications like flood risk assessment. Also, using web based BRB expert system helps to solve the problem of interoperability, which is one of the main issues taken into account now-a-days while developing real time systems. Moreover, a well established expert system for flood risk assessment is an urgent need for the decision makers today, and this research work implements a flood risk assessment system with the facility of multi-level analysis capability based on this. The use case selected for flood risk assessment is considered with sustainability in its core ideology and it favors all three pillars of sustainability, namely people, planet and profit, as shown in Fig. 1. There is a three-level sustainability impact of this work.

1. Reduction of environmental impacts such as damage to natural resources by providing valuable insights about different sub-scenarios so that decision makers can take necessary steps.

2. The application of flood risk assessment to serve as a risk assessment framework causing invaluable savings to resources, property and most importantly, human life.

3. Usage of web based BRB expert system to foster interoperability of the system and therefore reduce the effort of design and deployment of expert system.

Figure 1. Sustainability aspects of this work

The research presented in this thesis uses BRBESs methodology to develop flood risk assessment BRBES. This BRBES developed using open source tools and technologies and hence, ensures the interoperability overcoming the integration problems which are common in today's risk assessment systems. The use case study helps to understand the performance of the system in research domain and learning module allows to validate the expert system.

1.2 Research Challenges

The current experts systems for flood risk assessment consist of a plethora of methods and tools which only consider quantitative data and thus social and psychological aspects, which can not be expressed in terms of quantitative manner, are often ignored. This problem of integrity in terms of heterogeneous data is one of the key obstacles this research tries to address. Moreover, finding different factors related to the risk of flooding is difficult and uncertainties associated with these factors are mostly overlooked while building expert system which this research also tries to address. In addition, there are several challenges associated with the learning methodology of an expert system in variable conditions. This research work tries to address these challenges by using a BRB expert system along with a learning methodology and conducting proof of concept implementation as well as performance evaluation of the system.

1.3 Research Objectives and Questions

This research aims on the identification of factors of flood risk as well as their associated uncertainties. Eventually, it focuses on the development of a novel flood risk assessment system to help authorities and decision makers to evaluate different aspects of flood risk. Following objectives are identified to achieve the aims of this research.

1. Identification of the factors causing flood and exposed elements to the risk of flooding.
2. Find out the uncertainties associated with the factors of flood risk assessment.
3. Develop a web based BRBES for flood risk assessment along with an API which can be accessible for public use.
4. Fine tuning the system by using data collected from Bangladesh case study areas and developing the learning model to validate the system.

From the above mentioned objectives, the following research questions are identified:

1. What are the heterogeneous factors associated with flood risk assessment?
2. Which type of uncertainties these factors are associated with?
3. How to build an expert system which considers these factors along with their associated uncertainties?
4. How learning capability can be integrated in the expert system to make it intelligent?

This thesis aims to give the answers of these research questions using a research methodology.

1.4 Deliverables

This research studies and reviews the existing systems used for flood risk assessment and propose a web based BRBES to assess the flood risk taking uncertainties into consideration. The system is build up and tested with data taken from survey. A dynamic BRB tree traversal algorithm is introduced which can traverse any Belief Rule Base tree as well as a RESTful API is built by which researchers, who want to use the BRB algorithm without writing it from scratch, can evaluate the expert system result from their data. Optimal learning model for Belief Rule Base is also an integral part of this system which is used for optimizing the parameters to get better results. This feature makes the expert system intelligent as the system has knowledge base, inference engine and learning capability. This work aims to contribute in a novel way, by providing a combination of both web technologies and expert system with the ability of handling various types of uncertainties which could help potential researchers to use a fully functional high performance system. Decision makers can also get an overview of the flood assessment system which can help them to take necessary steps before the flooding and hence, can reduce the effect of flood in some ways such as saving human life. Such an approach of computing will open a new avenue of research agenda which we would like to call 'Sustainable Computing'.

1.5 Scopes and Delimitations

There are numerous fields where risk assessment expert systems can be utilized but this research work focuses on flood risk assessment. Experiments in different field with relevant data can be done in near future. Although experiment with different ways to implement the expert system to reduce energy, memory and resource consumption is an important concern for sustainability, it is also outside the ambit of this work and is the basis of future work. The flood risk assessment application is built using BRB based inference methodology, but its comparison with other similar methodologies is also beyond the scope of this work. Considering the limited time and resources, it was not possible to compare and benchmark between different methodologies of flood risk assessment in this research, instead it gives a detailed description of the expert system built in this research, which can be a basis of future work as well.

1.6 Organization of the Thesis

The rest of the thesis is organized as follows:

Chapter 2 - Background and Related Work

This chapter talks about flooding, flood risk assessment, different types of factors related to risk assessment and their associated uncertainties. The state of the art flood risk assessment expert systems are also described with real world scenarios followed by different risk assessment mechanisms.

Chapter 3 - Framework Design and System Implementation

Chapter 3 describes the methodology used for building the flood risk assessment system. It conveys explanations of the system architecture as well as various tools and technologies involved with the framework design. This chapter gives detailed explanation of the BRBES for assessing the risk of flooding along with it's different components. It also introduces a novel BRB tree traversal algorithm, JSON based data preparation, RESTful API and the web interface for the expert system.

Chapter 4 - Use Case: Flood Affected Neighborhood

This chapter presents an use case where the risk assessment framework has been implemented. It describes the method of surveying people from the flood affected neighborhood for collecting data. It also explains the way of preparing, validating and optimizing data collected from survey. The chapter ends with the system validation using real data along with the necessary results.

Chapter 5 - Discussion

This chapter presents the discussion of the risk assessment framework for analyzing and reducing the risk of flooding. It highlights the contribution and novelty of this research in different steps followed in previous chapters. It also tells about the sustainability aspect of this research.

Chapter 6 - Conclusion and Future Work

This chapter brings out the conclusions drawn from the result evaluated by using the expert system. Finally, it points out the limitations of this research and future directions of this work are hypothesized.

2 BACKGROUND AND RELATED WORK

This chapter presents the nature of flooding in terms of sufferings that it can bring by taking account of the flood data of the various parts of the world. It then introduces the concept of risk in general and flood risk in particular. This is followed by the discussion on the flood risk assessment framework. In the context of this framework, an investigation on the various consequences of flooding is presented. The uncertainty issue that need to be considered in measuring these consequences are also elaborated. A comprehensive literature review on flood risk assessment is also presented in this chapter.

2.1 Flooding

Flooding is a general and temporary condition of partial or complete inundation of dry land areas due to the overflow of inland or tidal water from the unusual and rapid accumulation [2]. Due to the rapid climate change and population growth, the risks of floods are expected to increase all over the world. Gradual migration of coastal areas is also a reason of increasing floods and flood risk. Flooding may occur as an overflow of water from river, lake, or ocean, or it may occur due to an accumulation of rainwater on saturated ground in an areal flood. The size of water sources will vary with seasonal changes in rainfall and snow melt and hence, it is unlikely to consider these changes significant unless they damage property or drown cattle or human.

There are two major dimensions of flood risk: flood depth and areal extension. Floods, such as flash floods, can develop in just a few minutes and without any visible sign or rain while some floods develop slowly over the time. Moreover, floods can be local which means the risk will be only on a community or neighborhood, or very large, affecting entire river basins.

Several devastating floods have been occurred during last century over the world. Some of the most notable floods in different regions along with the cause and aftermaths are described below.

Floods in USA

In 1948, the flood exceeded the capacity of river in Columbia, USA caused by the heavy rainfall in winter even with the presence of an extensive flood control system [7]. This flood mostly damaged a city named Vanport, which is close to the city of Portland, Oregon and situated on the floodplain of Columbia River, USA. It caused a financial damage of US 100 million. Twenty thousand people were affected from Vanport area while thirty died. This flood forced the authority to rebuilt almost the whole city as well as force migration of people from different part of USA.

During the summer of 1993, due to the over-flooding of the upper Mississippi River, nine Midwestern states; Iowa, South Dakota, Missouri, Illinois, Nebraska, Kansas, Minnesota, North Dakota and Wisconsin; experienced major damage [7]. This flood caused severe loss due to the fail of hundreds of levees built to protect the land. More than 75 towns along with millions of farmlands drowned under the water and approximately 50 people died. Thousands of people were temporarily or permanently evacuated from the area as thousands of homes were completely destroyed. The infrastructures and businesses were seriously affected and the economic damage was estimated around US 15 million to US 20 million [8]. This flood was so devastating that after the flood, a big reconstruction plan was taken by the government which included making an entire town to higher ground, rise up the foundations of buildings and change the building materials as well as the interior design which can handle further floods, deploy a new levee system to protect against floods [9].

Floods in Europe

In 1953, a huge sea storm created from the high spring tides combined with the strong winds destrcuted the southwestern quarter of the Netherlands and almost all the province of Zeeland was flooded [10]. Zeeland is the southwestern most province of Netherlands which is cut by three different river deltas (the Rhine, the Mass and the Schelde) and is mostly an agricultural area. The flood killed 1835 people and 72000 people were evacuated from Zeeland. Thousands of building were destroyed, over 200,000 heads of the livestock and large parts of farmlands drowned with the estimated loss of around 1.5 to 2.0 billion Dutch guilder (US 0.8 - 1.1 billion).

Floods in Asia

A devastating flood was occurred in China in 1998 due to the above average rainfall in the northern region of China for several months prior and during the summer of 1998. This heavy rainfall caused the over-flooding of Ynagtze River (also known as Changjiang River). Ynagtze is the third-longest river in the world, operating in the Quinghai-Tibet plain to the East China Sea. This river basin covers around 800,000 square miles, which is nearly one-fifth of China's territory and is the home for approximately 400 million people. The river is the sixth largest river in terms of discharge volume in the world which discharges about 34 trillion cubic feet which is 37 percent of total surface discharge in China annually. It was considered the worst Northern China flood in 40 years which resulted in 3,704 death, 15 million homeless and US 24 billion in economic loss. A staggering 100,000 square kilometers (25,000,000 acres) were evacuated, 13.3 million houses were damaged or destroyed [11]. 'Resettlement in the Stricken Areas Project' was introduced by the Chinese government with an estimated cost of RMB 10.1 billion (US 1.22 billion) which included family relocation, transforming agricultural lands into forest, destroying barriers to create floodplain etc.

The worst flood in the recent history of Bangladesh occurred in 1988 caused by the cyclone struck in several coastal districts of Bangladesh such as Bagerhat, Barguna, Bhola, Jessore, Khulna, Patuakhali and Stakhira, as well as the Sundarbans which submerged two-thirds of the country. Bangladesh is an alluvial plain delta created by three rivers; the Ganges, the Jamuna, and the Meghna. Increased upland flow of the rivers and the high tides in the Bay of Bengal during the rainy season results in an annual inundation affecting most of the 55,000 square-mile area of the country [12]. This flood was responsible for the death of 5708 people, injuries of millions of people, drowning of over 33,000 head of the livestock and damage of over 174,000 hectares (430,000 acres) of rice harvest land, mostly in the coastal area. The gross weight of crop losses was estimated in 200,000 tonnes (220,000 tons), which was accounted as 70% of Bangladeshi crops that were ready to harvest [13]. Due to the wide scope of the disaster, Bangladesh had got international aid from Japan, Canada, Netherlands and United Kingdom. Relief operations were conducted by Bangladesh Army as well as several non-government organizations in both air and water. To coordinate relief and rehabilitation efforts, a national disaster committee was composed of relief specialists. The whole disaster alert system was changed to minimize the effect of future disasters in coastal areas of Bangladesh.

Even in Twenty-first century, flood is causing lots of damage to life, livelihood, economy and hence, assessing the risk of flooding is really important. If it is possible to assess the risk of flood, it can save so many lives as well as natural resources. This enables the possibility of saving millions of dollars of damage caused by flood every year all over the world which tends to a sustainable society and hence, environment. Saving millions of dollars also ensures better management of financial resources and provides a sustainable financial system for the government.

2.2 Flood Risk Assessment

Risk is the potential of losing something of value such as physical health, social status, emotional well-being or financial wealth [14]. Risk can be termed as "what may go wrong". It can also be defined as the probability of occurrence of an event and its consequences. Alternatively, it can be defined as the combination of likelihood of event and its consequence as elaborated by Eq. 1 [15].

$$R = f(P(E), C) \tag{1}$$

where R is the risk, $P(E)$ is the probability of occurrence of an event E and C is the consequence of that event.

Risk can be alternatively defined as the probability of interaction or intersection or multiplication or integration between an event and exposed elements to it as elaborated in Eq. 2.

$$R = P(Event \wedge Consequence) \tag{2}$$

where R is the risk, P is the probability of occurrence of an event E and C is the consequence to the exposed elements of event E.

Therefore, in view of Eq. 1 and 2 flood risk can be defined as the product of the probability of flooding and the consequences caused by it [16]. It is possible to express flood risk mathematically using Eq. 3. Flood risk can also be defined as the degree of interaction between various dimensions (depth, areal dimension) of flood and exposed socio-economic elements to the flooding.

$$f(R) = P_F * C \tag{3}$$

where P_F is the probability of occurrence of flooding and C is the consequences of exposed elements of flooding.

Risk assessment is a scientific approach to deal with risks by possible accidental losses and designing and implementing procedures that minimize the occurrence of loss that do occur [17]. Flood risk assessment is an approach to measure flood risk using different factors and taking their associated uncertainties in concern. In order to provide a complete flood risk assessment model, existing methods of statistical techniques are used to assess flood risk in this research rather presenting any new approach. This research is particularly relevant as there are large global initiatives which approach to understand flood propagation in urban areas and its consequences in more details.

Flood risk assessment is related to the determination or measurement or evaluation of the intensity level of risk. This can be achieved if risk (which is the combination of likelihood of event and its consequence to the exposed elements) can be identified. So risk identification consists of identification of the factors those are responsible for the event as well as socio-economic objects exposed to the event. Only identification of risk is not enough since it is related to the factors causing events and the objects those are experiencing the consequences or damage. In this context, the determination of risk is necessary. This can be done either in a qualitative way or in a quantitative way. Finally, it needs to be evaluated to see the acceptance level of risk. From this context, the process of risk assessment consists of three steps including risk identification, risk determination and risk evaluation.

A risk assessment framework (RAF) is an approach for effectively identifying and assessing the causes of risk [18]. A RAF consists of three steps of risk assessment, as mentioned above, is illustrated in Fig. 2.

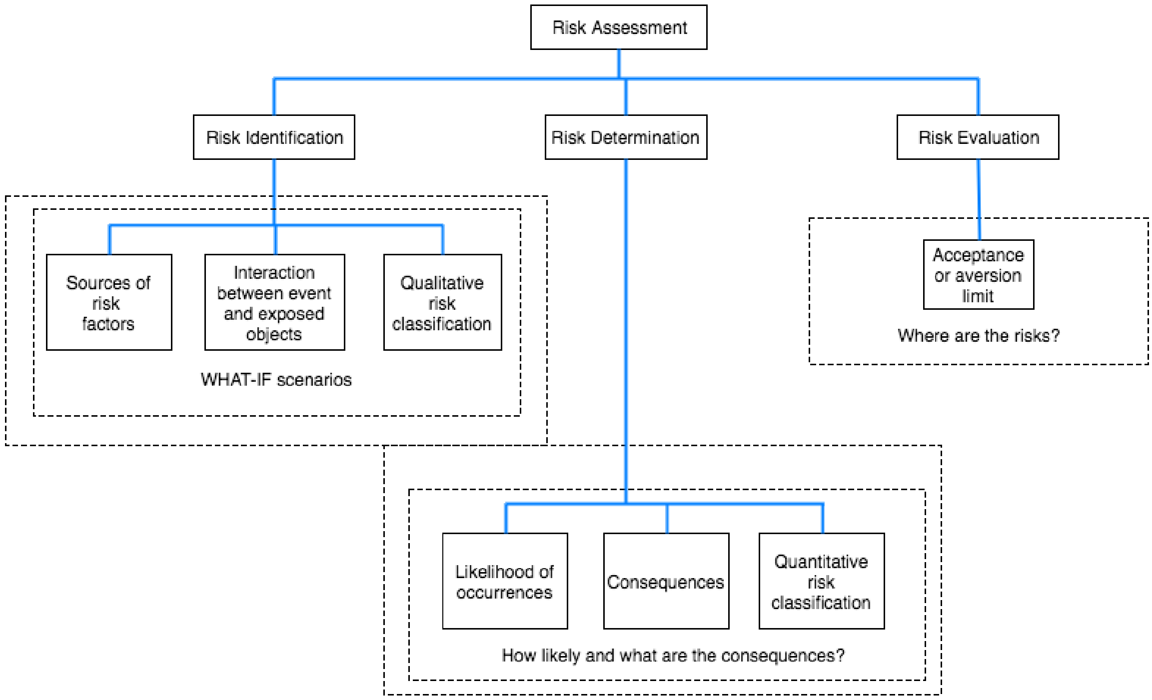


Figure 2. Risk Assessment Framework

By taking account of the risk assessment framework as illustrated in Fig. 2, a RAF for flooding is elaborated in Fig. 3, which consists of three steps.

1. Risk Identification: First step of flood risk assessment is to identify the causes of the flood risk and the exposed elements to the flooding.

2. Risk Determination: After identifying the factors of flood risk, next step is to determine the consequences of flooding. Consequences will lead us to understand the flood risk pattern as well as will be needed to collect necessary data.

3. Risk Evaluation: Final step of the risk assessment is concerned with identifying the acceptance limit of risk. This step gives the result of assessed risk.

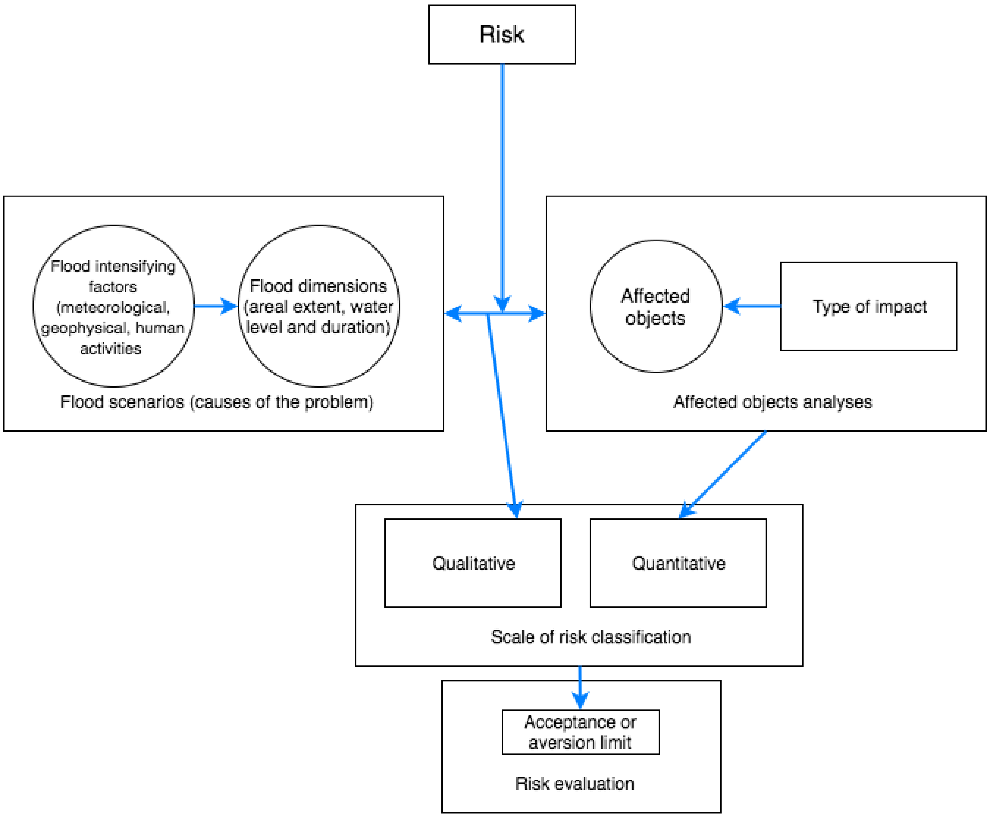


Figure 3. Flood Risk Assessment Framework

This research assesses the risk from flooding taking different types of factors into consideration.

2.2.1 Factors of Flood Risk Assessment

First step of the flood risk assessment is to figure out the factors causing flood as shown in Fig. 2 and 3. There are several factors which can be taken in consideration while assessing flood risk. To identify the factors of flood risk and to divide them in several categories for distinguishing, a model is followed. This model can classify the factors into four categories based on the degree of monetization, the degree of physical contact of the flood damages as well as the social impacts [19]. The factors are named as Direct Tangible, Direct Intangible, Indirect Tangible and Indirect Intangible as shown in four quadrants of Fig. 4. This model gives the domains (quadrants) by which factors can be selected and divided into categories and hence, this will provide a well organized structure for research.

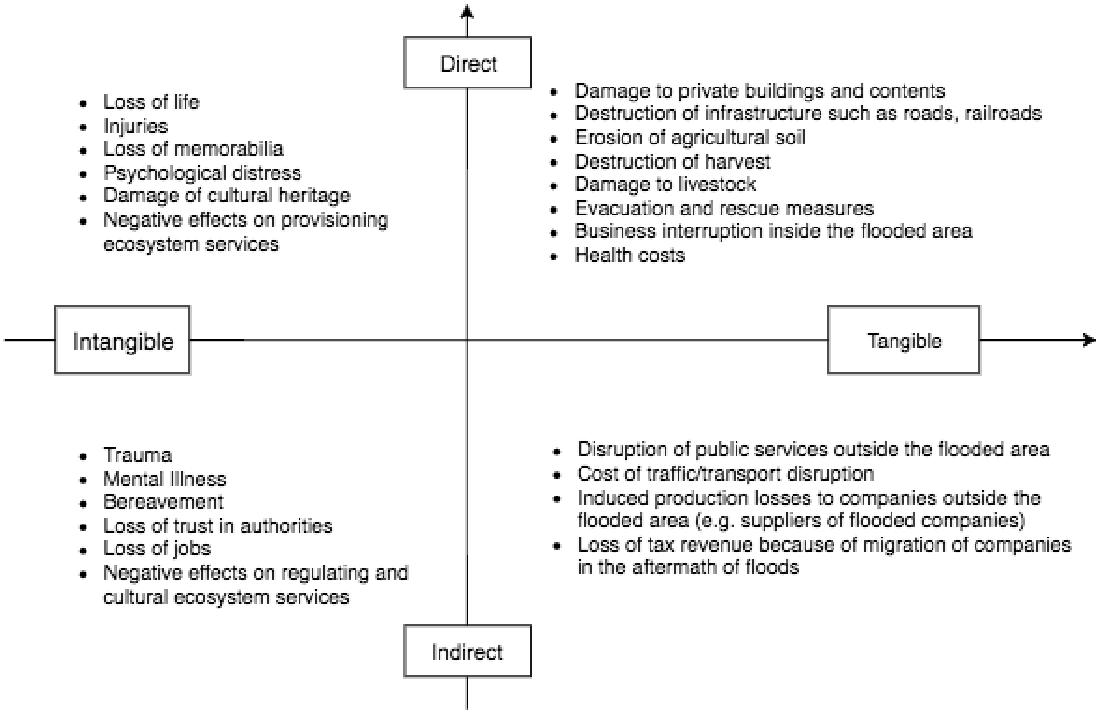


Figure 4. Damage categories of flood

2.2.2 Direct Tangible

Direct Tangible (DT) factors are the those which are directly affecting the flood prawned area and tangible in nature. Several factors can directly affect and cause damage due to flood such as, damage of buildings, properties and harvest, eradication of infrastructures such as roads, railroads, erosion of agricultural soil etc.

From the examples of flooding provided in section 2.2.1; it is easily noticeable that some of the common risks due to the flood was the damage of properties, destruction of infrastructure such as roads, houses, destruction of cultivable lands, evacuation and rescue actions caused by the flood. All of them can be considered as DT factors. Factors such as flood affected area, crop, water level, availability of cattle food, accommodation, availability of transport, road network are considered as DT factors.

2.2.3 Direct Intangible

Direct Intangible(DI) factors are those which are directly affecting the flood prawned area but intangible in nature. Example of such factors can be loss of life, damage to cultural heritage, loss of memorabilia etc.

The flood in Bangladesh in 1988 caused serious damage to life, nature and the social condition was worst after the flood. There were several injuries as well as death of human and cattle due to the flood. Epidemic diseases also spread out by water after the flood. Percentage of loss of cattle and social condition are two DI factors which have been considered for this work.

2.2.4 Indirect Tangible

Factors which are indirectly affecting the flood affected area and tangible in nature are said to be Indirect Tangible (IT). Though these factors are causing indirect effects, they are important to consider in terms of social, economical and psychological point of view. Cost of traffic or transport system, disruption of public services, loss of tax revenue from the companies due to floods etc. are some of the IT factors.

Indirect tangible factors are mostly those which do not cause direct damage but are important to consider due to the social and economical impacts of them. As per the examples given above, there might be disruption of public services such as transport, postal system etc. after the flood. These factors are not generally considered while measuring the flood risk as they are not directly affecting the flood but this research is taking them into consideration. For this case, transportation problem and availability of stuffs are considered at IT factors.

2.2.5 Indirect Intangible

Factors that are indirectly affecting the flood affected area and intangible in nature are called Indirect Intangible (II). Factors such as trauma, mental diseases, loss of jobs, loss of trust in authorities are examples of II factors.

Floods in developing countries like Bangladesh, India, China cause severe effects on the people as well as in the economy. Due to having large number of population living under poverty line, flood causes serious financial and social crisis in these areas. So many people become homeless, jobless and hence, they come in a situation of financial crisis which leads them to mental instability and causes serious issues such as mental illness, violence etc. Taking these socio-economic aspects in concern, financial condition and mental condition are the two II factors which has been taken in concern for this research.

While selecting the factors, it is important to keep in mind that efficiently choosing the factors is mandatory to get meaningful result for flood risk assessment. Availability of data is also an important issue and should be taken into consideration while developing the risk assessment model.

Different types of factors associated with flood risk assessment has been described in this subsection which gives the answer of **research question 1** as mentioned in section 1.3.

2.3 Uncertainties

There are several uncertainties associated with factors of flooding and hence, the next step is to identify the uncertainties associated with each factor mentioned in previous section.

Uncertainty is an unpredictable, and uncontrolled outcome while risk is a consequence of action taken in spite of uncertainty [20]. There are different types of uncertainties associated to each factor which has been identified by the risk factors classification model such as ignorance, incompleteness, ambiguity, imprecision and vagueness [21]. It is important to find out related uncertainties of the factors to get meaningful results for risk assessment avoiding inaccuracy.

Identification of uncertainties in flood risk assessment is important as uncertainties can cause inaccuracy. Some uncertainties of different factors were identified by literature review while others are found from the interview and questionnaires during the field visit at Bakkhali, Cox's Bazar. Some of these factors were expressed in qualitative term while others were expressed as quantitative. Table 1 illustrates all the factors of flood risk assessment as well as the uncertainties associated with each factor. Table 1 also gives an operational definition of each factor has been obtained.

While interviewing people from the case study area, DI factors such as "Percentage

Table 1. Factors of flood risk assessment and their related uncertainties

Factors	Uncertainty Type	Discussion
Direct Intangible		
Percentage Loss of Cattle	Imprecision, Incompleteness	Percentage of cattle were died during the flood.
Social Condition	Vagueness, Ignorance	Refers to the social condition of flood affected area.
Direct Tangible		
Area	Imprecision	Area of the case study.
Water Level	Incompleteness	Refers to the flood water level.
Availability of Cattle Food	Ignorance, Incompleteness	Food available for cattle after the flood.
Accommodation Problem	Vagueness	Problem with accommodation in flood affected area.
Availability of Transport	Incompleteness, Imprecision	Refers to the availability of transportation in flood affected area.
Length of Road Affected	Incompleteness	Length of the road network affected by flooding.
Road Damage	Imprecision	Percentage of road damage due to flood.
Duration of Standing Water	Incompleteness	How long the flood water is standing in the area.
Amount of Crop	Incompleteness	Amount of crop produced in the flood affected area.
Fertility	Imprecision	How fertile the lands are after the flood.
Availability of Labor	Vagueness, Incompleteness	How many labors are available after the flood.
Cost of Raw Materials	Imprecision	Increase of cost of raw materials due to flood.
Agricultural Wages	Incompleteness, Imprecision	What is the wage of agricultural workers.
Indirect Intangible		
Financial Condition	Vagueness	Financial condition of the flood affected people.
Mental Condition	Vagueness	Mental condition of people of the flood affected area.
Indirect Tangible		
Availability of Stuffs	Incompleteness, Ignorance	Stuffs available for work after the flood.
Cost of Transport	Vagueness, Incompleteness	Change in the cost of transportation due to flood.
Frequency of Travelers	Inconsistency, Incompleteness	Refers to the frequency of traveler.
Transportation of Goods	Vagueness, Incompleteness	Condition of the transportation of goods in flood affected area.
Transportation Delay	Ignorance	Delay in transportation due to flood water.

of Loss of cattle" and "Social Condition" had quantitative and qualitative data respectively. 30% people expressed the percentage of loss of cattle in range 30-35%, another 30% people answered it in the range 10-20% while 20% answer was 5-10% and other 10% had no idea about the perctange and hence, answered 'no idea'. Because of the variety in data, uncertainty due to incompleteness and imprecision arose for this factor. "Social Condition" qualitative result which was expressed in terms of high, medium and low and the answer varied in a wide range as well. This drives the data to uncertainties due to vagueness and ignorance.

II factors such as "Financial Condition" and "Mental Condition" were expressed in terms of qualitative data in a range high, medium, low. Data for "Financial Condition" was pretty stable as 96% people answered that "Financial condition" was "low" while other 4% was mentioned it as "medium". Data for "Mental Condition" was almost in the same pattern which caused uncertainty due to vagueness as "high", "medium" and "low" can be vague in pattern. The definition and weight of these terms may vary from person to person.

For IT factor such as "Availability of Stuff", data was represented as percentage and it varied in a very wide range. 35% people said the availability of stuff was around 50% while 30% people said it was only 10-15% and others said that almost 100% stuffs were moved away from the flood affected area which is not a normal case. This tends the data to have uncertainty due to incompleteness and ignorance as the data does not give a solid base for understanding the situation because of being incomplete as well as some people might have ignored this factor while answering. Another IT factor named "Frequency of Travelers" also gave data in a wide range instead of any solid insight. 30% people said there were about 2000-4000 people traveling during a day in the flood affected area while 30-32% people said the number was only 300-400 which drives this factor to uncertainty due to inconsistency and incompleteness.

In case of DT factors such as "Area" and "Fertility", due to the difference of expression in human knowledge, it was difficult to get a clear picture of the data. Some people said the "Area" in acres unit while others expressed it in kilometers or hector unit. For "Fertility", there were different ways data had been recorded. Some people expressed this factor in numerical term while others expressed it in subjective term such as "Very Fertile", "Fairly Fertile" and "Not Fertile" which caused the uncertainty due to imprecision. Same case happened for the factors "Cost of Raw Materials", "Road Damage" and hence, uncertainty due to imprecision occurred. Data for "Water Level", "Length of Road Effected", "Duration of Standing Water" and "Amount of Crops" were all quantitative in nature and they were in a wide range which caused uncertainty due to incompleteness. Data for "Agricultural Wages" and "Availability of Cattle Food" were quantitative and qualitative in nature, respectively. They caused uncertainty due to incompleteness and imprecision as they were very discrete in nature. Moreover, they did not give a solid base for reasoning the situation.

Uncertainties associated with different factors of flood risk assessment has been identified in this subsection. In a way, it gives the answer of the **research question 2** as mentioned in section 1.3.

2.4 Risk Assessment Methods

There are different existing methods which are used by various research groups and organizations throughout the world for accessing flood risk. Six widely used flood risk assessment methods are described below.

NFIP Hydrologic Method

The National Flood Insurance Program (NFIP) uses a hydrological method developed in the 1960's by U.S Department of Housing and Urban Development(HUD) of USA to asses flood risk for insurance purposes [22]. This method derives a water surface elevation probabilistic function to evaluate the flood risk in a geographical area. It transforms the probabilistic function as a function of depth inundation using a model of damage. This probability function is then integrated to compute the average annual loss.

NFIP also has a more comprehensive model for risk analysis that can evaluate site specific probabilities to represent risk. It takes the performance and reliability of flood protection as well as the effect of their failure on flooding into consideration. [23].

USACE Method

U.S. Army Corps of Engineers (USACE) follows the analysis procedures from Engineering Manual 1110-2-1619 to assess flood risk [24]. Methods from the engineering manual are used in flood damage reduction analysis software. It generates the results of economic analyses to assess flood risk. Compared to the approach of NFIP, this approach provides additional features such as site-specific flood hazard result and levee fragility.

Catastrophe Models

Many organizations and insurance companies assess the risk of flood including other catastrophic disasters such as earthquake, cyclone using catastrophe models [25]. These models are generally developed for large or geographically diverse areas. It is possible to assess individual or combined risk using these models. In general, catastrophic models include the following risk assessment components:

1. A probabilistic scenario of flood hazard.
2. How hazard is modified by mitigation and management measures which represents the likelihood of success or failure of the results.

3. A mathematical approach of exposed elements.

4. A model of the vulnerability caused by those elements to the hazard.

;

Comprehensive Risk Assessment Method

North Carolina Floodplain Risk Information System is another system which is used by North Carolina state authority to assess certain components of flood risk [26]. They have developed site-specific water surface elevation probability functions using hydrological and hydraulic studies to assess flood risk. This system can compute the annual probability of flood using the information collected from building footprints from the state. It provides a dynamic and accurate statewide mapping of floods.

As European floods are different in nature compared to the floods in North America and Asia, different approach is used in Europe for flood risk assessment.

Gillard and Givone Method

A flood risk assessment method, created by Gillard and Givone in late 90's [27], which creates land use map and price map to assess the vulnerability based on those maps is used for assessing flood risk, mainly in Europe. This method was used to assess the flood risk for different rivers across Europe for example the Turiec River, Slovakia [28]. It access the vulnerability as a product of weight and price of the land.

FOWM Method

Another flood risk assessment framework was used for assessing flood risk in different parts of Europe which was originally originated from Federal Office for Water Management (FOWM), Bern, Switzerland [29]. This framework is based on a risk matrix in combination with principles based on expressing the maximum acceptable risk. For example, flood risk in Rafina sub-urban area in Greece has been evaluated using FOWM method.

All six of the risk assessment methods for flooding described above address the different components of flood risk using methods tailored to their needs. Some methods such as the USACE and the comprehensive risk assessment methods explicitly consider uncertainty occurred due to levee performance but not all the uncertainties mentioned earlier while the NFIP method, Gillard method and FOWM method do not take any uncertainty of data in concern. Moreover, all these methods mainly relies upon quantitative measures but do not consider qualitative data which also have influence on the likelihood of flood occurrence.

Intelligent Risk Assessment Methods

There are several numerical and symbolic methods for representing and reasoning with uncertainty such as Fuzzy Logic (FL), Bayesian probability theory (BPT), MYCIN/EMYCIN and Dempster-Shafer Theory of evidence (DST) [30] [31]. Each of these frameworks has their own features and can be applicable in special application environment to handle uncertainties. For example, in FL, the procedure of constructing membership function it is not clear which often causes problems. Uncertainties due to ambiguity, vagueness and imprecision can be addressed by FL but it does not consider uncertainties due to incompleteness and ignorance which are also considered in flood risk assessment. In BPT, exponential number of prior probability is required which leads to computational complexity. It also does not consider the uncertainties, for example uncertainties related to flood risk. In case of MYCIN/EMYCIN, the combining functions area ad hoc. Moreover, the calculi does not perform well, due to that nature of the combining function, if the combination of several evidence is considered which is a case for this research. Both MYCIN/EMYCIN and DST assume that various pieces of evidence are independent of each other which is not appropriate for flood risk assessment. DST can handle uncertainty due to ignorance but not the others which may lead to inappropriate result for flood risk assessment. [30].

In real systems, all the uncertainties mentioned earlier may coexist and hence, in order to handle all of them, a mathematical model is needed. Moreover, attributes involved in the system can be heterogeneous in nature. All the above mentioned methods, used for the representation and reasoning with uncertainty, can only handle quantitative data, not the qualitative data which should also be considered while building an expert system for flood risk assessment. Hence, it is necessary for this research to use a hybrid framework that can handle heterogeneous input data along with uncertainties associated with each of them.

An expert system can solve this problem. Expert systems (ES) are branch of applied Artificial Intelligence (AI) which were developed in mid-1960s. The idea of ES is that simply expertise, which is the vast body of domain knowledge, is transfered from human to computer [32]. This domain knowledge is then stored to make inferences and derive a specific conclusion based on the specific advices given by the users. To develop such expert systems, a knowledge representation schema which can handle different factors of flood risk, is required. Additionally, the expert system should have the inference engine which can handle all the uncertainties mentioned above for the factors of flood risk assessment. As inference mechanisms such as forward chaining and backward chaining can not handle uncertainties, the uncertain data that exist in the evaluation of flood risk needs to be processed by using a refined knowledge representation schema along with an inference mechanism. For the design and development of flood risk assessment expert system, the employment of Belief Rule Base(BRB) inference methodology using the evidential reasoning approach (RIMER) is considered [6]. RIMER consists of two main parts: 1) belief rule base which is the knowledge representation schema and 2) evidential reasoning (ER) which is used as an inference mechanism.

2.5 Summary

This chapter has given a clear idea about flooding and flood risk assessment. It discussed about different factors using the damage framework of flooding and identified the factors of flood risk assessment. It also figured out the uncertainties associated with each of the factors. This chapter has ended with the description of several risk assessment methods along with different numerical and symbolic methods for representing and reasoning under uncertainty.

Next chapter will describe different components of the BRB expert system methodology introduced in this chapter as well as the system built for this research. It will also discuss about some of the novel contributions made as a part of this research.

3 FRAMEWORK DESIGN AND SYSTEM IMPLEMENTATION

This chapter discusses about the methodology used for this work. It describes expert systems, how expert systems are built as well as the architecture of the expert system for flood risk assessment. It also explains different components and tools used for building the expert system along with a novel dynamic BRB tree traversal algorithm and RESTful API.

3.1 BRB Expert System Methodology

Rules are one of the most common form of expressing various types of knowledge. Rule based expert systems, developed based on human knowledge, has become one of the fastest growing branch of AI [33].

A rule based expert system is consist of two essential components: a knowledge base and an inference engine. It uses the observation provided by users and the rules developed by experts to infer useful outcomes. It is important to deal with uncertainty while designing and implementing a rule-based system. Different types of uncertainty can be caused in real systems such as vagueness, imprecision, incompleteness and ignorance [34]. Therefore, it is necessary to build a framework which can process and represent these uncertainties.

Belief Rule Based expert systems (BRBES) is a hybrid rule based systems based on Dempster-Shafer theory of evidence, fuzzy logic and decision theory. It uses IF-THEN structure for modeling rule bases and ER approach for inference [6] [35] [36] [37] [38]. BRB is a generic rule based inference methodology which uses RIMER approach for evidential reasoning. In RIMER approach, a detailed analysis is done on the antecedent attributes as well as different types of uncertainties in data. Then a generic rule base is designed using a belief structure. A rule base is generated on the basis of this belief structure, namely belief rule base, which is used to represent nonlinear casual relationships as well as uncertainty. Different knowledge representation parameters such as attribute weights, rule weights and belief degrees are also considered by this scheme.

Based on the referential values, input of each antecedent is transformed into a dis-

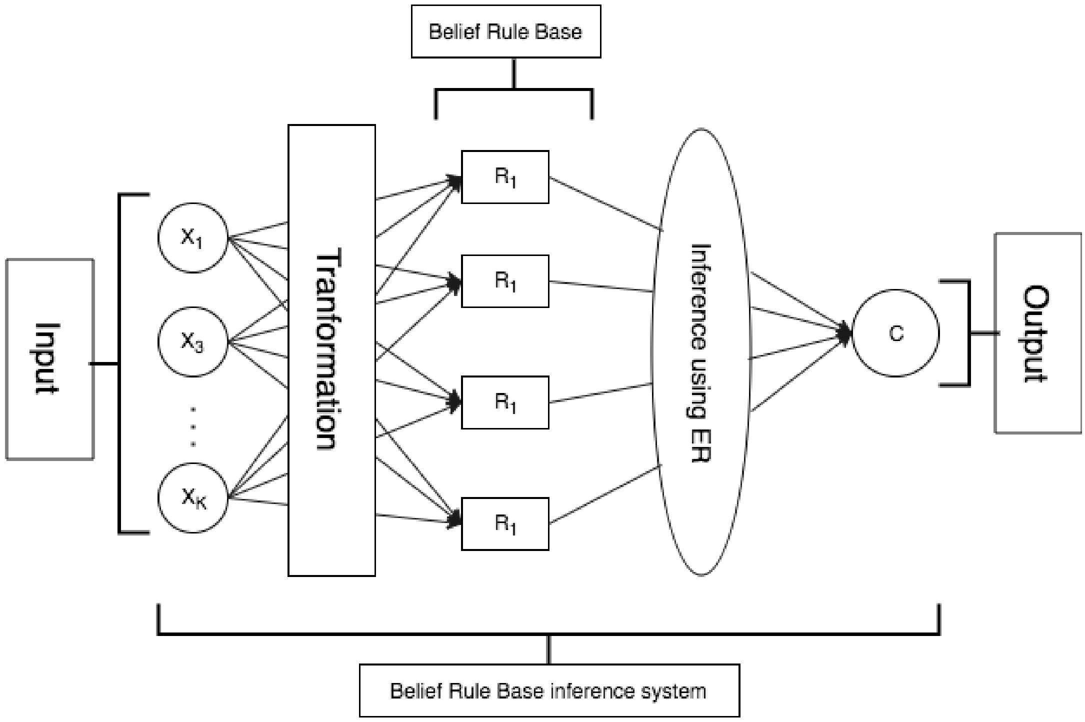


Figure 5. BRBESs inference procedures

tribution from the belief rule base. This distribution denotes the degree which has been activated for each antecedent. The activation weight of a rule is calculated by aggregating the degrees to which all antecedents are activated. Then, the output of consequent of the rule is obtained with a certain degree of belief. Finally, the inference mechanism is implemented for rule-based system using the ER approach.

Fig. 5 illustrates the architecture of BRBES inference procedures, consisting of four steps namely input transformation, activation weight calculation, belief update and rule aggregation.

Following section provides the basic understanding of the knowledge acquisition and representation procedures in BRB. This knowledge acquisition technique will be used to generate the initial BRB for the problem domain i.e., Flood Risk Assessment. The procedure of input transformation and rule update mechanism is also discussed. Rule aggregation in an initial BRB, which is one of the components of inference mechanisms, is also described in this section which will calculate the belief degree of consequent attribute for some input values of antecedent attributes.

3.1.1 BRBES Knowledge Acquisition Method

The design and implementation of conventional information systems usually consider that the problem is structured and complete information is available. However, this is not the case in many real case scenarios such as flood risk assessment. Moreover, the decision making process can be associated with uncertainty. Uncertainty may occur for various reasons. For example, it may occur due to partial or incomplete information coming from observations(e.g., interviews), or information which is not properly described, or due to inaccurate and poorly reliable instruments(e.g., sensors) used to make the observations.

There are four types of propositional statements which can be provided by the experts: crisp and certain, crisp and uncertain, fuzzy and certain or fuzzy and uncertain. Following examples elaborate the above phenomenon.

1. Bob is tall with certainty of 1 (fuzzy and certain)
2. Bob is tall with a certainty of 0.7 (fuzzy and uncertain)
3. Bob is 5 feet 10 inches with a certainty 1 (crisp and certain)
4. Bob is 5 feet 10 inches with a certainty of 0.8 (crisp and uncertain)

In the second example, 0.7 means that "we area 70% sure" that "Bob is tall" and the remaining 30% is ignorance. Ignorance can be caused in a rule based system by anemic implementation which may occur when experts are unable to precisely correlate between premise and conclusion, but only with degree of credibility or belief [39], [40].

A rule describes the casual relationship between antecedent attributes and their associated consequent. Propositions can be of three types in a rule-based system: boolean, fuzzy and random. Boolean propositions are assertive in nature which means they can be either true or false. Fuzzy propositions are related to vague concepts. For example, consider a IF-THEN rule for flood risk assessment:

$$\text{IF water level is high THEN flood is definite} \tag{4}$$

In this rule, water level is an example of fuzzy proposition. It can be calculated using the degree of membership of a numerical value to a fuzzy set which is modeled by defining a fuzzy set which takes a quantitative value in millimeters [41]. On the other hand, "raining causes flooding" is a probabilistic proposition rather than fuzzy or deterministic which can be evaluated using statistical method. There is a possibility to have heterogeneous types of attributes. Some attributes can be measured numerically(e.g., age) while others can only be represented subjectively(e.g., high). Continuous and numerical attributes are quantitative (e.g., Water Level, length of road affected) in general while symbolic and ordered symbolic attributes are qualitative (e.g., Financial Condition, Mental Condition).

Belief Rule is an extension of traditional IF-THEN rule. There are two main parts in a belief rule: antecedent and consequent. Antecedent attributes take the referential values while consequent of a belief rule is associated with belief degrees. To capture uncertainty is data, several knowledge representation parameters such as rule weight, antecedent attribute weight and belief degrees are used.

A belief rule can be defined in the following way:

$$R_k : \begin{cases} IF(P_1 is A_1^k) \cap (P_2 is A_2^k) \cap (P_3 is A_3^k) \cap .... \cap (P_N is A_{Nk}^k) \\ THEN\{(C_1, \beta_{1k}), (C_2, \beta_{2k}), (C_3, \beta_{3k}), ..., (C_N, \beta_{Nk})\} \end{cases} \tag{5}$$

$$\text{where } (\beta_{jk} \geq 0, \textstyle\sum_{j=1}^{N}(\beta_{jk}) \leq 1) \text{ with a rule weight}$$
$$\theta_k, attributeweights(\delta_{k1}, \delta_{k2}, \delta_{k3}, ...\delta_{kNk} k \in \{1, ...L\})$$

Here, P_1, P_2, P_3 ... are the antecedent attributes of the k_{th} rule. A_i denotes one of the referential values of the i_{th} antecedent attribute P_i while C_j is one of the consequent reference values of the rule. β_{jk} (j = 1, ... , k = 1, ... , L) is the degree of belief to which the consequent reference value is believed to be true. The k_{th} rule is said to be complete if $\sum_{j=1}^{N}(\beta_{jk}) = 1$. The rule is considered as incomplete if the summation is less than 1. It may be a case due to ignorance or incomplete information. N_k is total number of antecedent attributes used for k_{th} rule. Number of total belief rules is denoted by L and N is the number of all possible referential values of consequent.

IF Financial Condition is High AND Mental Condition is Low THEN
Indirect Intangible factor is {(High, 0.9), (Medium, 0.1), (Low, 0)} (6)

In this example, "Financial Condition" and "Mental Condition" are the antecedent attributes, while "High" and "Low" are their corresponding referential values. "Indirect Intangible" is the consequent attribute with referential values such as "High", "Medium", and "Low" where (High, 0.9), (Medium, 0.1) and (Low, 0) are the belief distribution of the consequent "Indirect Intangible" mentioned above. As the summation of belief degrees associated with referential values of the consequent attribute is one, the rule can be said complete. In traditional IF-THEN rule, he relationship between antecedents and the consequent attribute is linear, whereas it is non-linear in case of belief rule. Moreover, data collected from the surveys or interviews are non-linear in nature [42] and hence, it is possible to use belief rules to represent the data efficiently.

If the same rule is considered in Fuzzy Logic based expert system (FLBES), it can be written as follows.

IF Financial Condition is High AND Mental Condition is Low THEN Indirect Intangible factor is High (7)

It is clearly visible that the main different between FLBES and BRBES is that the degree of belief is not embedded with the consequent part of the rule in case of FL while n belief rule, the degree of belief is distributed over the referential values of the consequent. Since FLBES does not focus on the ignorance and incompleteness in consequent, it can not handle uncertainties due to ignorance and incompleteness and hence, BRBES is considered over FLBES in this research [43].

There are several methods for knowledge acquisition. Data can be collected by the literature review. Some literature may contain useful knowledge collected by researchers for their research in the same or different domain. Historical data is another source of data which can be used for further research. In an expert system, sometimes it is necessary to consider historical data of an event to predict the possibility of happening that event in near future. Data can also be collected using sensors. Since Internet of Things (IoT) is widely used over the world and different types of sensors became highly available as well as cheap, it is possible to collect different types of data using sensors with little investment [44]. Conducting interview is also a widely used method for data acquisition. A good set of questions can bring out reliable data with minimum effort. There are several cases where data may not be available via any source or collecting data using other methodologies are not possible. In those cases, interview is really a good method of data acquisition.

3.1.2 BRBES Inference Methodology

To generate belief rule base, inference mechanism is utilized. Inference mechanism is used to activate the rules, to update belief degree and to obtain the aggregated fuzzy values, which can be converted using the utility score into crisp value [45], [46], [47], [48].

Input Transformation

The main goal of input transformation is to distribute the input data over the referential values of the attribute of a rule [6]. The input transformation of a value of an antecedent P_i consists of the distribution of that input value into belief degrees of different referential values of that antecedent. The i_{th} value of an antecedent attribute can be transformed into a distribution across the referential values using their belief degrees given for the attribute. Following equation gives the assessment of the input value A_i.

$$H(A_i) = A_{ij}, a_{ij}, j = 1, ..., ji, i = 1, ..., N_K \tag{8}$$

Here, P_i is A_{ij} and the belief degree of the referential value with $a_{ij} \supset 0$ is a_{ij}. $\sum_{j=1}^{ji}(a_{ij}) \leq 1(i = 1, ..., N_k)$, H is the assessment of the belief degree assigned to the input value of antecedent attribute, referential value of the j_{th} input and j_i is the number of referential values.

The input value of an antecedent attribute is collected from the people residing in the flood affected area. Some of the values were given in qualitative terms while the others were given in quantitative terms. These values were distributed in terms of belief degree of different referential values of the antecedent attribute. It is possible to assign some utility values h_{ij} for the referential values A_{ij}. For example, the "High" referential value can be assigned a utility value $h_i 3 =$ '1', "Medium" can be assigned $h_i 2 =$ '0.5' and "Low" can be assigned $h_i 1 =$ '0'. The procedure for input transformation mentioned above can be elaborated using Eq. 9 and 10.

$$if h_{i3} \geq A_i \geq h_{i2}, then a_{i2} = \frac{h_{i2} - Ai_i}{h_{i2} - hi_2}, a_{i3} = 1 - a_{i2} \tag{9}$$

$$if h_{i2} \geq A_i \geq h_{i1}, then a_{i1} = \frac{h_{i2} - Ai_i}{h_{i2} - hi_2}, a_{i2} = 1 - a_{i1} \tag{10}$$

As the input is being distributed over the referential values of the attributes, each rule becomes dynamic after input transformation due to the changed values. Each transformed rule is generated from the initial rule base and kept in short-term memory such as Random Access Memory (RAM). These transformed rules are then used for further calculation in the next steps. Moreover, completeness of data is assured by the sum of all the consequent values to 1.

Activation Weight Calculation

Since the k_{th} rule of an initial rule base is constructed only by taking account of one of the referential values A_{jk} (an element of j) of an antecedent P_i in the initial rule base, determining the degree of belief a_{ik} of this referential value (A_{ik}) is necessary. It can be defined as the matching degree at which the belief is matched. a_{ij} can be obtained by the Eq. 8 - 9.

The combined matching degree α_k, to which the input matches the antecedent part of k^{th}rule, is evaluated using Eq. [49] .

$$\alpha_k = aggr((\delta_{k1}, \alpha_1^k), ..., (\delta_{kT_k}, \alpha_{T_k}^k)) \tag{11}$$

where *aggr* is an aggregation function which should be selected carefully. Following simple weighted multiplicative aggregation function can be used as an aggregation function [49].

$$\alpha_k = \prod_{i=1}^{T_k} (\alpha_i^k)^{\bar{\delta}_{ki}} \tag{12}$$

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

Recursive aggregation function can also be considered to calculate activation weight. The reason behind using multiplicative function instead of the recursive one is that recursive functions are computationally expensive in terms of time and memory. Recursion can also cause stack overflow in case of too many recursive calls when the rule base is substantially large.

When a matching degree is assigned to the referential values of the antecedent attributes, a rule is considered to be activated. This phenomenon is called Packet Antecedent of a rule. The activation weight of the k_{th} rule w_k is evaluated using Eq. 13 when the k_{th} rule is activated [6].

$$w_k = \frac{\theta_k \alpha_k}{\sum\limits_{i=1}^{L} (\theta_i \alpha_i)} \tag{13}$$

Here δ_{ki} is the relative weight of P_i. It is calculated by dividing the weights of P_i by the maximum weight of all antecedent attributes. The reason behind doing this is δ_{ki} becomes normalized in this way i.e. the range of it's value should be between 0 and 1.

Belief Update for Incomplete Data

For calculating the referential values of the consequent attribute, it is important to note that the rules have different weights. If a rule is not activated, activation weight of that rule will be 0. It is possible to have incompleteness in the consequent of a rule after the activation of a rule due to the insufficient information in the antecedents.

The belief degree associated with each rule in the rule base should be updated when an input data for any of the antecedent is ignored or missing. The belief degree of each of the rule is updated using Eq. 14 [49].

$$\beta_{ik} = \bar{\beta}_{ik} \frac{\sum\limits_{n=1}^{N_k} (\lambda(N,k) \sum\limits_{j=1}^{J_n} (\alpha_{nj}))}{\sum_{n=1}^{N_k} \lambda(n,k)} \tag{14}$$

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

Here, β_{ik} is the updated belief degree and $\bar{\beta_{ik}}$ is the original belief degree.

By updating the belief degree of each of the rules during belief update procedure, Eq. 14 can address the uncertainty in data due to ignorance and incompleteness.

Rule Aggregation

The ER approach was developed to handle multiple attribute decision analysis, which is a problem with heterogeneous attributes under uncertainty [35], [50], [51].

ER is an extension of Dempster-Shafer theory of evidence. It can be criticized that the computation complexity of reasoning for DST is a major problem if the combination rule is not properly used [52]. The aggregation of mass functions is NP-complete in case of DST while the it is linear in case of ER [35], [36], [37]. It should also be noted that uncertainty in data can be explicitly modeled in ER by using the normalized activation weight w_k and can be processed using evidential reasoning algorithm. In this way, ER is overcoming another drawback of DST in dealing with uncertain data.

This approach is used for aggregating all the packet antecedents of the rules and calculating belief degree of each referential value of the consequent attribute, taking account of input values P_i of antecedent attributes. The conclusion $O(Y)$ is evaluated using the analytical ER algorithm [53]. It consists of the referential values of the consequent attribute. Eq. 15 formalizes this approach.

$$O(Y) = S(P_i) = \big\{(C_j, \beta_j), j = 1, ..., N\big\} \tag{15}$$

Here, B_j is the belief degree associated with one of the consequent values such as C_j, β_j is calculated by analytical format of the ER algorithm [6]. The final belief degree β_j is evaluated, using the analytical ER algorithm, which is mentioned in Eq. 16 [53].

$$\beta_j = \frac{\mu \times [\prod_{k=1}^{L}(\omega_k \beta_{jk} + 1 - \omega_k \sum_{j=1}^{N} \beta_{jk}) - \prod_{k=1}^{L}(1 - \omega_k \sum_{j=1}^{N} \beta_{jk})]}{1 - \mu \times [\prod_{k=1}^{L} 1 - \omega_k]} \tag{16}$$

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

here, ω_k is the activation weight and $C_1\beta_1, C_2\beta_1, C_3\beta_1, ..., C_N\beta_N$ are the final combined output generated by ER. β_j is the final belief degree associated with the j_{th} referential value C_j of the consequent attribute, obtained by combining all the activated rules in the BRB using evidential reasoning. Analytical approach is preferable over recursive approach since it is efficient in terms of time and memory and do not cause failures, such as stack overflow, in case of too many data to handle [6], [36].

It is important to note that several types of uncertainties such as ignorance, incompleteness, vagueness, imprecision and ambiguity are addressed during the process of rule aggregation by Eq. 16 [53]. This step gives the overall picture of the scenario while considering both qualitative and quantitative data. The ability of handing heterogeneous data makes BRBES better and efficient compared to DST and FL and hence, it is preferable in such scenarios where both qualitative and quantitative data is considered, which is a common case for real world problem such as flooding.

Output of BRB Expert System

The output of the BRB system is non-fuzzy. This non-fuzzy or qualitative value can be transformed into crisp i.e., numerical value by assigning utility score to each referential value of the consequent attribute. Eq. 17 represents the calculation of numerical value from the non-fuzzy value.

$$y_m = \sum_{j=1}^{N}(u(C_j)\beta_j) \tag{17}$$

Here, $u(C_j)$ is the utility score for each referential value and y_m is the equivalent numerical value. Crisp value gives the quantitative picture of the whole scenario which is more readable and understandable for the end users.

BRBES methodology along with the procedure of handling various types of uncertainty associated with risk assessment factors has been described above. From this discussion, it can be argued that belief rule base expert system can identify all types of uncertainty associated with flood risk assessment. This subsection also gives a clear idea about how to build an expert system which considers the risk factors along with their associated uncertainties and hence, this subsection gives answer of the **research question 3** as mentioned in section 1.3.

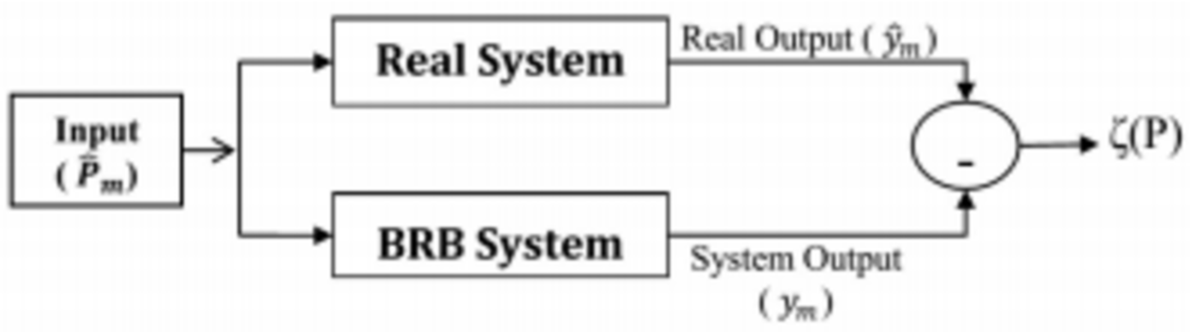


Figure 6. Optimal learning model for flood risk assessment

3.1.3 BRBES Learning Methodology

It is possible to acquire the learning parameters; rule weights, attribute weights and consequent belief degrees (θ_k, δ_i, β_{jk}); from domain experts or can be generated randomly for a BRB. However, these parameters may not be 100% accurate which can be obtained by training the BRB using historical data [54]. The objective of the BRB learning module is to obtain an optimal set of parameters (θ_k, δ_i, β_{jk}) which will minimize the discrepancy $\zeta(P)$ between the BRBES results (y_m) and system generated results ($\bar{y}_m$).

Several training models for BRB are available online and offline for qualitative, quantitative and mixed (both qualitative and quantitative) parameters. [54], [55], [56]. In this research, an optimization model for training BRBES in terms of mixed parameter has been developed. The reason behind selecting the one with mixed variable is, both qualitative and quantitative data have been considered in this research and hence, it is necessary to develop an optimization model which can handle heterogeneous data types. The optimal learning model for the flood risk assessment has been shown in Fig. 6.

This model consist of three steps, (a) construction of an objective function, (b) setting the constraints for the training parameters and (c) development of a training module to search the optimal set of parameters (θ_k, δ_i, β_{jk}), assuming that there are M possible combinations in a training sample and the input-output pairs of the M cases are $(\bar{P_m}, \bar{y_m})(m = 1, ..., M)$.

For minimizing the discrepancy $\zeta(P)$, following equation is used.

$$\zeta(P) = \frac{1}{M}\sum_{m=1}^{M}(\bar{y_m} - y_m)^2 \tag{18}$$

Minimizing the total mean squared error, $\min(\zeta_P)$, is the main objective here. Since both numerical and subjective values are present in the system, suppose that first M_1 pairs of training data are numeric values while last $M_2 = M - M_1$ pairs of training data are subjective. The optimization problem, formulated as the multiple-object optimization problem, can be used in this case for minimizing the difference between BRB system generated outputs and the corresponding observed outputs as shown in Eq. 19.

$$\min\{\zeta_1(P), \zeta_2(P), ..., \zeta_N(P); \zeta(P)\} \tag{19}$$

where ζ_j is the total mean squared error for the j_{th} referential term and P is the set of training parameters mentioned in Eq. 19.

The object evaluated from Eq. 19 has $N+1$ non-linear objective function. It can be solved using the minimax formulation as shown in Eq. 20.

$$\min(p)\max(\zeta_j)\{w_j\frac{\zeta_j(P) - \zeta_j^*}{\zeta_j^+ - \zeta_j^*}, j = 1, ..., N+1\} \tag{20}$$

Steps of calculating the minimax method for solving the multiple-objective problem are described here.

Step 1) Solving the single-object optimization problem:

$$\min\{\zeta_1(P)\} \tag{21}$$

$$\text{for } j = 1, 2, ..., N.$$

Step 2) Setting the relative weight vector $w = (w_1, ..., w_N; w_{N+1})$. One possible way to set the weight vector is as follows:

$$w_1 = ... = w_N = \frac{M_1}{N} w_{N+1} = M_2 \tag{22}$$

It means that the importance of one type of objective function depends on the number of training data sets.

Step 3) Reformulating the single-objective problem mentioned in Eq. 21 in an equivalent problem, shown in Eq. 23 and then get the solution of it.

$$min_r \tag{23}$$

$$s.t.\{w_j \frac{\zeta_j(P) - \zeta_j^*}{\zeta_j^+ - \zeta_j^*} \leq r, j = 1, ..., N+1 \tag{24}$$

If there is a need of interactive way of regulation for the relative weights, the above mentioned process can be repeated.

The learning module is described in this section. It is also explained that how this module can be integrated with the BRBES to make the expert system intelligent and hence, this subsection gives the answer of the **research question 4** as mentioned in section 1.3.

3.2 BRB Expert System for Flood Risk Assessment

This section gives the overview of the BRB expert system for assessing flood risk which includes the system architecture, knowledge base construction and JSON based data generation. It also describes a dynamic BRB tree traversal algorithm developed as a part of this research work as well as the web based BRBES implement methodology, RESTful API, graphical user interface and the learning module.

3.2.1 System Architecture

The choice of a development methodology is crucial while designing and building a system. In software engineering, high degree of optimality can be achieved by considering different factors such as development time, cost of development, product quality and maintainability. There are different methodologies for web development, but any model of software development is build in several different phases.

For this research, the Iterative and Incremental Development(IID) methodology has been chosen. The reason behind that is IID is an essential part of modern Agile Software Development methodology. The main purpose of IID is to design, implement and test the system incrementally, which means a little part of the whole system design is taken in concern each time and then that part is implemented and tested, until the product is ready [57].

The system architecture for the web based flood risk assessment system is represented in the Fig. 7.

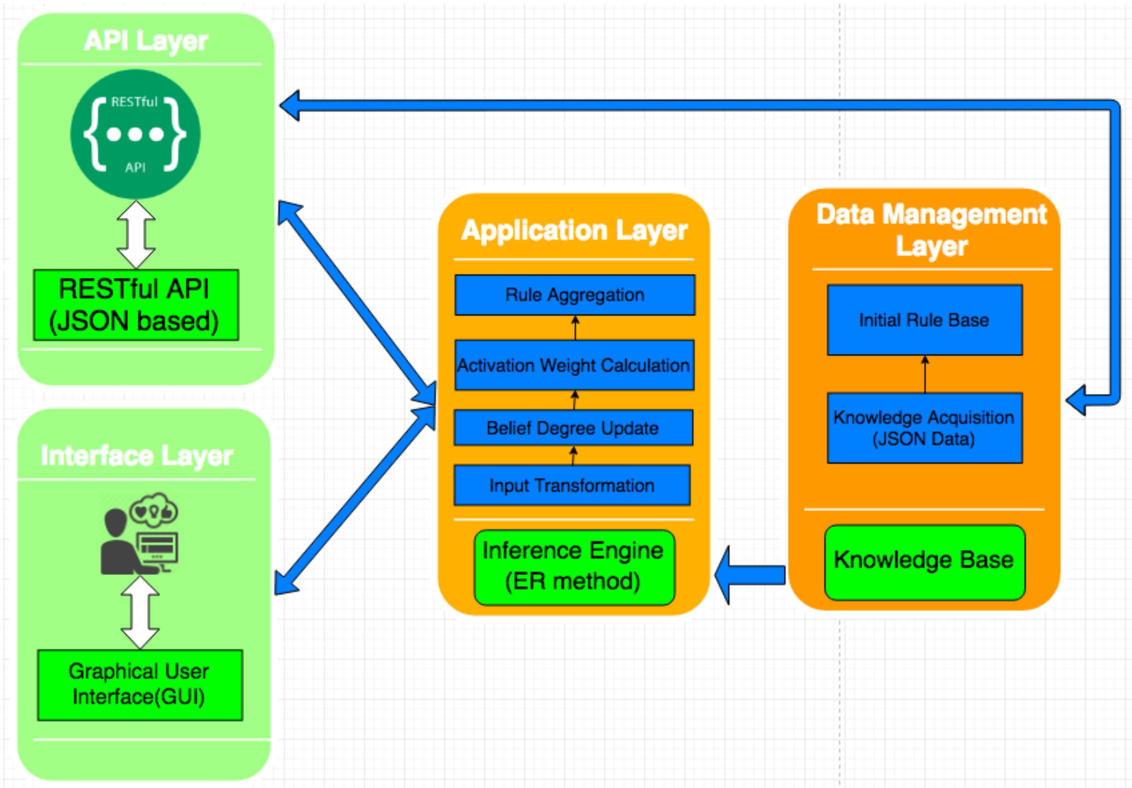


Figure 7. Architecture of BRB expert system

The architecture represents how its components are organized. The BRB expert system developed here adopt a four-layer architecture model, including data management layer, application layer, interface layer and Application Programming Interface (API) layer. Data management layer is responsible for the initial rule base creation using the data collected from the case study area during survey. The initial BRB is generated in the data management layer which is the knowledge base of the system. Application layer consists of inference engine with procedures including input transformation, rule activation, rule update and rule aggregation which are the parts of ER based expert system as mentioned earlier in this chapter. Application layer, which is consist of the inference engine, takes the initial BRB as input from data management layer. Interface layer is responsible for showing the system output in more human readable way which can be accessible by the users via a web interface. API Layer is a layer of abstraction of the underlying system for users which gives a simple programming interface for generating output of different levels of the BRBES using Uniform Resource Locator (URL), named as API endpoints. This API can communicate with the data management layer, application layer and provides a communication channel for the users using high level of abstraction.

3.2.2 Knowledge Base Construction

A BRB framework was developed to construct the knowledge base for this expert system while taking the factors associated with risk assessment of flood, as shown in Table 1.

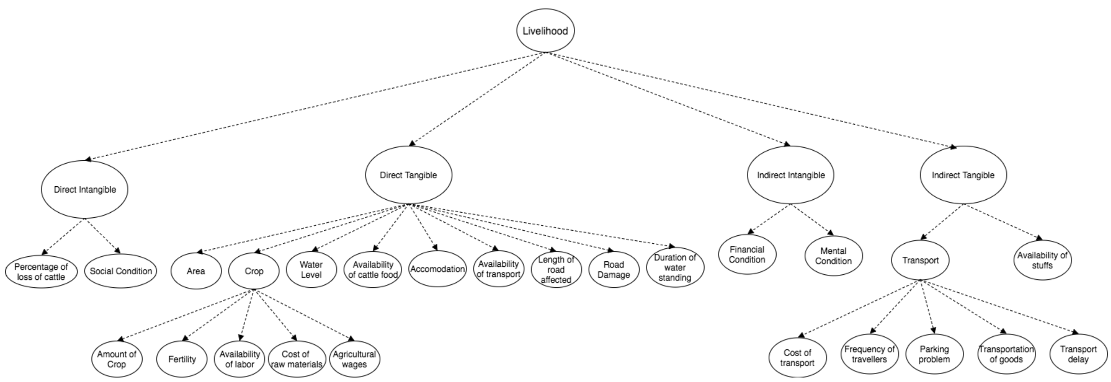


Figure 8. BRB Framework to assess the flood risk on livelihood

It can be observed that the framework includes all the input variables, which determine the evaluation of livelihood, mentioned in Table 1. To reduce the computational complexity, five variables under the category "Direct Tangible" were again categorized into a sub-group named "Crop" and five variables under "Indirect Tangible" category were categorized into a sub-group named "Transport" as shown in the Fig. 8.

There are four ways a BRB can be established: 1) by extracting belief rules from domain expert knowledge, 2) by extracting belief rules from historical data, 3) using random rules and 4) using previous rule bases, if available. if there is no prior knowledge [58]. The initial BRB was constructed by taking the knowledge of experts in this research. Equal weights has been assigned to all the belief rules as well as to all the antecedent attribute by the experts which is 1.

Tables 2, 3, 4, 5, 6 show all the initial beilef rules created for the nodes "Direct Intangible", "Direct Tangible", "Indirect Intangible", "Indirect Intangible" and "Livelihood". Referential values of the consequent were calculated for all combinations of referential values of the antecedent nodes to generate the initial rule base.

Table 2. Initial BRB for the factor "Direct Intangible"

		IF		THEN(Direct Intangible)		
Rule Identifier	Rule Weight	Percentage of Loss of Cattle	Social Condition	High	Medium	Low
R1	1	HIGH	HIGH	1.0	0	0
R2	1	HIGH	MEDIUM	0.5	0.5	0
R3	1	HIGH	LOW	0	1.0	0
R4	1	MEDIUM	HIGH	0.5	0.5	0
R5	1	MEDIUM	MEDIUM	0	1.0	0
R6	1	MEDIUM	LOW	0	0.5	0.5
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0	0	1.0

Table 3. Initial BRB for the factor "Direct Tangible"

		IF									THEN(Direct Tangible)		
Rule Identifier	Rule Weight	Area	Crop	Water Level	Availability of Cattle Food	Accommodation	Availability of Transport	Length of Road Affected	Road Damage	Duration of Water Standing	High	Medium	Low
R1	1	8.0	HIGH	3.0	0.33	1	0.33	10	43.75	2.5	1.0	0	0
R2	1	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.94	0.06	0
R3	1	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.88	0.12	0
..	..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..
R19681	1	10.0	LOW	3.0	0.33	1	0.88	10	43.75	3.0	0	0.12	0.88
R19682	1	9.0	LOW	3.5	0.5	1	0.88	12	43.75	2.5	0	0.06	0.94
R19683	1	10.0	LOW	3.5	0.5	1	0.33	12	43.75	3.0	0	0	1

Table 4. Initial BRB for the factor "Indirect Intangible"

		IF		THEN(Indirect Intangible)		
Rule Identifier	Rule Weight	Financial Condition	Mental Condition	High	Medium	Low
R1	1	HIGH	HIGH	1.0	0	0
R2	1	HIGH	MEDIUM	0.5	0.5	0
R3	1	HIGH	LOW	0	1.0	0
R4	1	MEDIUM	HIGH	0.5	0.5	0
R5	1	MEDIUM	MEDIUM	0	1.0	0
R6	1	MEDIUM	LOW	0	0.5	0.5
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0	0	1.0

Table 5. Initial BRB for the factor "Indirect Tangible"

		IF		THEN(Indirect Tangible)		
Rule Identifier	Rule Weight	Transport	Availability of Stuffs	High	Medium	Low
R1	1	HIGH	HIGH	1.0	0	0
R2	1	HIGH	MEDIUM	0.5	0.5	0
R3	1	HIGH	LOW	0	1.0	0
R4	1	MEDIUM	HIGH	0.5	0.5	0
R5	1	MEDIUM	MEDIUM	0	1.0	0
R6	1	MEDIUM	LOW	0	0.5	0.5
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0	0	1.0

Table 6. Initial BRB for the factor "Livelihood"

		IF				THEN(Livelihood)		
Rule Identifier	Rule Weight	Direct Intangible	Direct Tangible	Indirect Intangible	Indirect Tangible	High	Medium	Low
R1	1	HIGH	HIGH	HIGH	HIGH	1.0	0	0
R2	1	HIGH	HIGH	HIGH	MEDIUM	0.75	0.25	0
R3	1	HIGH	HIGH	HIGH	LOW	0.5	0.5	0
R4	1	HIGH	HIGH	MEDIUM	HIGH	0.75	0.25	0
R5	1	HIGH	HIGH	MEDIUM	MEDIUM	0.5	0.5	0
..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..
R77	1	LOW	LOW	MEDIUM	MEDIUM	0	0.75	0.25
R78	1	LOW	LOW	MEDIUM	LOW	0	0.5	0.5
R79	1	LOW	LOW	LOW	HIGH	0	0.25	0.75
R80	1	LOW	LOW	LOW	MEDIUM	0	0.5	0.5
R81	1	LOW	LOW	LOW	LOW	0	0	1.0

For example, all combinations of referential values of the antecedent nodes, "Financial Condition" and "Mental Condition", referential values of the consequent "Indirect Intangible" were calculated to generate the initial knowledge base which is visualized in Table 4. As there are two antecedent nodes and each antecedent node has three referential values, the total number of rules should be $3 * 3 = 9$ for this rule base. A belief rule, taken from Table 4, is illustrated below.

$$\begin{array}{l} \text{IF Financial Condition is High AND Mental Condition is Medium} \\ \text{THEN Indirect Intangible } \{(\text{High, 0}), (\text{Moderate, 0.5}), (\text{Low, 0.5})\} \end{array} \tag{25}$$

The belief degrees are attached to the three referential values of the consequent in this rule. Since the summation of degree of belief $(0 + 0.5 + 0.5)$ for this rule is 1, it can be said that the rule is complete. The consequent in traditional IF-THEN rules is assertive in nature and hence, it is difficult to represent a real world knowledge when such rule base is used. In addition, the belief structure provides the flexibility of representing knowledge of different structures and degrees of complexity. If the aforementioned rule is considered, the casual relationship between the consequents with three antecedents is complex, non-linear and uncertain as the relationship among them is not proportional.

3.2.3 JSON-based Data Generation

There are several data formats which can be used to make dataset for the BRB such as Comma Separated Values (CSV), eXtensible Markup Language (XML) and JavaScript Object Notation (JSON). Data can be delivered in any of these forms. CSV is a data representation form where data are delimited by comma(,) [59]. CSV files are hard to read when there is double commas or commas in unusual places. As the file only represents values separated by commas, if comma position is not right, it is impossible to read data from CSV. Moreover, validating a CSV file is tough as there is no convention of how the data will be kept in CSV files. XML is markup language which keeps data using tags which is similar to any other markup languages [60]. XML is too complex and need extra information to validate the data even for the simplest document which is problematic. Moreover, XML tags are not self-descriptive and the notion of "type" in XML adds extra layer of confusion and complexity while reading the data. JSON is an well-known data interchange format built on two possible structures: 1) ordered list of values (equivalent to array, vector, list or sequence in different languages) and 2) collection of key-value pairs (referred as object, record, dictionary, hash table, keyed list or associative array in different languages) [61].

In this research, JSON is considered as data format. The reason behind choosing JSON over CSV and XML is, JSON is a text format that is completely language independent, which means it does not depend on any language or framework as it has it's own convention of declaration. Moreover, due to the flat data structure, it is very easy to read and write JSON data for human. In addition, it is also easy for machines to parse and generate JSON data. Another reason of choosing JSON is that it is very lightweight compared to the other formats mentioned above.

An example of JSON data created for this research is illustrated in Fig. 9. It is created using key-value pair data structure as mentioned above and represents a node 'x8' of the tree. There are eleven keys for a node in this structure which are "antecedent_id", "antecedent_name", "attribute_weight", "rule_weight", "ref_val", "ref_title", "consequent_values", "crisp_val", "parent", "input_val" and "is_input". "antecedent_id" is the id of an antecedent which is visible in the tree where "antecedent_name" is the name of the antecedent in descriptive manner. "attribute_weight" and "rule_weight" are the value of attribute weight and rule weight, respectively, for the node which are set to "1" by default. "ref_val" is an array (also can be treated as vector, list or sequence) of referential values while "ref_title" is an array for their text representation. "consequent_values" is an array for the consequent values of a node while "crisp_val" represents the crisp value of the node. "parent" indicates the parent of the node in a BRB tree (for the top node the parent of the node is itself), "is_input" defines if this node is a leaf node or not and "input_val" is the input value for each node given by the user.

The data format for each node is same as mentioned and it comprises the whole BRB tree as JSON object.

```json
{
  "x8": {
    "antecedent_id": "x8",
    "antecedent_name": "Financial Condition",
    "attribute_weight": "1",
    "rule_weight": "1",
    "ref_val": [
      "1",
      "0.5",
      "0"
    ],
    "ref_title": [
      "High",
      "Medium",
      "Low"
    ],
    "consequent_values": [
      "[]"
    ],
    "crisp_val": "",
    "parent": "x3",
    "input_val": "",
    "is_input": "true"
  }
}
```

Figure 9. JSON data format for a node in BRB framework

3.2.4 A Dynamic BRB Tree Traversal Algorithm

To traverse the BRB tree dynamically, an algorithm has been developed which is illustrated in Algorithm 1. This algorithm uses a bottom-up approach for BRB tree traversal. This algorithm is a modified implementation of traditional Breadth First Search (BFS) algorithm which does the tree traversal in top-down approach. Both the space and time complexity of this algorithm is linear i.e. $O(n)$, where n is the number of nodes in the tree. It means that the algorithm is as efficient as BFS for bigger trees regardless the space and time complexity.

This algorithm starts with all the nodes of the tree kept in an array "Nodes". Then it takes the first element of the "Nodes" and names it "begin" and checks the parent of this node. As mentioned earlier, this information is available in the JSON data as the key "parent". Then it keeps the node in a array named "sibliings" and checks all the other nodes if any of the nodes has the same parent. If yes, that node is added in the "siblings" array. After getting all the siblings, it checks if all the siblings has "is_input" true or not. Again, this "is_input" key is also present in the BRB JSON. If all the siblings have "is_input" true that means this subtree can be send to the ER module to run BRBES inference procedures. After this step, all the siblings and the parent's (consequent) referential values are calculated.

After running the BRBES inference procedures on the children nodes, the "is_input" key of the parent is check as "True" and all the "siblings" nodes are removed from the "Nodes" array. This process is repeated until all the elements in the "Nodes" array are gone though the ER module to calculate the consequent referential values. The algorithm stops when the "Nodes" array become empty.

Algorithm 1 Dynamic BRB Tree traversal algorithm

procedure BRBTreeTraversal
  $Nodes \leftarrow \{x_i, x_{i+1}, ..., x_N\}$
  top:
  $begin \leftarrow \{Nodes_0\}$
  $parent = Parent_{begin}$
  $siblings = \{begin\}$
  $j \leftarrow i + 1$
  $loop$:
  if $parent = Parent_{Nodes[j]}$ **then**
    $siblings \leftarrow Nodes[j]$
    goto $loop$.
  $isAllInput = False$
  $loop$:
  if $isInput_{siblings[k]=True}$ **then**
    $isAllInput = True$
    goto $loop$.
  if $isAllInput = True$ **then**
    Run ER on the subtree to find out Parent referencial values
    $isInput_{Parent} = True$
    $Nodes = Nodes - siblings$
    goto top
  if $len(Nodes) = len(siblings)$ **then**
    Run ER on the subtree
    break

The beauty of the above described tree traversal algorithm is that it is generic and can be applicable for any BRB tree from any domain. No matter for which domain the BRB tree is created for, this algorithm can traverse the whole tree from bottom to top to generate the result of the top node by calculating the subtrees step by step.

3.2.5 Web-based BRBES Implementation Strategy

Web-based applications allow users to easily access and use the system without any additional software to install on users machine. Web-based applications are deployed and maintained in one location, namely web server, which can be accessed via specific Internet Protocol Address (IP address) and URL. They provide a standardized way of communication between different applications. Internet Protocols such as XML, Simple Object Access Protocol (SOAP), Hypertext Transfer Protocol (HTTP), Web Service Definition Language (WSDL) and Universal Description, Discovery, and Integration (UDDI) can be used for the communication [62], [63], [64], [65], [66]. Web services can be of two types: 1) Simple Object Access Protocol (SOAP) based and 2) Representational State Transfer (REST)-compliant [67], [68]. Both of these service types provide the flexibility of communication in cross platforms via simple service endpoints. It allows other applications to communicate with the system using the aforementioned protocols. Therefore, developing a web-based expert system will provide easy deployment and maintenance of applications as well as better user accessibility.

In this research, REST compliant web services over SOAP is chosen. REST compliant web services are more lightweight than SOAP and it is easier to implement by any modern programming language such as Python, PHP, Java and Javascript. Moreover, REST has more flexibility in terms of output format as it supports to provide output in CSV, JSON and XML while SOAP only supports XML. Moreover, REST uses HTTP over WSDL and UDDI, which makes it simpler [69]. A RESTful Application Programming Interface (API) has been developed as a part of this research which is built on the "uniform interface" constraint. This API shall be available in a web server so that anyone can use the API for their convenience without even programming a single line for the Belief Rule Base expert systems.

3.2.6 RESTful API-based BRB Expert System

Previously, BRBES was developed using Visual Basic and Microsoft SQL which runs on a user machine [70]. Due to the concreteness of the system, only limited number of people had the access to it. Moreover, user machines have limited computational capability. For these reasons, it was essential to develop a web based expert system application which will allow to run the expert system on a web server with more computational power and memory as well as will provide easy access for multiple users instead of running on local machine. This API will also allows other applications or users to run their own BRB framework without reinventing the wheel.

RESTful API's communicate with HTTP verbs. These verbs provide the action counterpart to the noun-based resource. The most common HTTP verbs are POST, GET, PUT and DELETE [71] which are used for create, read, update and delete (or CRUD in short-term) operations, respectively. The API initializes the BRB algorithm with a POST request which takes the JSON data file mentioned earlier as input with the URL "/v1/initiate_brb". This URL will return an access key which is a "sha1" based hash represented in hexadecimal format. After getting the access key, it can be used for further API calls. A sample of the initialization of the RESTful API with JSON data is illustrated in Fig. 10.

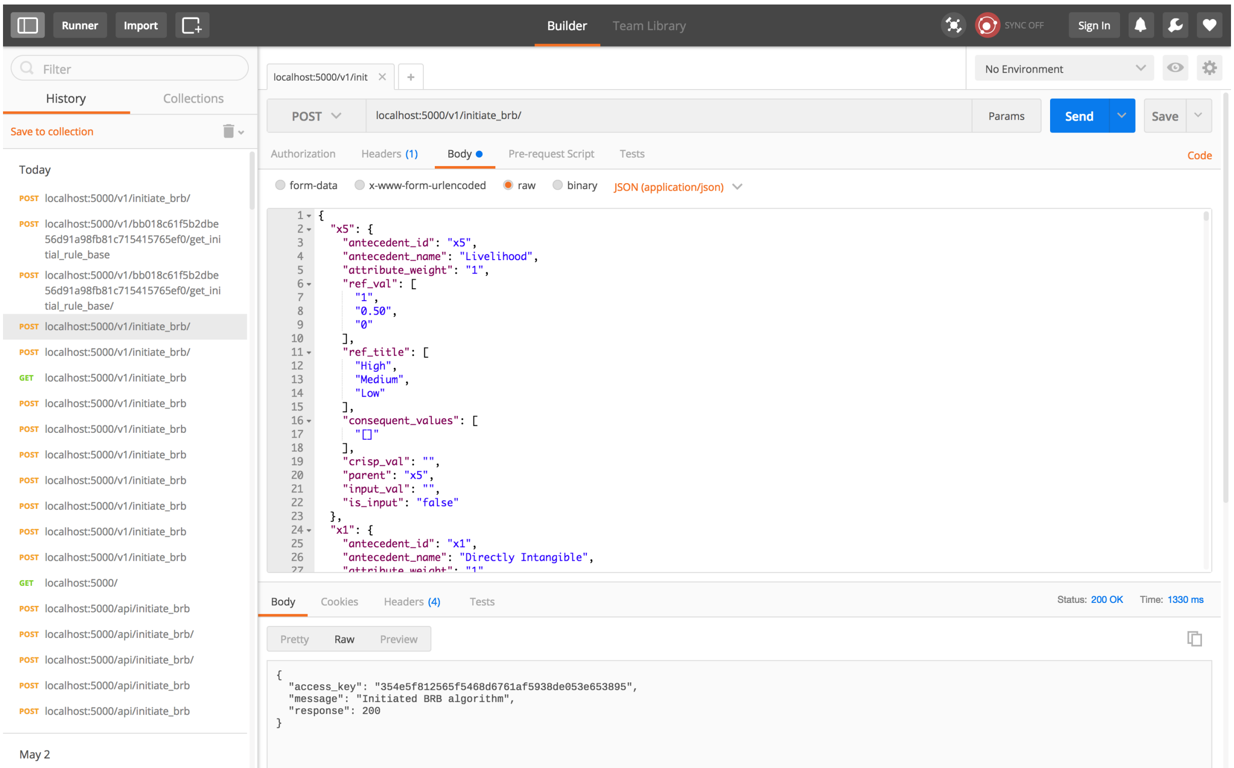
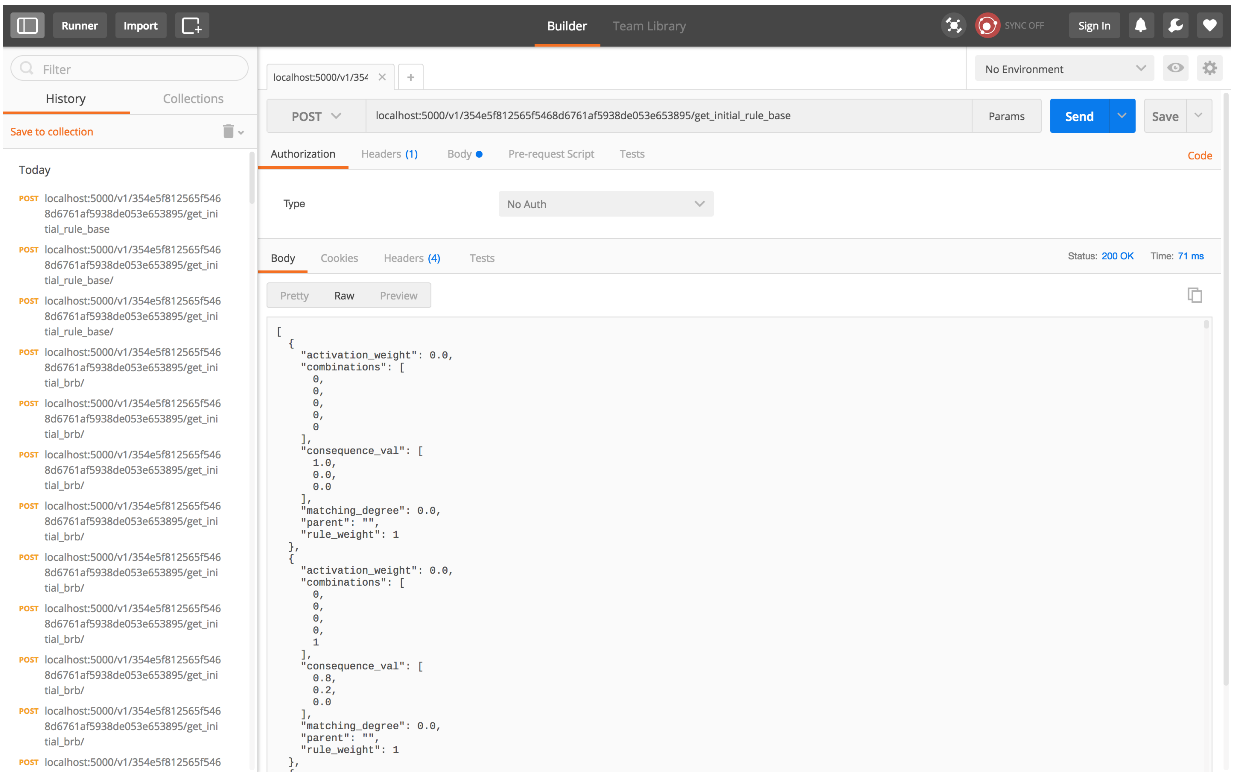


Figure 10. Initialization of BRB algorithm with JSON data for the RESTful API

The output of the API call is shown in the lower part of the Fig. 10 and there is an "access_key". After initializing the algorithm, the API can be accessed using different URL's, commonly known as API endpoints, using the access key. For example, to generate the initial rule base for the BRB algorithm, an API call with a GET request can be trigged using the URL "/v1/<access_key>/get_initial_rule_base" where "access_key" is the key received from the output of "/v1/initiate_brb". This sample API call and the output is shown in Fig. 11.



[
  {
    "activation_weight": 0.0,
    "combinations": [
      0,
      0,
      0,
      0,
      0
    ],
    "consequence_val": [
      1.0,
      0.0,
      0.0
    ],
    "matching_degree": 0.0,
    "parent": "",
    "rule_weight": 1
  },
  {
    "activation_weight": 0.0,
    "combinations": [
      0,
      0,
      0,
      0,
      1
    ],
    "consequence_val": [
      0.8,
      0.2,
      0.0
    ],
    "matching_degree": 0.0,
    "parent": "",
    "rule_weight": 1
  },

Figure 11. Output of the RESTful API for the initial rule base in JSON format

The output consists of all the initial rule bases created for BRB where there are several fields such as "activation_weight", "combinations", "consequence_val", "matching_degree", "parent" and "rule_weight". "combinations" is the combination of antecedent referential values for which the rule is generated and "consequence_val" are the referential values of the consequent for each rule. "parent", "activation_weight" and "rule_weight" are the same fields described earlier.

This RESTful API is built in an interpreted programming language named python using a web application framework named Flask. Python is a general purpose, cross platform language, used for scripting as well as to build web, mobile and desktop applications for different operating system platforms (e.g., Linux, Mac OSX, Windows). Flask is a highly configurable micro-framework for web development in python which contains a web server and template engine at startup [72]. The reason behind choosing python for the API instead of other programming languages (e.g., Java, PHP, C#, etc.) is, it has lots of good libraries to work on and really easy to learn and implement. The reason behind choosing Flask over other python web application frameworks (e.g., Django, Tornado, Falcon, etc.) is, as it is a micro-framework, it is very lightweight as well as takes minimum effort to configure, program and test an API.

3.2.7 Graphical User Interface for BRB Expert System

A graphical user interface (GUI) has been created for the BRBES as a part of this thesis work. GUI provides a visual platform to enables the interaction between the user and the system. The GUI for the BRBES for flood risk assessment is illustrated in Fig. 12.

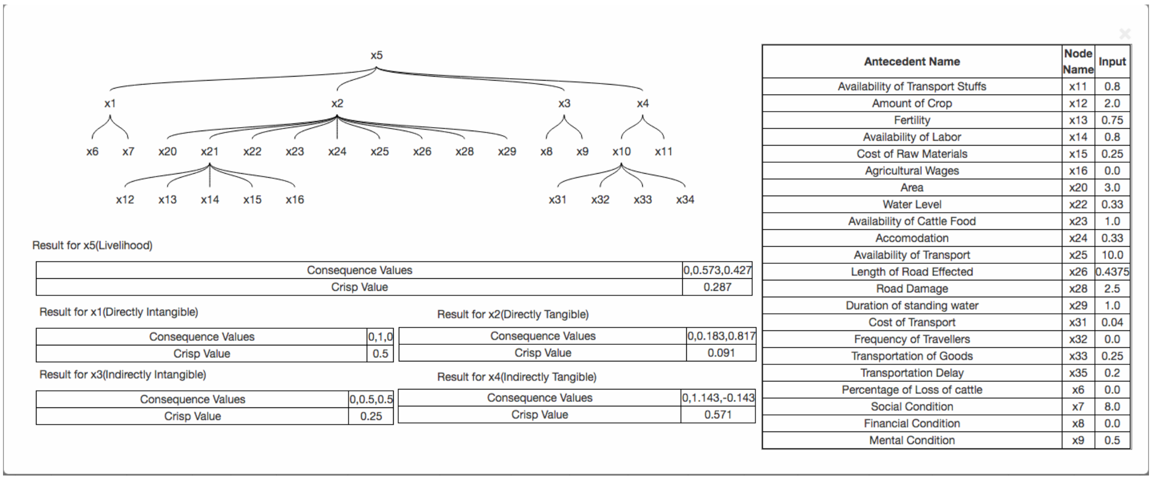


Figure 12. GUI of BRBES for flood risk assessment

This GUI displays the data for the antecedent attributes (leaf nodes) of the BRB framework shown in Fig. 8. These data are collected from the case study area by interviewing people. This interface enables the displaying of the evaluation result of the top node, x5 which is 'livelihood' as well as the result for the sub rule bases for four categories i.e. x1 (Direct Tangible), x2 (Direct Intangible), x3 (Indirect Intangible) and x4 (Indirect Tangible). There are two parts of the result for each node: consequence values and crisp value. Consequence value is obtained by applying Eq. 15 and 16 which are the referential values or belief degrees of the antecedent attributes. Crisp value is the quantitative value which is calculated from Eq. 17 using the consequent values and it returns one numerical value, which gives the quantitative picture of the whole scenario.

Fig. 12 also illustrates the whole flood risk assessment framework at the high or aggregated level.

3.2.8 Learning Module

The BRBES learning methodology is described above. The aim of the learning module is to find a optimal set of parameters, P, in a BRB system so that the difference between the calculated and estimated value of "livelihood" is minimized as illustrated in Fig. 6 as well as in Eq. 18. Fig. 13 illustrates the architecture of the proposed BRBES system along with the learning (training) module.

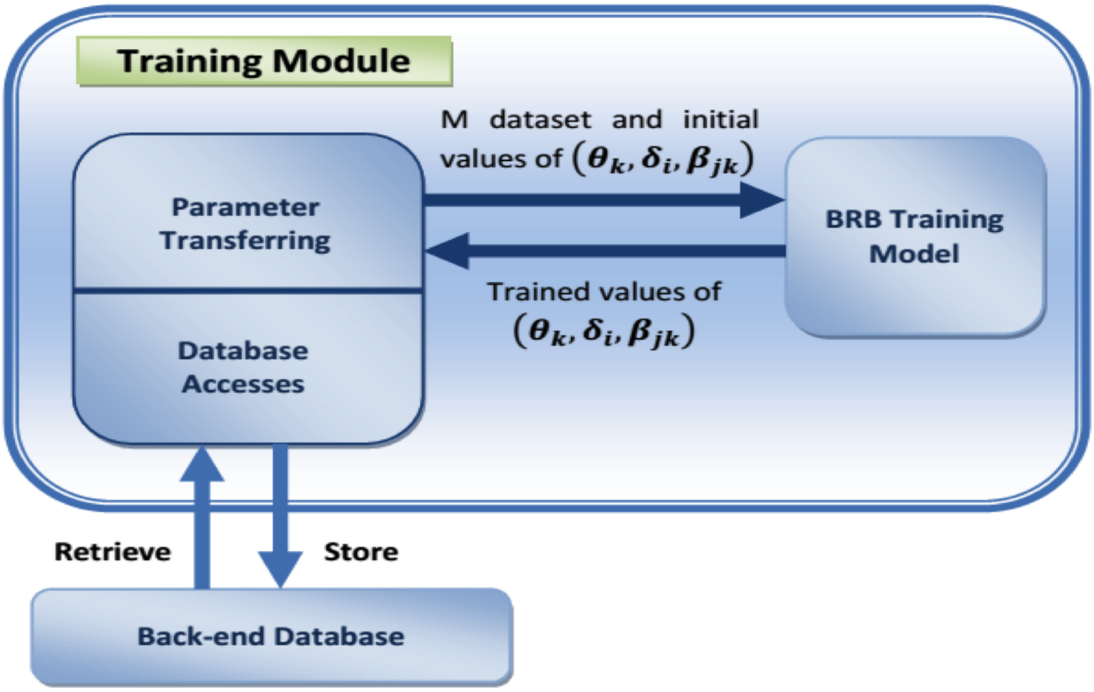


Figure 13. Architecture of the BRB learning module

If the example subtree of "Indirect Tangible" is considered from the belief rule base framework, then four different sets of parameters can be considered for training the BRB. The same objective function from Eq. 18 is used for each training round.

1. T_1: Train with different sets of rule weight θ_k.

2. T_2: Train with different sets of attribute weight δ_i.

3. T_3: Train with different sets of referential values of the two antecedents, "Financial Condition" and "Mental Condition".

4. T_4: Train with different sets of rule weight, activation weight, consequent belief degrees β_{jk} and referential values of the antecedents.

   Along with the belief degrees of the consequent, the best values of the parameters including attribute weight, rule weight and the referential values of the antecedent, which generate minimum errors are considered for T_4.

For learning, following constraints and initial values have been considered for each parameter mentioned above.

1. Rule Weights $\theta_k(k = 1, ...L)$:
   $1 \geq \theta_k(k = 1, ...L) \leq 0$;
   $\theta_k(k = 1, ...L) = 1$;

2. Attribute Weights $\delta_k(k = 1, ...L)$:
   $1 \geq \delta_k(k = 1, ...L) \leq 0$;
   $\delta_k(k = 1, ...L) = 1$;

3. Acuteness of the three referential titles.
   $\mu(C_j)(j = 1, ..., 3)$ where $1 \geq \mu(C_j) \leq 0$;
   $\mu(C_1(High)) > \mu(C_1(Medium)) > \mu(C_1(Low))$;
   $\mu(C_1(High)) = 1.0)$;
   $\mu(C_1(Medium)) = 0.5)$;
   $\mu(C_1(Low)) = 0)$;
   The range of the both antecedent attributes, "Financial Condition" and "Mental Condition" are in the range 0-1.

4. Belief degree of the consequent $\beta_{jk}(j = 1, ..., 3; k = 1, ..., L)$:
   $1 \leq \beta_{jk}(j = 1, ..., 3; k = 1, ..., L) \geq 0$;
   $1 \leq \sum_{j=1}^{3} \beta_{jk}(k = 1, ..., L) \geq 0$;

After setting all the initial values for the parameters, the optimal value of the training parameters are obtained. This learning module was developed using Matlab and random numbers generated by the *rand*() function of Matlab. For the other sub-rule bases of the BRB tree; shown in Fig. 8, the rule weight and attribute weight of the antecedent and the belief degrees of the consequents were trained in the same way.

To optimize the parameters, several optimization tools, such as *FMINCON* or *FMINMAX* functions, are available in Matlab. *FMINCON* is used for single-objective model while *FMINMAX* is used for optimizing multi-objective model. The optimization problem for this research is a multi-object optimization problem, as described earlier, and hence, *FMINMAX* function was used. The learning module of BRB can be understood from the flowchart illustrated in Fig. 14.

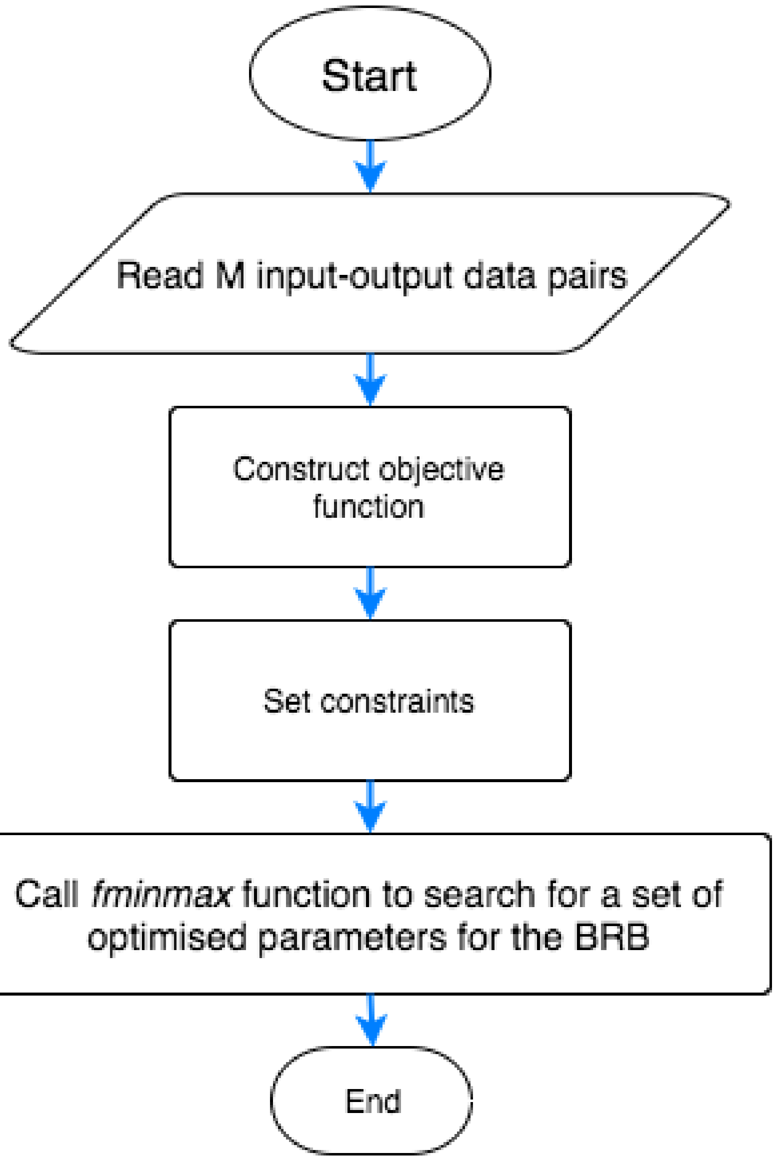


Figure 14. Flowchart of the BRB learning module

3.3 Summary

This chapter focused on the methodology for this research. It described how to build an expert system as well as different components and tools used for building the expert system. This chapter also discussed about the BRBES built in this research. Different parts of this expert system such as knowledge base, inference mechanism are also described along with the web based implementation of the expert system. It also explained a novel dynamic tree traversal algorithm for traversing BRB tree, the GUI created for the web based implementation and the RESTful API. Next chapter will take a case study area where this expert systems is going to be implemented to assess flood risk.

4 USE CASE: FLOOD AFFECTED NEIGHBORHOOD

This chapter presents the case study area to apply the BRBES to assess the risk of flooding on livelihood. The procedures of data collection are also presented in this chapter. The reliability of the flood risk assessment BRBES is also elaborated. The validity of the learning module is also discussed using the result evaluated from the system using collected data.

4.1 Survey of People

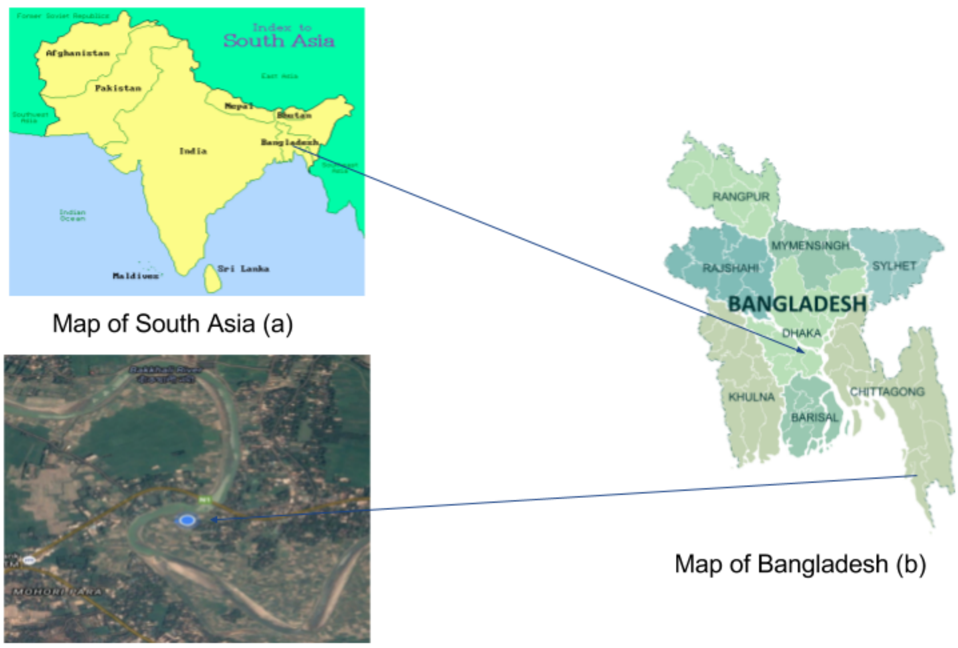
To feed the developed BRBES with real world data a survey was conducted by taking a case study area, located in the Cox's Bazar district of Bangladesh.

4.1.1 Case Study Area

Bangladesh can be treated as a land of natural disasters, since it is one of the countries in the world that suffers from frequent flooding often followed by loss of lives and economic damage [73]. For this reason, Bangladesh has been chosen as the case study area for this research. Moreover, the socioeconomic impacts of floods in Bangladesh are multidimensional. The occurrence of flooding in Bangladesh is a regular annual phenomena. Floods affect crops, human settlements, livestock, physical structures and the infrastructures of the country and cause enormous damage to the livelihood of the people annually, especially in the areas near the rivers and sea side.

Four severe river floods was occurred in Bangladesh in 1987, 1988, 1998 and 2004 in the past. The magnitude of these floods were equivalent to 10-, 70-, 50- and 20-year return period of flood, respectively. Due to flooding up to a depth of just over 2 meters, in 1988, around 60% of land area in Bangladesh experienced severe damage while the low-lying areas were up to 5 meters under water. This flood caused severe damages and the total damage estimated was equivalent to BDT (Bangladeshi Taka) 781.2 million (US 11.16 million). In 1987, Bangladesh experienced the financial damage estimated BDT 347.2 million (US 4.96 billion) in 1987. In addition, several less severe flooding have been recorded in Bangladesh in 1954, 1955, 1958, 1970, 1979 and 1980 based on flood record, since 1945.

Cox's Bazar is one of the 64 districts of Bangladesh, located 150 kilometers south of Chittagong, the second largest city of Bangladesh. Cox's Bazar is situated along the Bay of Bengal and has the world's longest unbroken sandy sea beach (120 km). The district of Cox's Bazar covers an area of 2491.85 km^2 including 8 upazilas and 140 villages. The population of Cox's Bazar is 2,289,918, according to the census of 2011. Bakkhali is a seaside hamlet located in Maheshkhali upazila of Cox's Bazar. Bakkhali has a 7 km long beach starting from Maheshkhali, Bangladesh to Fraserganj, India. The total area of Bakkhali is 62 km^2 and the estimated population is 62,583. A detailed map of the case study area is illustrated in Fig. 15.



Map of South Asia (a)

Map of Bangladesh (b)

Map of Bakkhali (c)

Figure 15. Location of the case study area

Being a sea side island, Bakkhali experiences flooding in regular interval. Due to the severity of the flooding each year in the adjacent area of Bakkhali river, it has been chosen to study the consequence of flooding by applying BRBES developed as part of this research.

4.1.2 Survey Method

Survey is a method of data collection used to describe, compare or explain individual and social knowledge, feelings, values, preferences, and behavior [74]. Surveys can be of two types: 1) self-administered (conducted through email or web), 2) Interviews (conducted through phone or in person). In person interview was considered as the method of survey to collect data in this research which means that the survey was done by doing the field visit and the respondents were responsible for their activities, without others help.

Surveyors prefer online survey over offline survey because it is easy to conduct and processing data from online survey is easier as the responses can be downloaded to a spreadsheet for further processing. However, online survey needs electronics devices to connect with internet. Due to the unavailability of reliable internet access in the case study area, offline interview has been chosen as the method of surveying in this research. Fig. 16 shows two of the images taken during the interviews.

(a) (b)

Figure 16. Interviewing people in Bakkhali, Cox's Bazar, Bangladesh.

Cross-sectional study methodology was followed for this survey. Cross-sectional study is an observational study methodology that analyzes data collected from a population at a specific point of time [74]. These collected data allow cross-sectional studies with little or no expense which is a prominent advantage of this methodology. In a cross-sectional survey, a specific group is taken into consideration to see if any activity, for example flooding, is related to the effect being investigated, for example the livelihood. If flooding can be correlated with livelihood, this should support the hypothesis that flooding affects the livelihood.

4.1.3 Data Collection

In order to collect data from the case study area, a set of questionnaires has been prepared. The questions were prepared to get data related to the factors of flood risk assessment. Table 7 illustrates some of the questions asked during the interview as a part of data collection. The full list of questionnaires can be found in Appendix 1.

Table 7. Interview questionnaires fo the factors of flood risk assessment

Factor	Question	Discussion
Direct Intangible		
Percentage of Loss of Cattle Social Condition	What is the percentage of cattle washed away and/or died because of the flood? How was the social condition after the flood?	Refers to the percentage of loss of cattle during flood. Refers to the social condition as good, average or bad.
Direct Tangible		
Water Level Accommodation Problem	What was the water level? Was accommodation a problem after the flood?	Water level in terms of feet. Refers to the accommodation problem due to flood.
Indirect Intangible		
Financial Condition Mental Condition	How was the financial condition after flood? How was the mental condition after flood?	Refers to the financial condition as good, average or bad. Refers to the mental condition of the flood affected people.
Indirect Tangible		
Availability of Stuffs Frequency of Travelers	Was there availability of stuffs after the flood? What was the frequency of travelers after the flood?	Stuffs available for work after the flood. How frequently people traveled after the flood.

For example, to get data for the factor "Percentage of loss of cattle", the question was "What is the percentage of cattle washed away and/or died because of the flood?". This question gives a value expressed in percentage (e.g., 60%) which is quantitative in nature. On the other hand, while asking the question "How was the social condition after the flood?" for the factor "Social Condition" gives a value within a set of possible answers ["Good", "Average", "Bad"] which is qualitative in nature. Data for other factors have been collected in similar way using questionnaires.

Two hundred people from Bakkhali were interviewed as a part of this survey. Among them, 81% were male (162 persons) while the other 19% were female (38 persons). The reason behind the less percentage of female is, it is difficult to interview woman in rural areas due to the conservative social system in most of the villages in Bangladesh. The percentage could have been much higher if the interview would have been conducted in the city areas. The age of the respondents also varied. 31% (63 out of 200) people had age in the range of 18 to 30 years while 27.5% (55 out of 200) people were of age in range 31 to 44 years, 28.5% (57 people) were in the range of 45 to 65 years and the rest 12.5% (25 people) were of age between 66 to 90 as shown in Fig. 17 as a pie-chart.

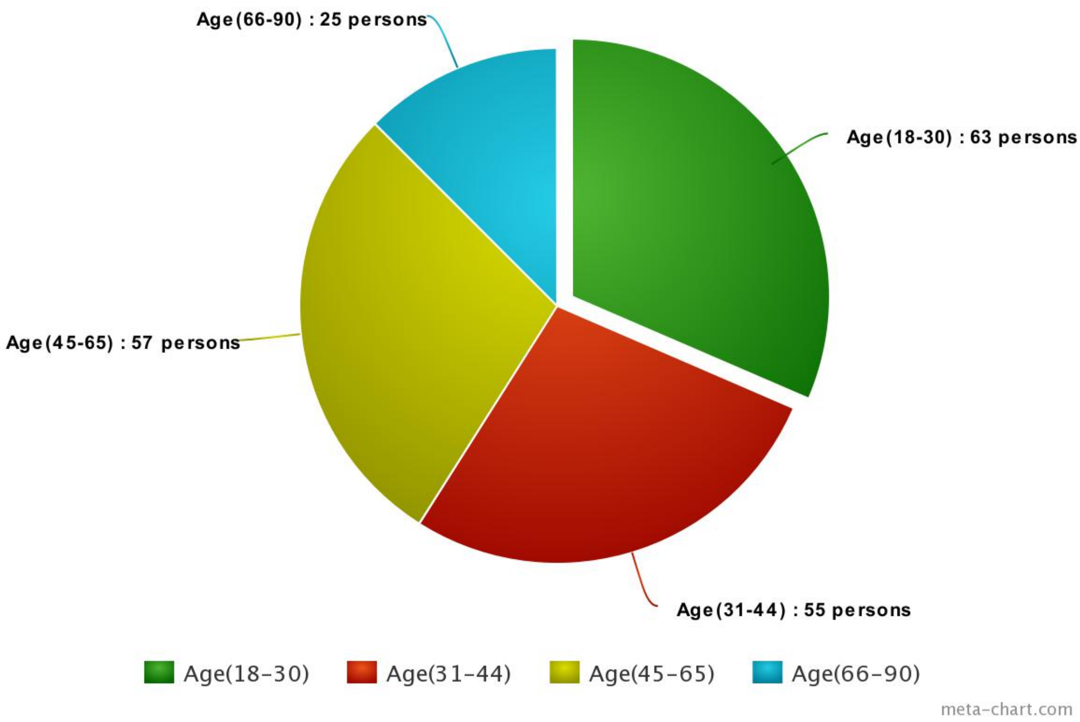


Figure 17. Age of the respondents being interviewed

Different age groups have been taken into consideration as it is believed that age is an important factor while considering expert opinions. People with age between 18-44 have less experience of being affected by flooding than the older ones and hence, considering aged people for expert opinion is a good idea to get reliable result.

These interviews were conducted with the same questionnaires by a group of four people in Bakkhali at the same day and before taking the interviews, the respondents were informed that these data should be used for research purpose. Due to the remoteness of the case study area, it was not possible to collect the data using any electronic devices such as mobile phones or laptops. Therefore, data was first collected in plain paper (by writing in hand) and then transfered into excel file for further processing.

4.2 System Validation

Data for the input nodes of the BRBES framework, mentioned in Fig. 8, have been collected from the people of the case study area by conducting interviews. These data have been used to assess the flood risk using the BRBES. Expert opinion has also been collected by using the same technique as elaborated in the previous section. The comparison, prediction, evaluation and assessment of the accuracy of the results obtained from the different models are also considered, which as an important aspect to measure the reliability of a system. Receiver Operator Characteristic (ROC) curves are widely used to assess, evaluate, predict and compare the performance of different methods. It provides a comprehensive and visual method of summarizing the accuracy of evaluation, assessment, prediction and comparison of a system. Thus, ROC curves have become a prominent tool for evaluating different models, algorithms and techniques in various domains such as biomedical engineering, machine learning, health informatics and many others [72], [75], [76]. ROC curves were used to measure the accuracy of the flood risk assessment system using BRBES in this research. In ROC curves, the accuracy can be measured by calculating the size of Area Under Curve (AUC) [77]. The accuracy of the system is directly proportional to the AUC. If the AUC is 1 then it means that the system result is 100% accurate.

The validation or the reliability of the results have been carried out for the BRBES itself as for the optimal learning model proposed in this research. At first, the performance of initial BRBES is evaluated using ROC curves. Then the trained BRB expert system is evaluated using ROC curves to compare the initial and trained BRB expert system result. Both parts of the system validation is described in the following parts.

4.2.1 Initial BRB Expert System

In order to evaluate the performance of the initial BRBES, five factors, namely "Direct Tangible", "Direct Intangible", "Indirect Tangible", "Indirect Intangible" and "Livelihood" (Overall BRBES for flood risk assessment), have been considered. Experts perception on the flood risk assessment data have been considered as the standard for the comparison between BRBES system generated result and expert opinion. A total number of data points collected from the case study areas is 200 for this work. These sample data is considered to be sufficient as sample sizes of more than 30 and less than 500 are appropriate for most of the research [78]. Table 8 illustrates some of the data points for the factor "Direct Intangible" from the dataset. Column 4 of the Table 8 shows the BRBES generated output in percentage which is evaluated using the Eq. 17 while column 5 illustrates experts opinion in percentage for a set of data collected from the case study area during interviewing.

Table 8. System validation for the factor "Direct Intangible"

Serial	Percentage of Loss of Cattle	Social Condition	System Result (%)	Expert Opinion (%)	Benchmark
1	20	Low	0.5	1	1
2	23	Medium	0.75	0.5	0
3	10	Low	0.5	0	0
4	17	Low	0.5	1	1
5	19	Low	0.5	1	1
.	..	..	..	..	..
.	..	..	..	..	..
196	15	Low	0.5	0.5	0
197	20	Low	0.5	0.5	0
198	19	Low	0.5	1	1
199	25	Low	1	1	1
200	27	Low	0.5	1	1

Table 9. System validation for the factor "Direct Tangible"

Serial	Area	Crop	Water Level	Availability of Cattle Food	Accommodation	Availability of Transport	Length of Road Affected	Road Damage	Duration of Water Standing	System Result (%)	Expert Opinion (%)	Benchmark
1	8.0	HIGH	3.0	0.33	1	0.33	10	43.75	2.5	0.091	0.52198	0
2	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.1	0.46448	0
3	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.1	0.46448	1
..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..
198	10.0	LOW	3.0	0.33	1	0.88	10	43.75	3.0	0.105	0.49605	1
199	9.0	LOW	3.5	0.5	1	0.88	12	43.75	2.5	0.092	0.52203	0
200	10.0	LOW	3.5	0.5	1	0.33	12	43.75	3.0	0.092	0.52203	1

Table 10. System validation for the factor "Indirect Intangible"

Serial	Financial Condition	Mental Condition	System Result (%)	Expert Opinion (%)	Benchmark
1	Low	Low	0.25	1	1
1	Low	Low	0.027	0.5	0
1	Low	Low	0.107	1	1
1	Low	Low	0.058	0.5	0
.	..	..	..	..	..
.	..	..	..	..	..
196	Low	Low	0.939	0.5	0
197	Low	Low	0.571	1	1
198	Low	Low	0.027	0.5	0
199	Low	Low	0.571	1	1
200	Low	Low	0.558	1	1

Table 11. System validation for the factor "Indirect Tangible"

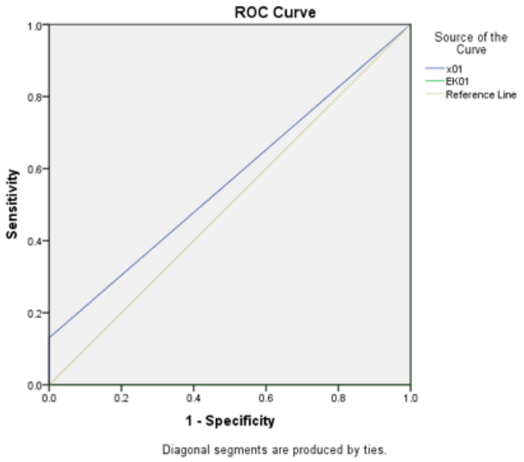
Serial	Transport	Availability of Stuffs	System Result (%)	Expert Opinion (%)	Benchmark
1	0	0.5	0.571	0.5	0
2	0	0	0.374	1	1
3	0.3	0.3	0.486	1	1
4	0.5	0.3	0.697	1	1
5	0	0	0.571	0.5	0
.	..	..	..	..	..
.	..	..	..	..	..
196	0	0.2	0.444	0.5	0
197	0.3	0.5	0.697	1	1
198	0	0.3	0.486	1	1
199	0.5	0.4	0.713	1	1
200	0	0.5	0.571	0.5	1

Table 12. System validation for the overall BRBES for flood risk assessment

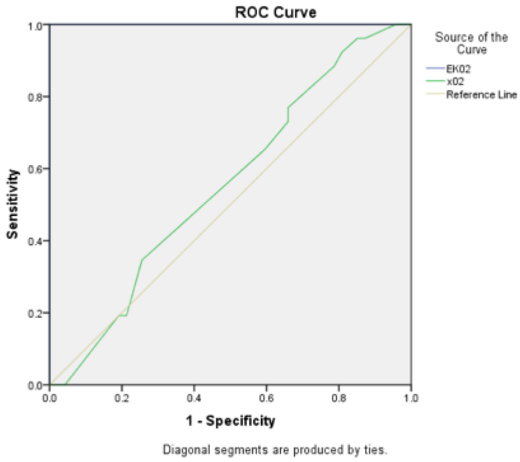
Serial	Direct Intangible	Direct Tangible	Indirect Intangible	Indirect Tangible	System Result (%)	Expert Opinion (%)	Benchmark
1	0.5	0.91	0.25	0.571	0.287	1	1
2	0.75	0.1	0.027	0.374	0.243	0.5	0
3	0.75	0.1	0.027	0.374	0.243	1	1
4	0.5	0.083	0.107	0.486	0.221	1	1
5	0.75	0.097	0.027	0.374	0.243	0.5	0
.	..	..	..	..	..	..	..
.	..	..	..	..	..	..	..
196	0.5	0.105	0.027	0.558	0.215	1	1
197	0.75	0.101	0.027	0.374	0.243	0.5	0
198	0.5	0.105	0.0393	0.546	0.214	1	1
199	0.5	0.092	0.027	0.583	0.221	0.5	1
200	0.5	0.092	0.558	0.571	0.349	1	1

Tables 9, 10, 11 and 12 illustrate the dataset for the factors "Direct Tangible", "Indirect Intangible", "Indirect Tangible" and "Livelihood" (overall BRBES for flood risk assessment) respectively with both system generated values and expert opinions.

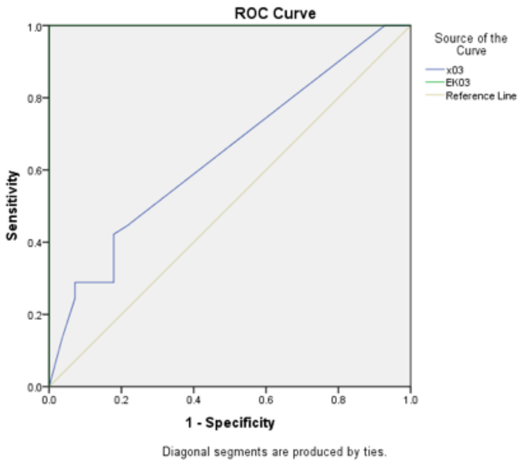
ROC curves for "Direct Intangible" (x1), "Direct Tangible" (x2), "Indirect Intangible" (x3), "Indirect Tangible" (x4) and "Livelihood" (x5), as mentioned earlier, are illustrated in Fig. 18 (a - e). There are two curves with two different colors in the ROC curve as shown in Fig. 18. The curve with green line illustrated the expert opinion and curve with the blue line represents the BRBES system generated result.



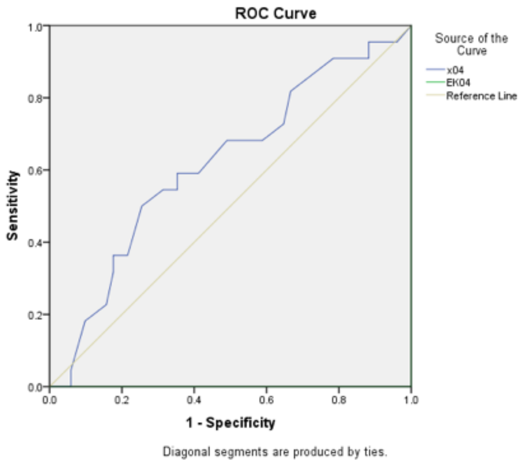
(a)



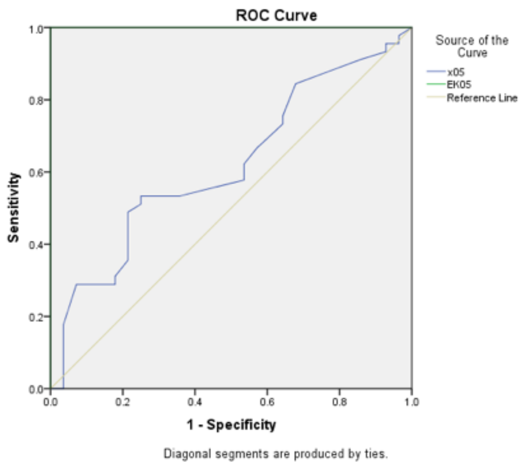
(b)



(c)



(d)



(e)

Figure 18. ROC curves of the initial BRB expert system for (a) x1, (b) x2 (c) x3, (d) x4 and (e) x5 (Overall BRBES for flood risk assessment)

The AUC and the confidence interval (CI) for above mentioned factors for initial BRBES are represented in Table 13. The AUC of "Indirect Intangible" (x3), "Indirect Tangible" (x4) and "Livelihood" (x5) are in a close range of each other which are 0.643, 0.617, 0.619, respectively, while the AUC for the node "Direct Intangible" (x1) and "Direct Tangible" (x2) is 0.565 and 0.556, respectively, as shown in Table 13. The AUC values demonstrate the reliability of the initial BRB inference procedures. However, the AUC values, as shown in Table 13, are not much closer to "1", meaning the performance of the initial BRBES needs to be improved. For this reason, the research proposed the addition of optimal learning component with the initial BRBES. This will allow the improvement of the BRBES performance and will be demonstrated in the next section.

Table 13. AUC and CI evaluated for initial BRB for the factors x1, x2, x3, x4 and x5

Results	Direct Intangible (x1)	Direct Tangible (x2)	Indirect Intangible (x3)	Indirect Tangible (x4)	Livelihood (x5)
AUC	0.565	0.556	0.643	0.617	0.619
CI	0.417 - 0.713	0.422 - 0.690	0.515 - 0.771	0.476 - 0.759	0.488 - 0.750

The reason behind considering mid-level nodes for ROC is that mid-level components can be useful to create 'what-if' scenarios which will help to understand the risk of flooding on those components during real time case study.

4.2.2 Optimization

In order to evaluate the performances of the trained BRBES, same five factors, namely "Direct Intangible", "Direct Tangible", "Indirect Intangible", "Indirect Tangible" and "Livelihood" (Overall BRBES for flood risk assessment), have been considered. Experts perception on the flood risk assessment has been considered as the standard for the comparison between BRBES system generated result and expert opinion as it was used for the initial BRB.

The BRB framework, mentioned in the previous chapter, has been modified for the training module. The modification of the initial BRB framework is considered due to the excessive time complexity of the training module for all the factors of "Direct Tangible". To reduce the time complexity, four children of the node "Direct Tangible" have been taken and shifted as children of a new node name "Road" which is now a child of "Direct Tangible" after modification as shown in Fig. 19.

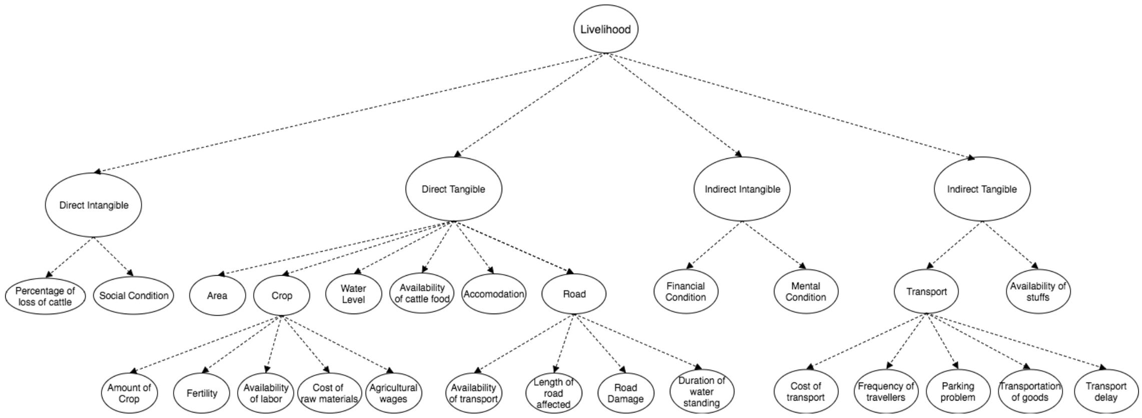


Figure 19. BRB framework for optimization

Table 14 illustrates the trained BRB for the factor "Direct Intangible" from where column 5, 6 and 7 illustrates the generated output (referential values) which is calculated using the Eq. 17.

Table 14. Trained BRB for the factor "Direct Intangible"

		IF		THEN(Direct Intangible)		
Rule Identifier	Rule Weight	Percentage of Loss of Cattles	Social Condition	High	Medium	Low
R1	1	HIGH	HIGH	1.0	0	0
R2	1	HIGH	MEDIUM	0.08	0.5	0.42
R3	1	HIGH	LOW	0.77	0.16	0.07
R4	1	MEDIUM	HIGH	0.5	0.5	0
R5	1	MEDIUM	MEDIUM	0	1.0	0
R6	1	MEDIUM	LOW	0	0.5	0.5
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0	0	1.0

Table 15. Trained BRB for the factor "Direct Tangible"

		IF									THEN(Direct Tangible)		
Rule Identifier	Rule Weight	Area	Crop	Water Level	Availability of Cattle Food	Accommodation	Availability of Transport	Length of Road Affected	Road Damage	Duration of Water Standing	High	Medium	Low
R1	1	8.0	HIGH	3.0	0.33	1	0.33	10	43.75	2.5	1.0	0	0
R2	1	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.94	0.06	0
R3	1	9.0	HIGH	5.0	0.33	1	0.66	8	37.5	3.0	0.88	0.12	0
..	..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..	..	..	..	..	..
R19681	1	10.0	LOW	3.0	0.33	1	0.88	10	43.75	3.0	0	0.12	0.88
R19682	1	9.0	LOW	3.5	0.5	1	0.88	12	43.75	2.5	0	0.06	0.94
R19683	1	10.0	LOW	3.5	0.5	1	0.33	12	43.75	3.0	0	0	1

Table 16. Trained BRB for the factor "Indirect Intangible"

		IF		THEN(Indirect Intangible)		
Rule Identifier	Rule Weight	Financial Condition	Mental Condition	High	Medium	Low
R1	1	HIGH	HIGH	1.0	0	0
R2	1	HIGH	MEDIUM	0.5	0.5	0
R3	0.24	HIGH	LOW	0.92	0.08	0
R4	1	MEDIUM	HIGH	0.5	0.5	0
R5	1	MEDIUM	MEDIUM	0	1.0	0
R6	0.48	MEDIUM	LOW	0.56	0.55	0
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0.77	0	0.23

Table 17. Trained BRB for the factor "Indirect Tangible"

		IF		THEN(Indirect Tangible)		
Rule Identifier	Rule Weight	Transport	Availability of Stuffs	High	Medium	Low
R1	0.23	HIGH	HIGH	0	0	1.0
R2	1	HIGH	MEDIUM	1.0	0	0
R3	1	HIGH	LOW	0	1.0	0
R4	0.45	MEDIUM	HIGH	1.0	0	0
R5	0.09	MEDIUM	MEDIUM	0.42	0.26	0.32
R6	1	MEDIUM	LOW	0	0.5	0.5
R7	1	LOW	HIGH	0	1.0	0
R8	1	LOW	MEDIUM	0	0.5	0.5
R9	1	LOW	LOW	0	0	1.0

Table 18. Trained BRB for the factor "Livelihood"

		IF				THEN(Livelihood)		
Rule Identifier	Rule Weight	Direct Intangible	Direct Tangible	Indirect Intangible	Indirect Tangible	High	Medium	Low
R1	1	HIGH	HIGH	HIGH	HIGH	1.0	0	0
R2	1	HIGH	HIGH	HIGH	MEDIUM	0.75	0.25	0
R3	1	HIGH	HIGH	HIGH	LOW	0.5	0.5	0
R4	0.91	HIGH	HIGH	MEDIUM	HIGH	1.0	0	0
R5	0.99	HIGH	HIGH	MEDIUM	MEDIUM	1.0	0	0
..	..	..	..	..	..	..	..	..
..	..	..	..	..	..	..	..	..
R77	1	LOW	LOW	MEDIUM	MEDIUM	0	0.5	0.5
R78	1	LOW	LOW	MEDIUM	LOW	0	0.25	0.75
R79	1	LOW	LOW	LOW	HIGH	0	0.5	0.5
R80	1	LOW	LOW	LOW	MEDIUM	0	0.25	0.75
R81	1	LOW	LOW	LOW	LOW	0	0	1.0

Trained BRB for "Direct Tangible", "Indirect Intangible", "Indirect Tangible" and "Livelihood" are illustrated in Tables 15, 16, 17, 18 respectively.

ROC curves for "Direct Intangible" (x1), "Direct Tangible" (x2), "Indirect Intangible" (x3), "Indirect Tangible" (x4) and "Livelihood" (x5) generated using the trained BRBES and are illustrated in Fig. 20 (a - e). The ROC curve with green line illustrates the real system result and the blue line illustrates the curve for BRBES result as shown in Fig. 20.

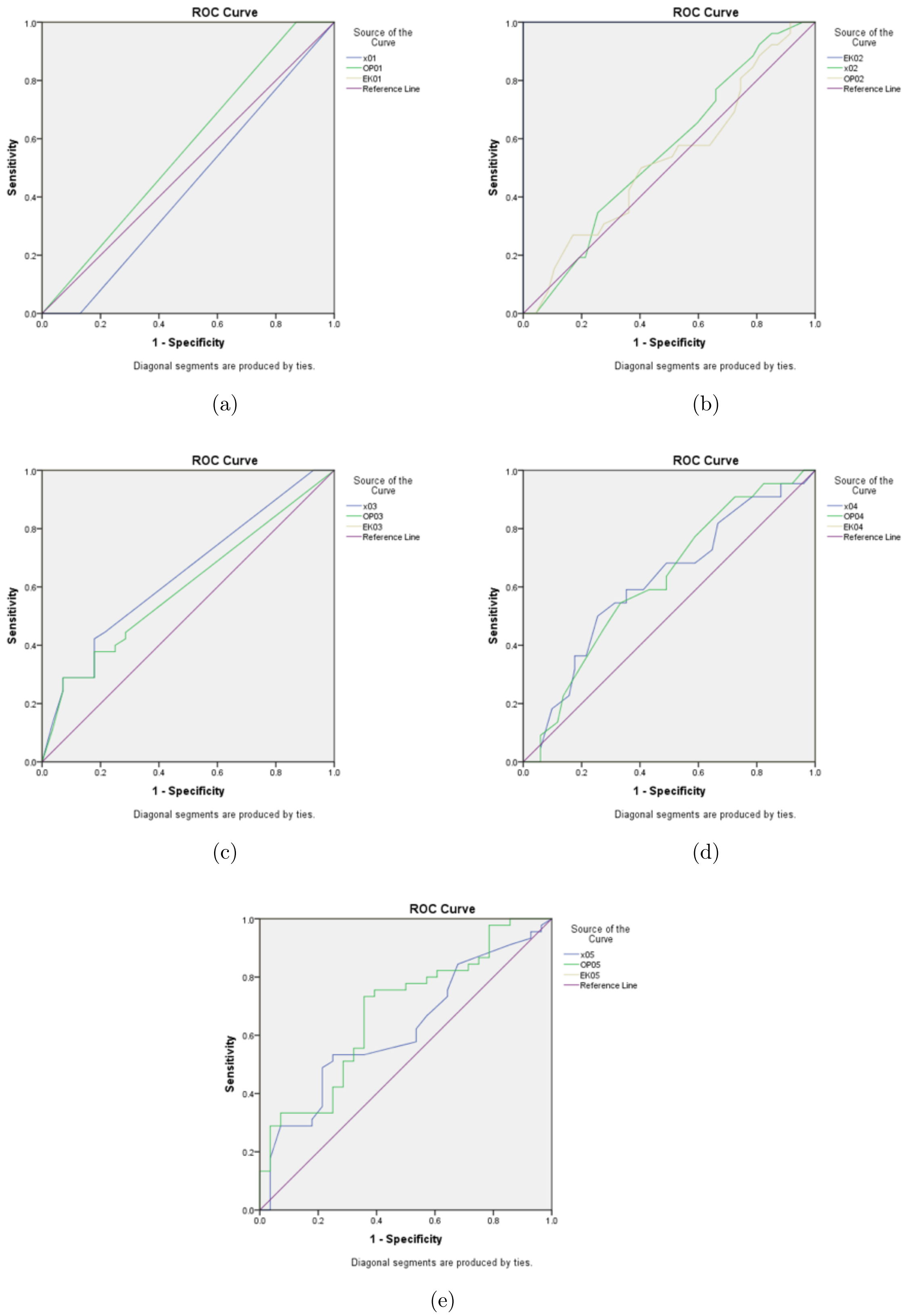


Figure 20. ROC curves of the trained BRB expert system after optimization for (a) x1, (b) x2, (c) x3, (d) x4 and (e) x5 (Overall BRBES for flood risk assessment)

The AUC and CI for the five factors for trained BRBES are shown in Table 19. It can be observed from the results shown in Table 19 that the AUC of "Direct Intangible" (x1), "Direct Tangible" (x2), "Indirect Intangible" (x3), "Indirect Tangible" (x4) and "Livelihood" (x5) are 0.571, 0.612, 0.664, 0.693 and 0.651, respectively, for the trained BRBES as shown in Table 19.

Table 19. AUC and CI evaluated for trained BRB for the factors x1, x2, x3, x4 and x5

Results	Direct Intangible (x1)	Direct Tangible (x2)	Indirect Intangible (x3)	Indirect Tangible (x4)	Livelihood (x5)
AUC	0.571	0.612	0.664	0.693	0.651
CI	0.417 - 0.713	0.390 - 0.668	0.466 - 0.728	0.487 - 0.758	0.555 - 0.807

Table 20 shows the comparison of the AUC's calculated for initial BRB and trained BRB. It is observable that the AUC of "Direct Intangible" (x1) was 0.565 in initial BRB which has slightly been improved to 0.571 in trained BRB. AUCs for other factors i. e., "Direct Tangible" (x2), "Indirect Intangible" (x3), "Indirect Tangible" (x4) and "Livelihood" (x5) have been improved significantly in trained BRB compared to the initial BRB. This improvement will continue if the training module train with more training dataset.

Table 20. Comparison of AUC calculated for initial and trained BRB for the factors x1, x2, x3, x4 and x5

AUC	Direct Intangible (x1)	Direct Tangible (x2)	Indirect Intangible (x3)	Indirect Tangible (x4)	Livelihood (x5)
Initial BRB	0.565	0.556	0.643	0.617	0.619
Trained BRB	0.571	0.612	0.664	0.693	0.651

From the results in Table 20, it can be clearly visible that the trained BRB performs better than the initial BRB and hence, it is possible to come to a conclusion that the learning module incorporates intelligence in the BRBES to perform better.

4.3 Summary

This chapter described how the case study area was selected. It also discussed about the methodology of data collection for the input variables of the flood risk assessment expert system. It represented the way of validating the data along with the procedure of validating the system, using the results evaluated from the BRBES, using ROC curves. Next chapters discusses about the different perspectives of this research.

5 DISCUSSION

This chapter presents a discussion of the risk assessment framework for analyzing and reducing the risk of flooding. It also highlights the achievements and the novelty along with the sustainability aspect of this research.

5.1 Risks of Flooding

Flooding is one of the most devastating forms of natural disaster for its frequency of happening as well as the greater death toll and destruction of property than any other form of disaster [2]. The risk of flooding on an object acts as a source for the effect of flooding on a series of other objects in a society, acting in a ripple fashion. In this way multiplication of the flood risk takes place.

Livelihood is an integral part of a society and is often termed as 'socio-economic life-line'. Flooding causes several damages in the regular life of the people and hence, this research focused on livelihood. Factors of flood risk assessment has been identified from the flood damage model described in section 2.2.1. This model has classified the factors into four different categories, namely Direct Tangible (DT), Direct Intangible (DI), Indirect Tangible (IT) and Indirect Intangible (II), depending on various factors as shown in Fig. 4. This model has given the domains by which factors can be selected and classified into categories.

Different types of uncertainties associated to the factors has been identified. Uncertainties such as ignorance, incompleteness, ambiguity, imprecision and vagueness can be caused by lack of human knowledge or incompleteness in data. It was also necessary to identify the uncertainties as they cause inaccuracy in an expert system. Some uncertainties were identified by literature review while others were found from the interview data and they were illustrated in Table 1 with operational definition of each factor.

Different flood risk assessment methods have been described in section 2.4 which were already been used for assessing and reducing flood risk in different parts of the world such as in USA, Europe and south Asia. Examples of the real flood has been provided while describing these methodologies along with the damage scenario and aftermaths. As most of those methods dealt with numerical data and did not handle any uncertainty associated with the data, several numerical and symbolic methods for representing and reasoning with uncertainties has also been discussed. From that discussion, BRB has been selected over other methodologies because of it's ability to handle heterogeneous (qualitative and quantitative) data as well as all types of uncertainties associated with the factors of this research.

5.2 Expert System Components for Flood Risk Assessment

The aim of this thesis was to develop a solution to the risk of flooding on livelihood by visualizing and analyzing flood scenarios, and supporting the decision-making processes related to the reduction of flood risk. This aim was achieved by developing a belief rule based expert system that has the ability to generate various 'what-if' scenarios of floods and assess their risk on livelihood. This expert system had two main components, a knowledge base and inference engine as described in section 3.1. Several methods for knowledge acquisition such as literature review, historical data and interviews have also been described in section 3.1.1. Interview has been chosen over literature review and historical data due to the lack of available information from these methods of knowledge acquisition. Different components of the inference procedure have been described in section 3.1.2 along with all necessary equations and explanations. BRBES learning methodology has also been described in section 3.1.3 with equations and the learning module framework.

A system is said to be intelligent if it has knowledge as well as the learning capability at the same time [79]. This expert system for flood risk assessment is said to be intelligent as it has the knowledge base and due to the incorporation of the learning module, it can learn itself to increase it's performance.

5.3 System Architecture

Different components of the system architecture for the BRB expert system has been described in section 3.2. Architecture of the system has been shown in Fig. 7 which consists of four different layers namely data management layer, application layer, interface layer and API layer. BRB framework in form of a BRB tree has been illustrated in Fig. 8 while initial knowledge bases for different sub-trees has been shown in Tables 2, 4, 5 and 6.

Different data formats such as CSV, XML and JSON have been described in section

3.2.3 along with their advantages and disadvantages. JSON has been chosen as the data format for the BRBES over CSV or XML due to it's lightweight and flexible data interchange ability for web based system. JSON data for a node in BRB framework has been illustrated in Fig. 9 along with description of each of the variables needed for a node in the tree.

A dynamic tree traversal algorithm has been introduced and described with a pseudo code in section 3.2.4. The algorithm has been developed as a variant of traditional top-down BFS algorithm approach and it works in a generic way that it can be used for any BRB tree from any domain.

The system was then converted into a web-based application. A GUI for the web-based application has also been developed which was described in section 3.2.7. A RESTful API has been built for the portability of the system. RESTful API's are very popular and widely used now-a-days and give a way to interact with an application without even knowing much about the back-end system. The creation and application of this API were described in section 3.2.6. Both RESTful API and the web-based GUI have been developed using python programming language. The architecture of the learning module of the BRBES, introduced in section 3.1.3, has been described along with a Fig. 14 in section 3.2.8 which has been developed using Matlab. The full list of tools used and built in this research along with the necessary URL's can be found in Appendix 2.

5.4 Case Study

For implementing the system in real world scenario to evaluate the system performance, Bakkhali area in Cox's Bazar district of Bangladesh has been chosen as the case study due to the frequency of flooding in that area. Interviews have been chosen over online survey as the method of data collection due to the remoteness and lack of availability of internet in the case study area. Questionnaires have been created to collect data for each of the factors mentioned earlier and the interview was conducted for 200 people in the case study area. Data was collected as hand written text and then transfered in to excel sheet for further processing. The full list of the questionnaires can be found in Appendix 1.

5.5 System Validation

Data for the input nodes of the BRB framework along with the expert opinions has been collected from the interviews in the case study area and then BRBES has been applied on the data to evaluate result. The system validation methodology had two parts, namely initial BRB validation and trained BRB validation. First, the performance of initial BRBES has been evaluated in section 4.2.1. The dataset for the system testing with initial BRB can be found in Tables 8, 9, 10, 11, 12. Then the learning module has been applied on the data to train the BRB and it's performance has been evaluated. The trained BRB's are shown in Tables 14, 15, 16, 17 and 18 in section 4.2.2.

ROC curves have been used to evaluate the result which can be found in Fig. 18 (a - e) for initial BRB and in Fig. 20 (a-e) for trained BRB. In the end, the output of the evaluation from these two parts has been compared to validate the system accuracy which can be found in Table 20.

5.6 Sustainability Aspects

Sustainability and green IT concepts are gaining attention in the modern era. According to Zadeh [80], sustainability implies that you are happy with the situation as it is and you want it to continue. Any natural disaster hinders the regular activity of human life and hence, hampers sustainability.

It seems that the current flood risk assessment systems, as described in section 2.4, focuses mainly on economic damages. Moreover, social and environmental effects of flooding are often not considered or considered in a very limited manner. This is partly because they are not easily measurable in quantitative terms and hence, not comparable with economic damages. Due to the over sighting of the intangible parts of flooding, flood risk assessment is often incomplete and hence biased. For this reason, an aggregation of social, economic and environmental factors is important to access both tangible and intangible consequences of flood. Moreover, existing flood risk assessment systems do not take different sustainability aspects in concern other than economic sustainability. This research has focused on these problems.

If the flood risk assessment scenario is considered, both tangible and intangible factors have been taken into consideration in this research. The livelihood of flood affected neighborhood is affected badly due to flooding and it hinders the sustainability of that area. In this research, social and environmental aspects of flooding have been taken into consideration along with the economical damage and hence, it gives a combination of social, economic and environmental perspective of flooding which is necessary for social well-being of the people. Due to the flexibility of having the "what-if" scenarios of different types of risk factors, it is convenient for decision makers to focus on different aspects of the flood risk and in this way, it can be used to save numerous number of life as well as reduce the damage of flooding.

The idea of sustainable development is to use the available resources in such a way then it can efficiently fulfill the needs of present generation without compromising the needs of future. The Erasmus Mundus masters program in PERvasive Computing and COMmunication for sustainable development (PERCCOM) also focuses mainly on sustainable development [81] [82]. Therefore, the flood risk assessment system provided by this research helps in achieving sustainability. By using the flood risk assessment system, it is possible to save life, nature and this will also decrease the effect of flooding. Furthermore, proper steps can be taken beforehand by the authorities which also leads to financial savings. This ultimately leads to reduce the costs for the infrastructure building as well as urban planning. Thus it can contribute to sustainable and energy efficient cities.

Summary

This chapter discussed about the works done to build the risk assessment framework for analyzing and reducing the risk of flooding. It has also highlighted the achievements and novelty of this research along with the sustainability aspects. Next chapter will conclude this research keeping the limitations into consideration along with some suggestions for possible future works.

6 CONCLUSION AND FUTURE WORK

6.1 Conclusion

The main purpose of this research was to develop an expert system to assess the risk of flooding while taking the uncertainties in data into consideration. It has been demonstrated that the BRBES can be used to analyze, visualize and support decision making processes that can reduce the flood risk on livelihood. The case study has given a clear idea about how to implement an expert system in real world. The methodology utilized by the risk assessment framework differs significantly from the existing flood risk assessment methodologies. Moreover, considering heterogeneous data for the expert system made it more reliable and widely usable compared to the other existing flood risk assessment expert systems. Different mid level nodes have been considered for analyzing the expert system which also has given the flexibility to emphasize on different sub-domains, as well as the overall livelihood of the area, for the decision makers. This flood risk assessment system is intelligent as it has the knowledge base, inference engine as well as the training module which gives the system the capability of learning itself.

Students from Bangladesh, who are using BRB for different domain, has already used the tree traversal algorithm for their BRB trees as well as the API built as a part of this research and hence, it can be said that this research contributed in the field of risk assessment in a great manner. Data used for this risk assessment framework was collected by conducting interviews due to the lack of data from other resources such as historical data. Moreover, some mid levels nodes have not been considered for the evaluation of result due to lack of time.

6.2 Future Work

All the mid-level nodes can be considered in near future while evaluating the result which will give more clear view of the risk scenario. Based on the limitations of the system, as mentioned earlier, other data collection methodologies can also be used for the system as a part of further research. Only age group has been taken in concern while concerning expert opinions due to the time limitations. Other factors can also be considered in near future to bring more effective and valuable outcomes.

Pervasive computing is becoming an integral part of today's information system so it can be introduced as a source of data for the BRBES. IoT devices such as sensors can also be used as a part of data collection. Sensors data can be used for some of the factors instead of using data from interviews to improve the system performance. If the volume of data is too big to handle for single machine, big data technologies such as Hadoop or Spark can also be integrated with the expert system to analyze big volume of data.

REFERENCES

[1] A. Askew, “Water in the international decade for natural disaster reduction,” *IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences*, vol. 239, pp. 3–12, 1997.

[2] L. Cunha, W. Krajewski, and R. Mantilla, “A framework for flood risk assessment under nonstationary conditions or in the absence of historical data,” *Journal of Flood Risk Management*, vol. 4, no. 1, pp. 3–22, 2011.

[3] M. Shahjahan, “Flood disaster management in deltaic plain integrated with rural development,” *Bangladesh floods: Views from home and abroad*, 1998.

[4] T. Estrela, M. Menéndez, M. Dimas, C. Marcuello, G. Rees, G. Cole, K. Weber, J. Grath, J. Leonard, N. B. Ovesen *et al.*, *Sustainable water use in Europe. Part 3: Extreme hydrological events: floods and droughts.* European Environment Agency, 2001.

[5] A. K. Gain, C. Giupponi, and F. G. Renaud, “Climate change adaptation and vulnerability assessment of water resources systems in developing countries: a generalized framework and a feasibility study in Bangladesh,” *Water*, vol. 4, no. 2, pp. 345–366, 2012.

[6] J.-B. Yang, J. Liu, J. Wang, H.-S. Sii, and H.-W. Wang, “Belief rule-base inference methodology using the evidential reasoning approach-rimer,” *IEEE Transactions on Systems, Man, and Cybernetics-part A: Systems and Humans*, vol. 36, no. 2, pp. 266–285, 2006.

[7] J. P. Kahan, *From Flood Control to Integrated Water Resource Management: Lessons for the Gulf Coast from flooding in other places in the last sixty years.* Rand Corporation, 2006, vol. 164.

[8] S. A. Changnon, *The great flood of 1993: Causes, impacts, and responses.* Westview Press, 1996.

[9] B. Taylor, “Lessons from last floods hold water today,” 2001.

[10] J. De Vries, “A different approach to water: water management policy in the 21st century,” Ministerie V&W/Verkeer & Waterstaat, Tech. Rep., 2000.

[11] Y. Zong and X. Chen, “The 1998 flood on the yangtze, china,” *Natural Hazards*, vol. 22, no. 2, pp. 165–184, 2000.

[12] J. Seaman, S. Leivesley, and C. Hogg, *Epidemiology of natural disasters.* Karger, 1984, vol. 5.

[13] M. Ahmad, *Flood in Bangladesh.* Community Development Library, 1989.

[14] P. Kungwani, "Risk management-an analytical study," *IOSR Journal of Business and Management*, vol. 16, no. 3, pp. 83–89, 2014.

[15] P. Cline, "The merging of risk analysis and adventure education," in *Wilderness Risk Management Conference: Banff, Canada*, 2004.

[16] E. P. Evans and E. C. Penning-Rowsell, *Future flooding and coastal erosion risks.* Thomas Telford, 2007.

[17] E. J. Vaughan and T. Vaughan, *Fundamentals of risk and insurance.* John Wiley & Sons, 2007.

[18] D. D. Hart, T. E. Johnson, K. L. Bushaw-Newton, R. J. Horwitz, A. T. Bednarek, D. F. Charles, D. A. Kreeger, and D. J. Velinsky, "Dam removal: challenges and opportunities for ecological research and river restoration: we develop a risk assessment framework for understanding how potential responses to dam removal vary with dam and watershed characteristics, which can lead to more effective use of this restoration method," *BioScience*, vol. 52, no. 8, pp. 669–682, 2002.

[19] B. Merz, H. Kreibich, R. Schwarze, and A. Thieken, "Review article" assessment of economic flood damage"," *Natural Hazards and Earth System Sciences*, vol. 10, no. 8, p. 1697, 2010.

[20] R. Antunes and V. Gonzalez, "A production model for construction: A theoretical framework," *Buildings*, vol. 5, no. 1, pp. 209–228, 2015.

[21] D. P. Thunnissen, "Uncertainty classification for the design and development of complex systems," in *3rd annual predictive methods conference*, 2003, pp. 16–17.

[22] S. Mohammadi, M. Nazariha, and N. Mehrdadi, "Flood damage estimate (quantity), using hec-fda model. case study: the neka river," *Procedia Engineering*, vol. 70, pp. 1173–1182, 2014.

[23] N. R. Councils, *Levees and the National Flood Insurance Program: Improving Policies and Practices.* National Academies Press, 2013.

[24] I. Kelman and R. Spence, "An overview of flood actions on buildings," *Engineering Geology*, vol. 73, no. 3, pp. 297–309, 2004.

[25] P. Grossi, *Catastrophe modeling: a new approach to managing risk.* Springer Science & Business Media, 2005, vol. 25.

[26] J. S. Holladay and J. A. Schwartz, "Flooding the market: The distributional consequences of the nfip," *Institute for Policy Integrity, Policy Brief*, no. 7, 2010.

[27] G. H. Leavesley, *DESTRUCTIVE WATER: Water-Caused Natural Disasters, their Abatement and Control.* IAHS Press, 1997, no. 239.

[28] S. Ruman, "Comparison of two flood risk assessment methods in the case of the turiec river, slovakia," *AUC GEOGRAPHICA*, vol. 49, no. 2, 2014.

[29] K. Karagiorgos, "Flood hazard assessment validation based on the floods risk directive 2007/60/ec–a case study in rafina (attica, greece) catchment."

[30] S. J. Henkind and M. C. Harrison, "An analysis of four uncertainty calculi," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 18, no. 5, pp. 700–714, 1988.

[31] S. Parsons, "Current approaches to handling imperfect information in data and knowledge bases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 8, no. 3, pp. 353–372, 1996.

[32] E. Turban and J. Aronson, "Expert systems and intelligent systems," 2001.

[33] R. Sun, "Robust reasoning: integrating rule-based and similarity-based reasoning," *Artificial Intelligence*, vol. 75, no. 2, pp. 241–295, 1995.

[34] L. A. Zadeh, "Knowledge representation in fuzzy logic," in *An introduction to fuzzy logic applications in intelligent systems.* Springer, 1992, pp. 1–25.

[35] J.-B. Yang and M. G. Singh, "An evidential reasoning approach for multiple-attribute decision making with uncertainty," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 1, pp. 1–18, 1994.

[36] J.-B. Yang and P. Sen, "A general multi-level evaluation process for hybrid madm with uncertainty," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 10, pp. 1458–1473, 1994.

[37] J.-B. Yang, "Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties," *European Journal of Operational Research*, vol. 131, no. 1, pp. 31–61, 2001.

[38] J. Liu, J.-B. Yang, J. Wang, H.-S. SII, and Y.-M. Wang, "Fuzzy rule-based evidential reasoning approach for safety analysis," *International Journal of General Systems*, vol. 33, no. 2-3, pp. 183–204, 2004.

[39] A. P. Dempster, "A generalization of bayesian inference," in *Classic works of the dempster-shafer theory of belief functions.* Springer, 2008, pp. 73–104.

[40] E. Binaghi and P. Madella, "Fuzzy dempster–shafer reasoning for rule-based classifiers," *International Journal of Intelligent Systems*, vol. 14, no. 6, pp. 559–583, 1999.

[41] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Information Sciences*, vol. 8, no. 4, pp. 301–357, 1975.

[42] R. Ul Islam, K. Andersson, and M. S. Hossain, "A web based belief rule based expert system to predict flood," in *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services.* ACM, 2015.

[43] M. S. Hossain, K. Andersson, and S. Naznin, "A belief rule based expert system to diagnose measles under uncertainty," pp. 17–23, 2015.

[44] R. U. Islam, M. S. Hossain, and K. Andersson, "A novel anomaly detection algorithm for sensor data under uncertainty," *Soft Computing*, pp. 1–17, 2016. [Online]. Available: http://dx.doi.org/10.1007/s00500-016-2425-2

[45] M. S. Hossain, P. Zander, M. S. Kamal, and L. Chowdhury, "Belief-rule-based expert systems for evaluation of e-government: a case study," *Expert Systems*, vol. 32, no. 5, pp. 563–577, 2015.

[46] M. S. Hossain, M. E. Hossain, M. S. Khalid, and M. A. Haque, "A belief rule-based (BRB) decision support system for assessing clinical asthma suspicion," in *Proceedings of the Scandinavian Conference on Health Informatics*, 2014, pp. 83–89.

[47] S. Rahaman and M. S. Hossain, "A belief rule based clinical decision support system to assess suspicion of heart failure from signs, symptoms and risk factors," in *2013 IEEE International Conference on Informatics, Electronics & Vision (ICIEV)*, 2013, pp. 1–6.

[48] 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*, 2017. [Online]. Available: http://dx.doi.org/10.1109/TSUSC.2017.2697768

[49] M. S. Hossain, K. Andersson, and S. Naznin, "A belief rule based expert system to diagnose measles under uncertainty," in *Proceedings of the 2015 International Conference on Health Informatics and Medical Systems (HIMS'15)*, 2015, pp. 17–23.

[50] D. Tang, J.-B. Yang, D. Bamford, D.-L. Xu, M. Waugh, J. Bamford, and S. Zhang, "The evidential reasoning approach for risk management in large enterprises," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20, pp. 17–30, 2012.

[51] M. S. Hossain, M. S. U. Chowdury, and S. Sarker, "An intelligent tender evaluation system using evidential reasoning approach," *International Journal of Computer Applications*, vol. 61, no. 15, 2013.

[52] P. Orponen, "Dempster's rule of combination is# p-complete," *Artificial Intelligence*, vol. 44, no. 1-2, pp. 245–253, 1990.

[53] Y.-M. Wang, J.-B. Yang, and D.-L. Xu, "Environmental impact assessment using the evidential reasoning approach," *European Journal of Operational Research*, vol. 174, no. 3, pp. 1885–1913, 2006.

[54] 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.

[55] Z.-J. Zhou, C.-H. Hu, J.-B. Yang, D.-L. Xu, and D.-H. Zhou, "Online updating belief rule based system for pipeline leak detection under expert intervention," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7700–7709, 2009.

[56] Z.-J. Zhou, C.-H. Hu, J.-B. Yang, D.-L. Xu, M.-Y. Chen, and D.-H. Zhou, "A sequential learning algorithm for online constructing belief-rule-based systems," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1790–1799, 2010.

[57] C. Larman and V. R. Basili, "Iterative and incremental developments. a brief history," *Computer*, vol. 36, no. 6, pp. 47–56, 2003.

[58] 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.

[59] Y. Shafranovich, "Common Format and MIME Type for Comma-Separated Values (CSV) Files," Internet Requests for Comments, RFC 4180, October 2005. [Online]. Available: http://www.rfc-editor.org/rfc/rfc4180.txt

[60] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau, "Extensible markup language (xml)." *World Wide Web Journal*, vol. 2, no. 4, pp. 27–66, 1997.

[61] D. Crockford, "The application/json media type for javascript object notation (json)," 2006.

[62] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, "Hypertext transfer protocol–http/1.1," Tech. Rep., 1999.

[63] G. Mein, S. Pal, G. Dhondu, T. K. Anand, A. Stojanovic, M. Al-Ghosein, and P. M. Oeuvray, "Simple object access protocol," Sep. 24 2002, uS Patent 6,457,066.

[64] E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana, "Web services description language (wsdl) 1.1," 2001. [Online]. Available: http://www.w3.org/TR/2001/NOTE-wsdl-20010315

[65] L. Clement, A. Hately, C. Riegen, and T. Rogers, "Universal description discovery & integration (UDDI) 3.0.2 (2004)," *IEEE Network*, vol. 15, no. 6, pp. 30–39, 2001.

[66] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S. Weerawarana, "Unraveling the web services web: an introduction to SOAP, WSDL, and UDDI," *IEEE Internet Computing*, vol. 6, no. 2, pp. 86–93, 2002.

[67] R. Fielding, "Representational state transfer," Ph.D. dissertation, University of California, Irvine, 2000.

[68] D. Box, D. Ehnebuske, G. Kakivaya, A. Layman, N. Mendelsohn, H. F. Nielsen, S. Thatte, and D. Winer, "Simple object access protocol (SOAP) 1.1," 2000. [Online]. Available: https://www.w3.org/TR/2000/NOTE-SOAP-20000508/

[69] M. Zur Muehlen, J. V. Nickerson, and K. D. Swenson, "Developing web services choreography standards the case of REST vs. SOAP," *Decision Support Systems*, vol. 40, no. 1, pp. 9–29, 2005.

[70] 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)*, 2015, pp. 133–137.

[71] B. J. Deen, J. M. Soderberg, V. C. Van, and H. L. Sanders, "Methods and systems for processing HTTP requests," Sep. 30 2003, US Patent 6,629,127.

[72] M. Grinberg, *Flask web development: developing web applications with Python*. O'Reilly Media, Inc., 2014.

[73] A. Yusof, A. Siddique, A. Baqui, A. Eusof, and K. Zaman, "1988 floods in Bangladesh: pattern of illness and causes of death," *Journal of Diarrhoeal Diseases Research*, pp. 310–314, 1991.

[74] A. Fink, *How to Conduct Surveys: A Step-by-Step Guide*. Sage Publications, 2012.

[75] C. E. Metz, "Basic principles of ROC analysis," in *Seminars in nuclear medicine*, vol. 8, no. 4, 1978, pp. 283–298.

[76] M. Gönen, *Analyzing receiver operating characteristic curves with SAS*. SAS Institute, 2007.

[77] R. C. Gagnon and J. J. Peterson, "Estimation of confidence intervals for area under the curve from destructively obtained pharmacokinetic data," *Journal of Pharmacokinetics and Pharmacodynamics*, vol. 26, no. 1, pp. 87–102, 1998.

[78] J. T. Roscoe, *Fundamental research statistics for the behavioral sciences*, 1975.

[79] B. G. Buchanan, D. Barstow, R. Bechtal, J. Bennett, W. Clancey, C. Kulikowski, T. Mitchell, and D. A. Waterman, "Constructing an expert system," *Building Expert Systems*, vol. 50, pp. 127–167, 1983.

[80] A. Zadeh, "Long term perspective of sustainability," https://www.youtube.com/watch?v=NTmihXF4WNA, 2014, [Online; accessed 25-May-2017].

[81] A. Klimova, E. Rondeau, K. Andersson, J. Porras, A. Rybin, and A. Zaslavsky, "An international master's program in green ict as a contribution to sustainable development," *Journal of Cleaner Production*, vol. 135, pp. 223–239, 2016.

[82] J. Porras, A. Seffah, E. Rondeau, K. Andersson, and A. Klimova, "Perccom: A master program in pervasive computing and communications for sustainable development," in *Software Engineering Education and Training (CSEET), 2016 IEEE 29th International Conference on*. IEEE, 2016, pp. 204–212.

Appendix 1. Questionnaires of Interview

Table A1.1. Questionnaires for collecting data for each factors of flood risk assessment.

Factor	Question
Direct Intangible	
Percentage of Loss of Cattle	What is the percentage of cattle washed away and/or died because of the flood?
Social Condition	How was the social condition after the flood?
Direct Tangible	
Area	How much area have been effected due to flood?
Water Level	What was the water level?
Availability of Cattle Food	How much cattle food was available?
Accommodation Problem	How was the accommodation problem during the flood?
Availability of Transport	What there enough transport available at flood time?
Length of Road Effected	How much road was affected during the flood?
Road Damage	How much road was damaged?
Duration of Standing Water	How long was the water standing on road?
Amount of Crop	What is the amount of crop affected?
Fertility	How fertile was the land after the flood?
Availability of Labor	How many labors were available in the area after the flood?
Cost of Raw Materials	What was the cost of raw materials after the flood?
Agricultural Wages	What was the wage of labors in agriculture?
Indirect Intangible	
Financial Condition	How was the financial condition after flood?
Mental Condition	How was the mental condition after flood?
Indirect Tangible	
Availability of Stuffs	Was there availability of stuffs after the flood?
Frequency of Travelers	What was the frequency of travelers after the flood?
Cost of Transport	How costly the transportation was after the flood?
Transportation of Goods	How much goods are transported during flood?
Transportation Delay	How much delay was incurred in transportation due to flood?

Appendix 2. Web Sites, Software Repositories and Files

1. Implementation of BRB expert system - https://github.com/iamrafiul/lib_brb
2. RESTful API for BRB expert system - https://github.com/iamrafiul/lib_brb/tree/master/api
3. Web based BRB expert system - http://130.240.134.31/
4. Flask API - http://flask.pocoo.org/docs/0.12/api/
5. Excel data parser for python - http://www.python-excel.org/
6. Postman for testing API - https://www.getpostman.com/
   Google Chrome plugin available https://chrome.google.com/webstore/detail/postman/fhbjgbiflinjbdggehcddcbncdddomop?hl=en