International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: 6.614
Available Online at www.journalijcar.org
Volume 8; Issue 03 (A); March 2019; Page No.17593-17605
DOI: http://dx.doi.org/10.24327/ijcar.2019.17605-3344

Research Article

EXPERIMENTAL STUDY OF A HYBRID INTELLIGENT SYSTEM FOR FLOOD RISK MANAGEMENT

[1]Akpan, Emem Etok, [2]Inyang, Udoinyang Godwin and [3]Akinyokun Oluwole Charles*

[1]Department of Computer Science, Federal Polytechnic, Nassarawa, Nigeria
[2]Department of Computer Science, University of Uyo, Uyo, Nigeria
[3] Department of Software Engineering, Federal University of Technology, Akure, Nigeria

ARTICLE INFO

Article History:
Received 13th December, 2018
Received in revised form 11th January, 2019
Accepted 8th February, 2019
Published online 28th March, 2019

Key words:

Experimental Study of a Hybrid Intelligent System for Flood Risk Management

ABSTRACT

This work carries out the experimental study of a neuro-fuzzy-genetic hybrid framework and demonstrates its potential, strengths and capabilities in flood risk management. A six layered neuro-fuzzy inference engine was formulated where the first, second and fifth layers consist of adaptive nodes while the third, fourth and sixth layers are fixed nodes. The perception of emergency risk management is very important in modern society; therefore this work demonstrates its practical application, data mining techniques and tools for emergency risk management. The implementation of knowledge extraction by rule discovery, rule evaluation and rule pruning is carried out. The work made used of MatLab programming language and Tanagra data mining software as front engines while Ms-Access database served as back engine. The evaluation of the proposed hybrid intelligent system using flood data obtained from Nigeria Emergency Management Agency (NEMA) is carried out. The research examined the application of hybrid intelligent for descriptive and predicative analytic framework in the domain of flood risks management with good result in flood data clustering, visualization, classification and predication.

INTRODUCTION

An emergency condition creates potential threats to life, property and the entire ecosystem. Hence, emergency condition requires prompt attention and emergency response plan which includes gathering of resources and acting upon the problems as soon as the incident occurs. Disasters produces chaos; those affected and those responding to hazard plunge into an unknown, frightening and often dangerous situations where they must urgently perform some unfamiliar activities[1]. Every aspect of any nation economy is vulnerable to threats posed by natural and human-induced disasters, such as flood, typhoon, hurricanes, heavy rains, volcanic explosion, sea level rise, drought, earthquake, oil spillage, terroism, fire and so on. Annually, these natural disasters have created serious negative impact on economic, social, financial, property, cultural heriatage, critical infrastructure and tragic loss of human lives.

Advances in Information Technology (IT) have enabled robust and dynamic approaches at capturing and analyzing large volumes of data from all sectors of the world economy[2,3]. A key requirement of knowledge driven society is the efficient management of data and their transformation into information and knowledge[4].

Corresponding author:* **Akinyokun Oluwole Charles
Department of Software Engineering, Federal University of Technology, Akure, Nigeria

This forms the foundation for new techniques and tools to support organization in automatic and intelligent sourcing and analysis of large databases. This emerging demand delivers new and promising research area known as knowledge discovery, knowledge mining and knowledge warehousing.

In knowledge discovery, a comprehensible view of the revealed patterns is considered more useful than their predictive competence[5,6]. Though, Neural Networks (NN) can acquire patterns from corrupted or noisy data, there are no general methods to determine the optimal weights of nodes. The rules extracted from a predefined rule-set in the knowledge base are incomprehensible, inconsistent and solely depend on the knowledge acquisition method.

Knowledge workers rely on data warehouses for information that allows them to make sound decisions based on a concrete fact[7]. However, only a part of the required information resides on computers, while the vast majority are scattered throughout the organization as intellectual assets among its knowledge workers[8]. Therefore, a knowledge warehouse does not necessarily provide adequate support for knowledge intensive queries in an organization. What is needed is a knowledge enabled systems that provide the infrastructure needed to capture, enhance, store, organize, create and disseminate not only data and information but also knowledge which may include objects, relationships, mathematical models, meta-models, text streams and so on.

IT offers society with advance modelling tools and techniques for dealing and controlling disasters. The flood risk response system interactions with robustness of a mitigation plan are highly contingent upon an IT tools. Such infrastructure facilitates information gathering, creation, storage, integration and organization during disasters handling[9,10]. Intelligent expert systems, decision support systems, knowledge discovery and data mining are some core tools that can support several fields of environmental risk management. Failure of traditional analysis techniques to reveal hidden patterns relationship from large and diverse datasets is taken care by the use of knowledge discovery techniques and knowledge mining system[11,12].

Artificial Intelligence (AI) tools such as Neural Network (NN), Fuzzy Logic (FL) and Genetic Algorithm (GA) aim at extracting useful information and patterns from large database for prediction and modelling[13]. NN are used for classification and prediction and applied to an increasing number of real life problem of high complexity[14,15]. They offer ideal solutions to a variety of classification tasks such as speech, character and signal recognition as well as risk assessment and treatment. Back propagation is a gradient search techniques widely used in optimization for training Neural Networks (NNs), however, it suffers some shortcomings such as inability to find global solutions and increases cost of computation[16]. GA approaches need a priori information and provides a wide ranging optimization method. It is an explorative method used to find approximate solutions to complicated problems through application of the principles of survival of fittest. The major strength of GA is that, incomplete or corrupted data do not affect the end solution but produces desired results[17].

FL is a superset of the conventional Boolean logic which handles the concept of partial truth[18,19]. FL is a powerful problem solving methodology with wide range applications in industrial control and information processing[20]. FL treats inference mechanisms in consideration of cognitive uncertainty, whereas NN is characterized by learning, adaptation, fault tolerance, parallelism and generalization for data processing. These attributes enable the systems to deal with cognitive uncertainty in a human like manner; one may infuse the theories of FL into neural network. Human operators can enhance NNs by incorporating their knowledge with fuzzy membership functions, which are fine-tuned by a learning process. On the contrary, GA is a robust tool used to implement structure and weights optimization of NNs. Hence, it is imperative to blend NNs, FL and GA techniques in complementary manner[21,22].

Flood is a recurring global phenomenon and seemingly getting worse due to climate change and man-made activities[23,24,25]. All over the world, floods are highly familiar natural disaster; their frequency, magnitude and damaging effect are on the increase. Flood is a natural phenomenon that results in the temporary submerging with water of a land that does not occur under normal conditions. Floods are natural events which have high frequent occurrence and serious consequences on the environment. Factors responsible for inducement of floods are anthropogenic activities and human interference with the natural processes. The human interferences with the natural processes include increase in settlement areas, population growth and economic assets over low lying plains prone to flood leading to alterations in the natural drainage and river basin patterns, deforestation and climate change.

Public and private sectors all over the world have integrated flood risk mitigation plans into their organizational plans. The plans seek to achieve the following benefits such as providing a master plan for flood control and relief measures for victims; mitigate floods through the relevant land use laws and edicts; improve institutional capacity for flood prediction and public awareness programmes and minimize the impact of floods through the provision and maintenance of appropriate infrastructure. Flood risk is generally used to associate potential consequences, positive or negatives, linked to a specific decision, act, fact or hazard. Scientific literature largely agreed that flood is not a risk on its own: the concept of risk involves at least two aspects, a hazardous phenomenon and the vulnerable systems exposed to it[26,27]. The hazardous phenomenon is the presence of water and its characteristics in a specific place and time. The vulnerable systems are the assets, human beings, goods, public or private infrastructure, environment and valuables disclosed to the hazard. Therefore, the flood incidence is a natural process that is taken as a risk only if human added values are possibly affected by flood. Flood risk management creates a serious challenge in terms of scientific knowledge as well as socio-economic efficiency improvement. Flood risk assessment is characterized by lack of objective measures of acceptable risk, scarcity of data and an abundance of unknown probability distributions. Flood risk analyses are essential tools to support territory organization; land uses policies, flood management projects, recovery budget and insurance rates determination. Risk has been classified into three groups namely; risk for which statistics of the identified causalities are known; risk for which there may be some evidences, but where the connection between suspected cause and damage cannot be established and the estimates of the probabilities of the events that have not been occurred[28]. Many approaches of risk analysis can be adopted with different objectives. The evaluation of potential flood damage is a useful type of analysis that is growing in popularity. These analyses allow for quantification of the risk considering different criteria and approaches according to the objectives of the evaluation.

This research examined the application of hybrid intelligent for descriptive and predicative analytic framework in the domain of flood risks management with good result in flood data clustering, visualization, classification and predication. A six layered neuro-fuzzy inference engine was developed. The first, second and fifth layers consist of adaptive nodes while the third, fourth and sixth layers are fixed nodes. The implementation of knowledge extraction by rule discovery, rule evaluation and rule pruning. The work made used of MatLab programming language and Tanagra data mining software as front engines while Ms-Access database served as back engine. The evaluation of the proposed hybrid intelligent system using flood data obtained from Nigeria Emergency Management Agency (NEMA) is carried out. The research examined the application of hybrid intelligent for descriptive and predicative analytic framework in the domain of flood risks management with good result in flood data clustering, visualization, classification and predication.

Experimental Study of the Proposed System

The implementation of the neuro-fuzzy-genetic Hybrid Intelligent System for Flood Risk Management (HISFRM) which relied on the design proposed in[29] is carried out in a computing environment characterized by the following:

Hardware requirements: Pentium IV computer system, 2.0 GHz Duo Core processor and 2Gb of Random Access Memory (RAM). Pentium IV computer system is chosen as the mini Central Processing Unit (CPU) for HISFRM application because it has Quad Data Rate (QDR) feature which enhances local system bus performance four times better than the actual clock rate thereby facilitating quick transfer of data, speed processing of instruction and quick response to users request. Duo core processor is the integration of two central processing units (CPUs) on one chip on the motherboard which enhances speed and performance of the system. It supports multi-tasking and multi-threading which facilitates high performance of a heavy load scenario. A duo core with a clock speed of at least 2,0 GHz is suitable for (HISFRM) application. It allows multi-tasking of HISFRM program with other program and multi-threading of HISFRM program on the cores thereby providing speedy processing of program. Random Access Memory (RAM) device is a volatile memory that allows data items to be accessed (read or written) in almost the same amount of time irrespective of the physical location of data in the memory. RAM size of at least 2.0GB is needed to serve as a temporary store and work space for operating system and HISFRM program.

Software requirements: The implementation of the HISFRM is carried out in software environment characterized by Matrix Laboratory (MatLab) Version 8.1 (R2013a), Microsoft Excel 2007, Microsoft Access 2010 Relational Database Management System, Tanagra version 1.4.50 data mining software and Microsoft Windows 8 pro operating system as the platform.

Features Analysis of National Emergency Management Agency (NEMA)

In 1976, the Federal Government of Nigeria established the National Emergency Relief Agency (NERA) to coordinate its disaster response activities. Although its mandate was widened in 1993 to encompass all aspects of disaster management. The increase in deaths from natural and man-made disasters makes mitigation and prevention of disasters an urgent priority. This led to establishment of the National Emergency Management Agency (NEMA) in 1999. The National Emergency Management Agency was established via Act 12 as amended by Act 50 of 1999, to manage disasters in Nigeria. The same decree also empowers the thirty-six States Government to establish State Emergency Management Agencies (SEMA's) which operate in conjunction with NEMA. From inception, NEMA has been tackling disaster related issues through the establishment of concrete structures and measures. Such measures as the education of the public in order to raise their level of awareness and reduce the effects of disasters in the Country. The Agency has also put in place structures that enable it detects, respond and combat disasters in a timely manner. With continuous overwhelming support from the Federal Government and other stake-holders in Disaster Management, NEMA has continued to improve in its capability and effectiveness to discharge its duties passionately.

For the purpose of operational efficiency, NEMA created six zonal offices and three operations offices in the country to take emergency management to the grassroots where it matters most. The zones are North central, with responsibility for Benue, Nasarawa and Plateau states; North East covering Adamawa, Borno and Yobe; North West for the seven states of Kaduna, Kano, Kebbi, Jigawa, Katsina, Kebbi and Sokoto aand South East for Abia, Anambra, Ebonyi and Enugu states. Others are South South zone which covers Akwa Ibom, Bayelsa, Cross River, Delta, Edo and Rivers states. The Abuja operations office has responsibility for the Federal Capital Territory and Kogi state. Minna operations office is responsible for Kwara and Niger state, while Gombe operations office covers Bauchi, Gombe and Taraba states. In case of any flood disaster occurrence, Zonal offices lead others stakeholder to carry out assessment and present report to head office. Their report gives detailed account of how the flood disaster affected the various states, local governments and communities and recommendations on prevention or mitigation of future flood disaster occurrence.

The primary Responsibilities of NEMA Amongst Others Includes

a. formulate policy on all activities relating to disaster management in Nigeria and coordinate the plans and programmes for efficient and effective response to disasters at national level.
b. monitor the state of preparedness of all organizations or agencies which may contribute to disaster management in Nigeria.
c. collate data from relevant agencies so as to enhance forecasting, planning and field operations of disaster management.
d. educate and inform the public on disaster prevention and control measures.
e. coordinate and facilitate the provision of necessary resources for Search and Rescue and other types of disaster curtailment activities in response to distress calls.
f. coordinate the activities of all voluntary organizations engaged in emergency relief operations in any part of Nigeria.
g. receive financial and technical aid from international organizations and non-governmental agencies for the purpose of disaster management in Nigeria.
h. collect emergency relief supply from local and foreign sources such as international and nongovernmental agencies
i. distribute emergency relief materials to victims of natural or other disasters and assist in the rehabilitation of the victims, where necessary.
j. liaise with State Emergency Management Agencies to assess and monitor, where necessary the distribution of relief materials to disaster victims.

Since inception in 1999, NEMA has been active in providing timely relief assistance to disaster victims both at national and international levels. The Agency has provided relief materials to disaster victims from various states across the country in response to various emergencies within the country. The Agency, has also undertaken the resettlement of the Internally Displaced Persons (IDPs). The Federal Government through NEMA has a mandate to assist States and Local Governments in disaster response and recovery. Recovery personnel at the National Emergency Coordination Centre (NECC) in line with this mandate are required to closely monitor response activities and to obtain valuable data regarding the severity and intensity of the event, the affected geographic area and the potential unsatisfied critical needs of the affected population.

Description of Operations Dataset of NEMA

Secondary data were collected and used from NEMA, Abuja. The secondary data were flood events that took place in Nigeria and recorded in NEMA administrative flood disaster and annual reports. Nigeria is located between latitude 4^oN to 14^oN and longitude 3^oE to 15^oE. It has a land extent of about 923,769 km^2; a north-south length of about 1,450-km and a west-east breadth of about 800 km. It is a country with diverse biophysical characteristics, ethnic nationalities, agro-ecological zones and socio-economy. The Federal Capital Territory (FCT) is Abuja and Lagos is the largest city and main commercial centre. The Country has 36 States, FCT and 774 Local Government Areas. The States had experienced flood disaster occasionally or frequently. A total of 918 flood instances were from December, 2009 to December, 2015. The record contains the following tuple, namely: State, Local Government Area, date of flood occurrence, number of affected communities, number of people affected, source of flood, relief materials distributed to victims, number of Internal Displaced Person (IDP), number of persons injured, number of deaths and government response to serve as input variables. Other input variables such as a projected local government area population density, political zone and average household size were collected from the National Population Commission (NPC), Abuja. These input variables were used to give full description of the flood incidence location. The detail attributes of flood dataset is depicted in Table 1. Based on expert opinion, the sources and mitigation measures of flood were recorded and presented in Table 2 and 3 respectively.

Table 1 Attributes Description of Flood Dataset

S/n	Attribute	Description	Attribute Type
1	State	State	Discrete
2	Aff_LGA	Affected Local Government Area	Discrete
3	State_Zone	State Geopolitical Zone	Numeric
4	Fc_Code	Federal Constituency Code	Numeric
5	Pd_LGA	Population Density of LGA	Numeric
6	Hh_SizeAve	HouseHold Size Average	Numeric
7	Sea_Occ	Season of Occurrence	Numeric
8	So_Flood	Source of Flood	Numeric
9	Mi_measure	Mitigation measure	Numeric
10	Co_Relief	Cost of Relief in Million of Naira	Numeric
11	No_AffComm	Number of Affected Communities	Numeric
12	Day_Occ	Day of flood	Numeric
13	Mo_Occ	Month of flood	Numeric
14	Ye_Occ	Year occurred	Numeric
15	No_ Death	No of Death	Numeric
16	No_InjuPer	No of injured persons	Numeric
17	No_ HDam	No of House damaged	Numeric
18	No _ AffPerson	No of Affected Person	Numeric
19	No_ IDP	No of Internal Displaced Persons	Numeric

Table 2 Source of Flood

Code	Source of flood
1	Blockage of Canals , Filled/Silted/Dirty Drainage Channels and Collapsed Bridges/Culverts.
2	Constructions and Reconstructions, Encroachment/Land , Reclamation and Farming along Flood Plains.
3	Illegal Channelization of Drains, Illegal Structure on Drainage, Channels, Inadequate Drainage Channel and Unregulated development of disaster prone areas.
4	Nature of Terrain, Global Warming and Climate Change, Ocean/Lagoon Surge (Rise in sea level), Overflowing from River, Base Water Flow and Spring Water Flow.
5	Poor heeding to predictions, Poor Physical Planning, Negligence, Neglects of building codes, Substandard housing and physical structures and Inadequate land-use planning.
6	Torrential/Heavy Rainfall Severe Wind Storm with Rainfall.
7	Release of Excess water from Dams.
8	Improper disposal of waste and refuse and Social Cultural Activities.

Table 3 Mitigation Measure for Flood

Code	Mitigation Measures (set)
1	Construction of Embankment, Construction of Drains, Construction of Culverts, Dredging of River and Canals and Proper refuse disposal.
2	Proper land use Planning and Public Enlightenment against Danger of Flood.
3	Provision of Adequate Drainage System, Proper Land use Planning Construction of Drains, Proper Refuse Disposal, Dredging of River and Canals.
4	Construction of Embankment, Public enlightenment against Danger of Flood, Dredging of River and Canals.
5	Provision of Adequate Drainage System, Proper land use Planning, Dredging of River and Canals, and Public Enlightenment against Danger of Flood.
6	Public Enlightenment against Danger of Flood.
7	Public Enlightenment against Danger of Flood and Proper Refuse Disposal.

Implementation Techniques

The neuro-fuzzy-genetic model for flood risk system is implemented with MatLab and Tanagra software packages, which serves as the core programming tool and front end engine. Microsoft Access 2007 Version serves as the database system for flood incidence records, which is the back end engine. Microsoft Excel 2007 Version was used to pre-process the required dataset into a format that could be exported to MatLab and Tanagra workspace. Micro Excel serves as an interface and means of interoperability between Tanagra and M atLab as presented in Figure 1. The configuration allows Matlab and Tanagra software packages to launch automatically, once Microsoft Excel is loaded.

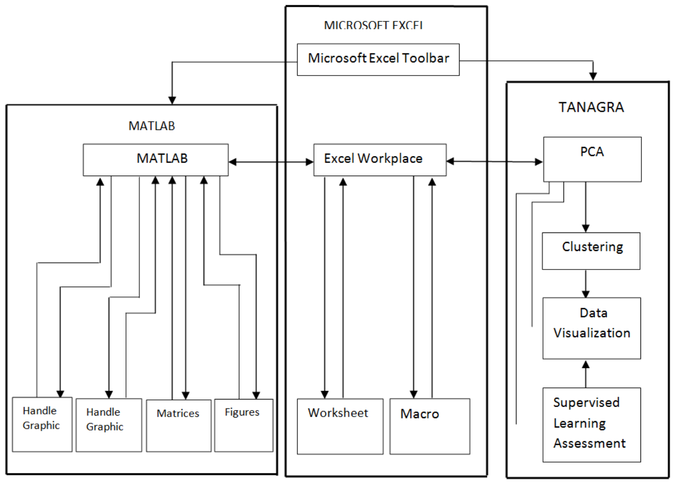

Figure 1 Interaction of Matlab, Tanagra and Microsoft Excel

The Mat Lab Tabs includes the Following

a. startmatlab, responsible for initializing a new Matlab application or connecting to an existing MatLab session.
b. putmatrix, allows the transfer of data from Microsoft Excel to MatLab workspace.
c. getmatrix, permits the retrieval of data from MatLab workspace into Microsoft Excel Workshop Worksheet.
d. evalstring, is a function used in executing MatLab command in Excel.
e. getfigure, imports a current MatLab figure into Microsoft Excel workspace.

f. preference, allows MatLab to be configured in Microsoft.

Tanagra data mining software was used for input variable rank analysis. The aim is to produce a few numbers of factors which summarize as better as possible the amount of information in the data. The Tanagra has the following tabs and components as reported in Table 4

Table 4 Tanagra Tabs and Components

Tab	Operator (Component)	Function
Data visualization	View dataset	View the content of the data file in a grid
Feature selection	Define status	Specify the attributes to use
Descriptive stats	Univariate continuous stat	Descriptive statistics for continuous attributes
Descriptive stats	Univariate discrete stat	Descriptive statistics for discrete attributes
Descriptive stats	Group characterization	Statistics for sub-population

The first step in using the Tanagra software is data importation. This involves importing and viewing the dataset from Microsoft Excel environment into Tanagra depicted in Figure 2. Then, selection of attributes is performed by clicking define status icon and indicating the desired attributes and attribute status whether target, input or illustrative. The outcome of these activities gives dataset description as shown in Figure 3.

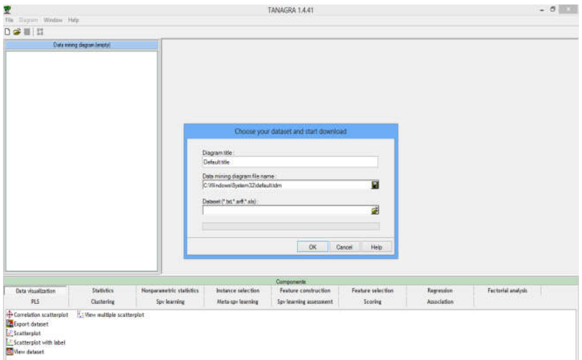

Figure 2 Tanagra Data Import Interface

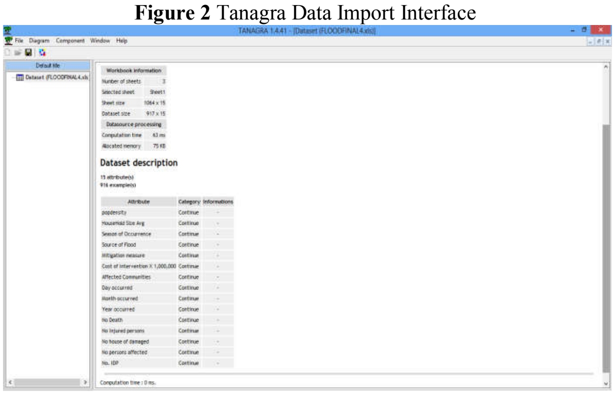

Figure 3 Tanagra Feature selection Windom

Two Features of Mat Lab are Used and They are

a. Self-Organizing Map (SOM) and
b. Adaptive Neural Fuzzy Inference System (ANFIS).

The SOM visualization was used as clustering tool, which involves partitioning flood risk data by similarity thereby determine the level of flood risk by inspecting the relation position between the nodes cluster in the map. The SOM start up window on the MatLab is presented in Figure 4. The use of SOM tool from the MatLab starts by invoking the Neural Network Clustering Tool startup window in the MatLab as follows:

a. Open the Neural Network Start Graphic User Interface (GUI) or with this command: nnstart.
b. Click Clustering Tool to open the Neural Network Clustering Tool, SOM window comes up as depicted in Figure 5.
c. Click select data window.
d. Load data set.
e. Click simple clusters and import.
f. Click next to continue.
g. Click Train network.
h. The training runs for the maximum number of Epochs, which is 200.

MatLab fuzzy logic toolbox is used to design Adaptive Neuro Fuzzy Inference System (ANFIS) for flood risk classification. Using the given training data set, the toolbox constructs an ANFIS structure using a combination of back propagation with least squares type of method (hybrid algorithm). ANFIS model can be generated either from the command line or through the ANFIS editor GUI. In this study, ANFIS Editor GUI is used to generate the ANFIS models with the chosen design parameters in construction phase. Figure.6 shows ANFIS interface ready for accepting parameter settings and loading of data set. Figure 7 displays outcome of the training session with Fuzzy Inference System.

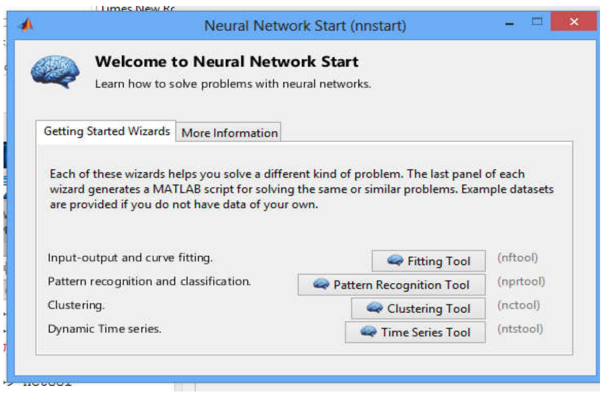

Figure 4 Welcome Inter face for Neural Network Tool in Mat Lab

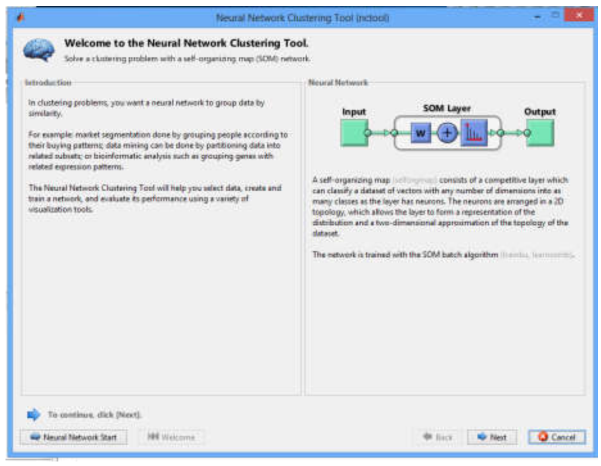

Figure 5 Mat Lab SOM Window

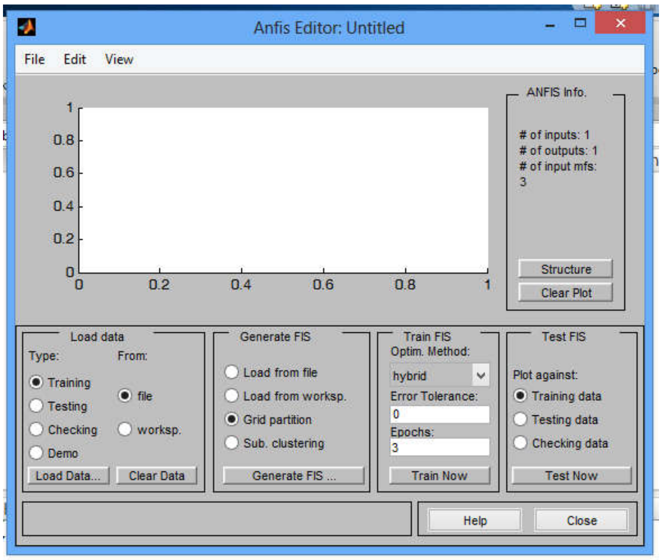

Figure 6 MatLab ANFIS Interface Ready for Data Loading and

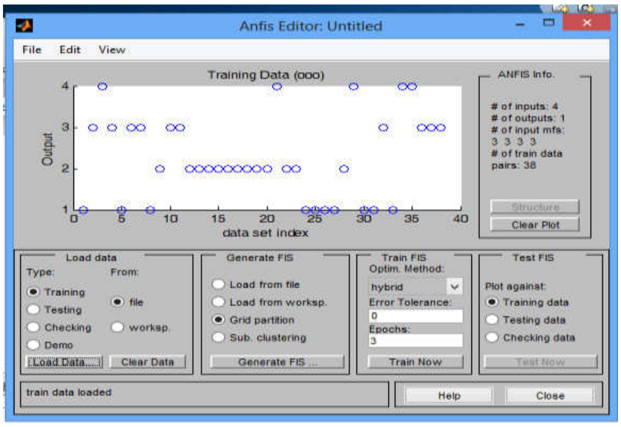

Figure 7 Outcome of the Training Session with Fuzzy Inference System

A customized written MatLab code is used to train the ANFIS structure in the training step. The use of the ANFIS editor GUI can be found in program help files. The steps in ANFIS for classification and prediction in this study utilizing the MatLab fuzzy logic toolbox are as follows:

a. Generated training data set that contained desired input/output data pairs of the target system to be modelled is loaded to the Editor GUI.
b. Design parameters, number of input MF, type of input and output Membership Function (MF), are chosen. Thus, initial ANFIS structure is formed.
c. The code for the training is run with the initial structure.
d. ANFIS structure constituted after training is saved to use as classifier.

Case Study of NEMA

Implementation of the proposed HISFRM was achieved with MatLab Version 8.1 (R2013a), Microsoft Access 2007 Version for Flood incidence records, Microsoft Excel 2007 Version was used to pre-process the required dataset into a format that could be exported to MatLab workspace. The implementation of neuro-fuzzy-genetic model for flood risk management was performed in the following stages:

a. Selection and Dataset Pre-processing.
b. Data Transformation.
c. Unsupervised learning for flood risk pattern assessment and classification.
d. Supervised learning Hybrid intelligent system for flood risk severity level prediction.

Selection and Database Pre-processing

Selection of data entails the identification of the variables on which the intelligent technique will be worked upon. In this work, there was no target variable. The target variable will be determined in subsequent operations. The selected attributes or variables are numeric data since unsupervised learning method is applied to numeric dataset.

Data Transformation

Data transformation, normalization, filtering, and filling of missing values is carried out. Categorization of attributes was also performed where textual attributes were converted to numeric codes. Composite attributes were broken-down into atomic attributes for example; date of occurrence was broken down into day, month and year. All the attributes of flood incidence dataset was considered input variables and none of the attribute was designated target variable. Table 5 contains the selected flood incidence attributes ready for training.

Table 5 Flood Incidence Attributes Ready for Training

s/no	Attribute	Description	Type
1.	Population Density of LGA	Population Density of LGA	Numeric
2.	House Hold Size Average	House Hold Size Average	Numeric
3.	Season of Occurrence	Season of Occurrence	Numeric
4.	Source of Flood	Source of Flood	Numeric
5.	Mitigation measure	Mitigation measure	Numeric
6.	Cost of Relief Intervention (in Million naira)	Cost of Relief Intervention (in Million naira)	Numeric
7.	No of Affected Communities	Number of Affected Communities	Numeric
8.	Day occurred	Day of flood	Numeric
9.	Month occurred	Month of flood	Numeric
10.	Year occurred	Year occurred	Numeric
11.	No of Death	No of Death	Numeric
12.	No of injured persons	Number of injured persons	Numeric
13	No of House damaged	Number of House damaged	Numeric
14	No of Person affected	Number of Person affected	Numeric
15	No of IDP	Number of IDP	Numeric

Unsupervised Learning Procedure for Pattern Assessment and Classification

In this section, a two stage unsupervised learning of flood hazard data with K-means and Self Organizing Map (SOM) is implemented. The implementation aims at determining the number of natural divisions in the dataset and then visualizing the patterns and assigning to each record of the dataset, a target class corresponding to the flood severity level.

Cluster Analysis of Flood Risk Pattern Assessment and Classification.

Some data contain natural divisions that indicate the actual number of clusters; while others do not contain natural divisions or the natural divisions are unknown. Cluster validation is the technique of finding a specific number of clusters that best fits natural number of clusters without any priori class knowledge. Several methods exist in literature such as gap statistic, silhouette, prediction strength, jump method and inference of Bayesian. This work adopts silhouette criterion. The silhouette criterion validates the clustering performance based on the pair wise difference between and within-cluster distances by maximizing the value of this index to determine the optimal cluster number. The Silhouette Value (SV) for a point gives a measure of how similar the point is to other points in its own cluster, when compared to data points in other clusters. The SV for the *ith* point denoted by S_i is given as:

$$S_i = \frac{b_i - a_i}{\max(a_i, b_i)}$$

where a_i is the average distance from the *ith* point to the other points in the same cluster as i, and b_i is the minimum average distance from the *ith* point to points in a different cluster, minimized over clusters. The experiment on optimum cluster number and validation was carried out by implementing K-means algorithm followed by determination of Mean Silhouette Value (MSV). The flood impact dataset used for the training is presented in Figure 8. Distance measure plays a very important role on the performance of K-means algorithm vis-a-vis silhouette criterion. The experiment was therefore conducted for 2 to 20 clusters. The performance of each cluster was assessed with four distance measure (squared Euclidean, cosine, correlation and cityblock). The SV in the various cluster number and distance measures is presented in Table 6, Figure 9 and Figure 10.

As shown in Table 6, Figure 11 and Figure 12, the average SV performance on varying cluster numbers decreases as the number of cluster increases, except in few instances. The top three performing cluster number is 2 clusters, 3 clusters and 4 clusters with averages SV of 0.7565, 0.6210 and 0.7040 respectively. The rank of clusters numbers also reveals the least performing number of clusters (15 and 19) with 0.3974 as average SV. However, in terms of distance metric, Squared Euclidean (Sqeuclid) performed the best with 0.6126 and the average SV followed by cosine with weight of 0.5053. Therefore, Squared Euclidean measure is best suitable for clustering and mining of flood dataset. In this thesis, the best performing number of cluster will be determined from the SV resulting from Squared Euclidean distance metric. Top three performing cluster numbers are 2, 4 and 3 in order of decreasing SV. However, this result is subjected to human experts' judgment which favoured 4 clusters. The silhouette plot for 4 clusters is presented in Figure 11. The figure shows well separated and compact clusters with Mean SV of 0.72. Cluster 1 (High has the highest number of members, while cluster 3 has the least number of data points. The K-means cluster of the flood dataset into 4 clusters is shown in Figure 12. To enable the visualization analysis of the flood attributes and of the cluster classes' assignment, SOM clustering using the results from the K-means algorithm was performed

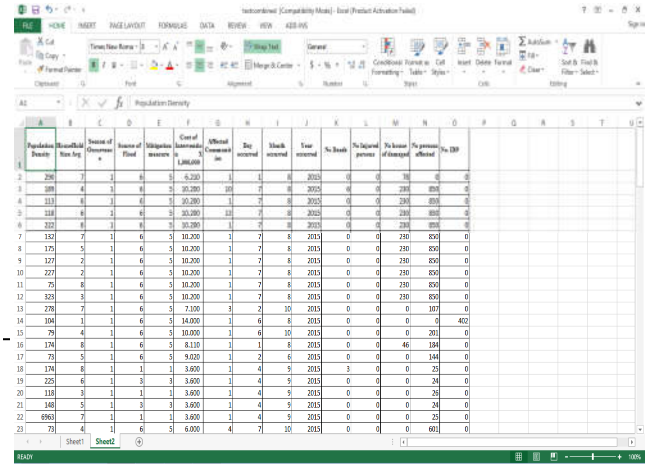

Figure 8 Snippet of Flood Impact Dataset Ready for K-mean and SOM training

Table 6 Performance of K-means on Clusters and Distance Measures

Cluster Number	Correlation	Citybock	squared Euclidean	Cosine	Average SV
2	0.8302	0.568	0.8037	0.8241	0.7565
3	0.4249	0.4647	0.7312	0.8635	0.6210
4	0.8566	0.4209	0.6810	0.8575	0.7040
5	0.4690	0.3948	0.6974	0.6489	0.5525
6	0.4996	0.4275	0.6590	0.5139	0.5250
7	0.2269	0.4084	0.6226	0.5603	0.45455
8	0.5821	0.4003	0.6117	0.3521	0.48655
9	0.3839	0.3902	0.6165	0.5952	0.49645
10	0.4069	0.3722	0.6152	0.4328	0.456775
11	0.3701	0.3565	0.5843	0.4415	0.4381
12	0.3411	0.3978	0.6263	0.3942	0.4398
13	0.3480	0.3604	0.5575	0.4509	0.4292
14	0.4421	0.3288	0.5901	0.3903	0.4378
15	0.3314	0.3369	0.5663	0.3552	0.3974
16	0.3731	0.3455	0.5338	0.3510	0.4008
17	0.3943	0.3145	0.5298	0.4071	0.4114
18	0.4671	0.3309	0.5448	0.3961	0.4347
19	0.4033	0.3083	0.5176	0.3652	0.3986
20	0.3533	0.3018	0.5523	0.4015	0.4022
Average	0.4475	0.3804	0.6126	0.5053	

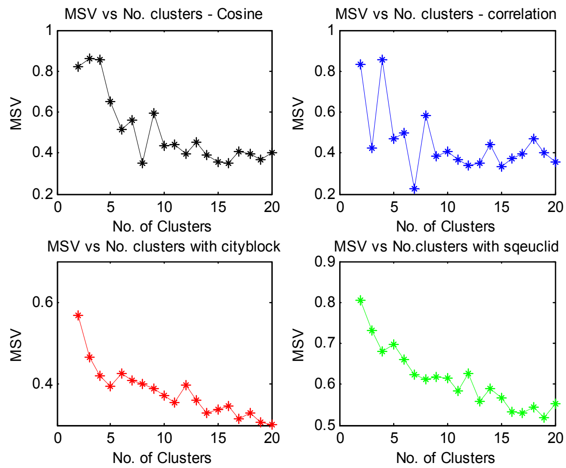

Figure 9 Plot of K-Means Performance on Clusters and Distance Measure

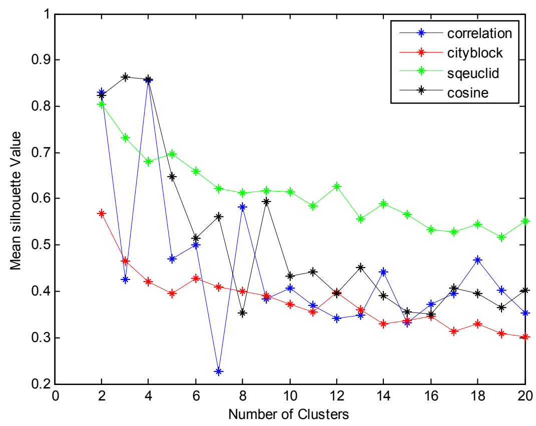

Figure 10 Average Silhousette Value Performance on Varying Cluster Numbers

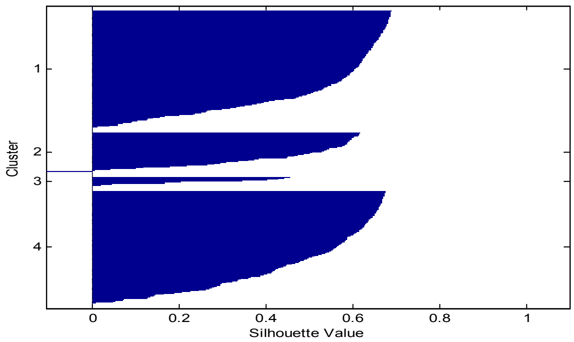

Figure 11 Silhousette Plot of Flood Dataset on 4 Clusters

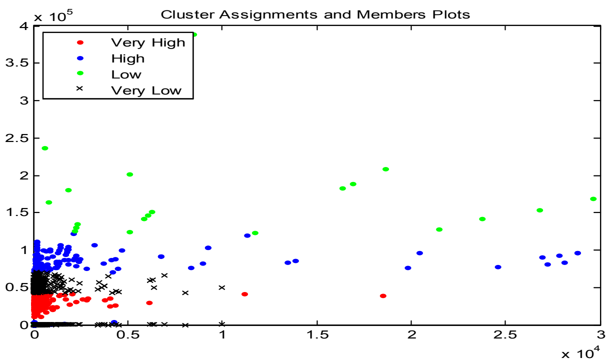

Figure 12 K-means Clustering of Flood Dataset t into 4 Clusters

SOM Visualization Analysis for Flood risk

The batch unsupervised weight/bias algorithm (*trainbu*), in which weights and biases are only updated after all the inputs and targets are presented to the network was adopted. The *trainbu* algorithm trains a network with weight and bias learning rules using batch updates, in two stages namely rough and fine training phases. At the rough training phase, which had 1000 epochs, an initial and final neighbourhood radius was set to 5 and 2 respectively while the learning rate was between 0.1 and 0.5. The fine training phase was performed in 500 epochs with a learning rate standardized at 0.2. The numbers of neurons in the output (Kohonen) layer was set to 4 (as obtained from the k-means analysis), the weight vector ($h_1^1, h_2^1, \ldots, h_n^j$) of the connections constitutes the prototype associated with each neuron and has the same dimension as the input vector. Thereafter, the best performing distance measure and Euclidean distance criterion was utilized to select the best representative (centroids) of the flood dataset feature within each cluster. SOM was applied to identify and classify the flood dataset into the segments with similar flood characteristics.The cluster map topology is as shown in Table 7. The map quality (Total Sum of Squares) of 78.6% is explained by partitioning the dataset into 4 classes. As shown in Table 7, cluster 1 has 373 data points while cluster 2 and cluster 3 has 130 data points and 30 data points respectively, cluster 4 has 385 data points. The tool gives opportunity, after training to evaluate and visual the network performances like a hexagonal grid topology: neighbour distances, input plane, sample hits and SOM weight position.

Table 7 SOM Cluster Map Topology

Cluster Number	Number of data points
1	373
2	130
3	30
4	385

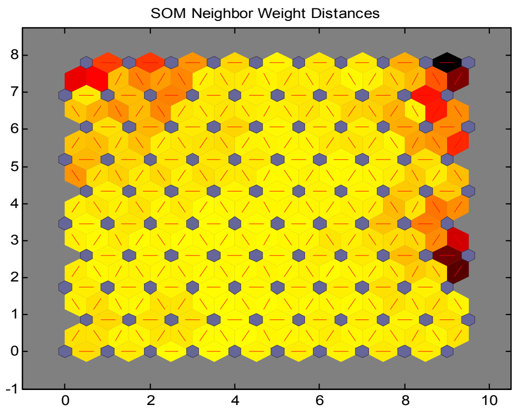

Figure 13 SOM Neighbour Distance Visualization

Neighbour Distance Visualization

Figure 13 displays the SOM weight distance. It is observed that a group of light segment dominates the matrix with band of darker segments separating them. The upper triangular regions contain small group of compacted clustered data point. The corresponding weights indicated by light colour are closer together in this region.

SOM Input Planes Visualization

The visualization of the inputs of flood hazards is presented in Figure 14 which shows that input attributes 4, 5, 8 and 9 exhibit similar configurations and are highly associated, this implies that the source of flood, the mitigation measures and the day occurred and month occurred are correlated. They revealed similar flood severity. This goes to confirm that all the location of the zones experienced flood risk on the same month. Inputs, 6, 12, 13, 14 and 15 shows resemblance and are highly correlated. This confirms that the cost of relief intervention was impacted on the number of person death, injured, affected person, IDP and number of house damaged. However input 1 (population density), produces a unique structure, this implies that it is independent of other variables in the description of flood risks and also applies to input 2 (Household Size Average).

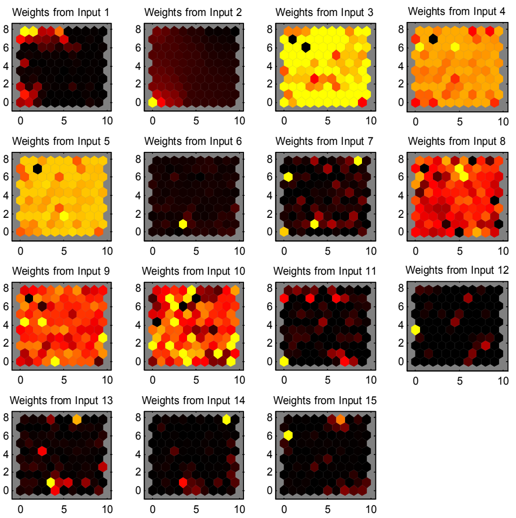

Figure 14 SOM Flood Incidence Input Plane

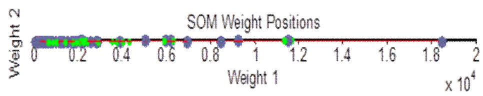

Figure 15 SOM Weight Positions

Weight Position

This SOM weight plots depicted in Figure 15 is a classification structure that displays the input vectors as green dots and shows how the SOM classifies the input space by showing blue-gray dots for each neuron's weight vector and connecting neighbouring neurons with red lines. It is observed that in 200 iterations, the map exhibited uneven distribution throughout the input space. This outlier effect created by the unconnected dots may not be unconnected with the year occurred vector, which shows weak or no correlation with other input vectors. This confirms that year occurred does not contribute to impact of flood risk in Nigeria. The SOM results were severity level of flood magnitude, this was assigned to each records of the flood dataset while a visualization of the input attribute yielded four different groups with similar patterns were discovered. The resultant dataset (flood dataset with target variable) is now set to be trained with intelligent supervised machine learning tools for flood severity level prediction.

Supervised Learning Procedure

Supervised learning is a process of training an artificial neural network by giving a data set consisting of input vectors and desired output associated with each input vector. The resultant dataset (flood dataset with target variable) generated from SOM is now served as input data in this stage. The procedure proceeds as follows:

a. Input ranking analysis.
b. Application of ANFIS model to Flood Risk Management.

Input Variable Rank Analysis

The input variable rank analysis is necessary to rank attributes based on their contributions to the target values and to reduce the size of input variables for better performance in prediction. The reduction in the size of input vector to a lower dimension that includes most of the useful information from the original variables given to the ANFIS model will increase speed and accuracy in the prediction and classification task. The common methods for data dimension reduction are the Linear Discriminate Analysis (LDA) and Principle Component Analysis (PCA). Principal Component Analysis (PCA) is a statistical method of dimension reduction that is used to reduce the complexity of a data set while minimizing information loss. It transforms a data set in which there are a large number of interrelated variables into a new set of uncorrelated variables, the principal components and which are ordered sequentially with the first component explaining as much of the variation as it can. Each principal component is a linear combination of the original variables in which the coefficients indicate the relative importance of the variable in the component. This study adopts PCA technique for the input variable ranking analysis. The PCA technique was implemented on flood data set with Tanagra data mining software and the results obtained are depicted in Table 8, Figure 16a and Figure 16b respectively.

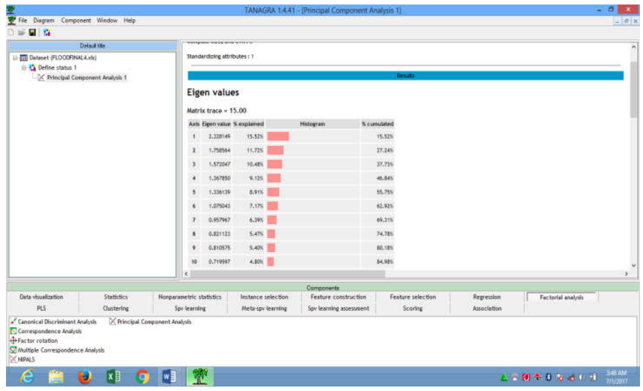

Figure 16a Screen Shot of PCA Result

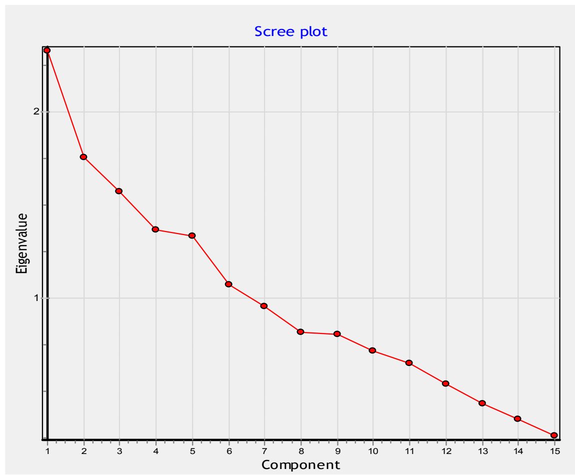

Figure 16b Screen Plot of Eigen Value Versus Principal Component

Table 8 Principal Component Analysis Results

S/n	Attribute	Eigen value	Proposition (%)	Cumulative (%)
1	Population Density of LGA	2.328149	15.52	15.52
2	Household Size Average	1.758564	11.72	27.24
3	Season of Occurrence	1.572047	10.48	37.73
4	Source of Flood	1.367850	9.12	46.84
5	Mitigation measure	1.336139	8.91	55.75
6	Cost of Relief intervention (in Million naira)	1.075043	7.17	62.92
7	No of Affected Communities	0.957967	6.39	69.31
8	Day occurred	0.821123	5.47	74.78
9	Month occurred	0.810575	5.40	80.18
10	Year occurred	0.719597	4.80	84.98
11	No of Death	0.655290	4.37	89.35
12	No of injured persons	0.540962	3.61	92.96
13	No of House damaged	0.439505	2.93	95.89
14	No of Person affected	0.352054	2.35	98.23
15	No of IDP	0.265135	1.77	100.00

15.000000 -

After the execution of the PCA program, the criteria for the selection of input variables for further analysis were based on eigenvalues, the proportion of variance account for and the scree plot. Table 8 shows the importance of each of the fifteen principal components based eigenvalues. The eigenvalue associates to a factor corresponds to its variance. Thus, the eigenvalue indicates the importance of the attributes. The higher is the eigenvalue, the higher is the importance of the factor. To determine the number of relevant factors, the study depends on their eigenvalues. Only the first six components have eigenvalues over 1.00 and together these explain over 62.92% of the total variability in the data. According to the feature reduction results, the first six principal component features having eigenvalue more than 1.0 were selected and then fed into the ANFIS tool for classification. This concludes that a six principal components solution will probably be adequate. On the other hand, a component with eigenvalue less than 1.00 is accounting for less variance than had been contributed by one variable. This conclusion is further supported by the scree plot depicted in Figure 16b. In Figure 16b, there is a "knee" or "elbow" in the scree plot at the eight principle component; therefore, according to a popular rule, the number of principal components to be considered should be 6 which occurred before the knee from top.

Application of ANFIS to Flood Risk Classification

Linguistics variables used and their model function are presented in Table 9. The ANFIS model was implemented in the Graphical User Interface (GUI) of MatLab R2013b. In this stage, ANFIS Editor GUI in MatLab was used to generate the ANFIS models with the default parameters. The ANFIS model for flood risk classification is a 5-layered structure consisting of a total 8224 nodes. The structure of the ANFIS model for flood impact predication and classification is depicted in Figure 17. The rule layer has 4096 nodes; each node represents a rule antecedent part. The normalization layer also has 4096 nodes; each node is the rule consequent part corresponding to the rule antecedent node of the rule layer. The defuzzification and output layer has one node each. The output of the system is the severity of flood disaster risks. Therefore, the proposed ANFIS model contains 8224 rules, the number of neurons in each layer equals to the number of rules. The rule viewer and final surface views of ANFIS are generated and presented in Figure 18 and Figure 19 respectively.

The flood training dataset input for the system were Population density, Household Size Avg, Season of Occurrence, Source of flood, Mitigation measure and Cost of Relief intervention. The output is the flood risk level. There are 6 nodes in the fuzzification layer, which represents linguistic values set Very Low, Low, High, Very High for each input node. During rule extraction phase, grid partition of the input space, hybrid algorithm and triangular membership were applied. Triangular membership is widely most suitably and commonly used in fuzzy inference system. This study employs triangular membership function. ANFIS systems produce different results depending on the type of MF and learning algorithm.

Table 9 Flood Incidence Linguistics Variables and Member Functions

S/n	Variables	Range	Member Functions (MF)	No. of MF	Linguistic variables
1	Population Density of LGA	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 - High 4 – Very High
2	HouseHold Size Average	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 – High 4 – Very High
3	Season of Occurrence	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 – High 4 – Very High
4	Source of Flood	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 - High 4 – Very High
5	Mitigation measure	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 - High 4 – Very High
6	Cost of Relief intervention (in Million naira)	[0-1]	Triangle (trimf)	4	1 – Very Low 2 - Low 3 – High 4 – Very High

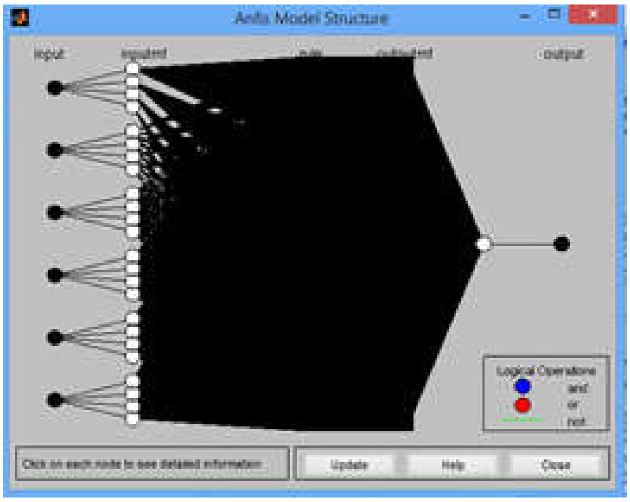

Figure 17 ANFIS Model for Flood Risk Classification

ANFIS Parameters Training and Updating

In this study, hybrid learning algorithm was adopted for determination of optimal values of ANFIS parameters. Since the hybrid learning approach converges much faster by reducing search space dimensions than the original back propagation method, it becomes more desirable and appropriate. Hybrid learning algorithms combines backward propagation and least square techniques. In the forward pass, the antecedent parameters are assumed fixed while the consequent parameters are identified by the Least Square Error (LSE) algorithm. In the backward pass, the consequent parameters are assumed fixed while the antecedent parameters are identified by the back propagation algorithm through gradient descent.

Parameters in the algorithm are epoch size (presentation of the entire data set), error tolerance, initial step size, step size decrease rate and step size increase rate. Since there is no exact method in literature to find the optimum of these parameters a trial and error procedure is used. In all trainings, they are taken as 10, 1x10-5, 0.01, 0.9, and 1.1, respectively as ** default constant value as proposed in MatLab. Training data set that containing desired input/output data to be modelled was brought into MatLab workspace in the prescribed format to initialize the parameterized model. The dataset is an array with the data arranged as column vectors and the output (target) data in the last column.

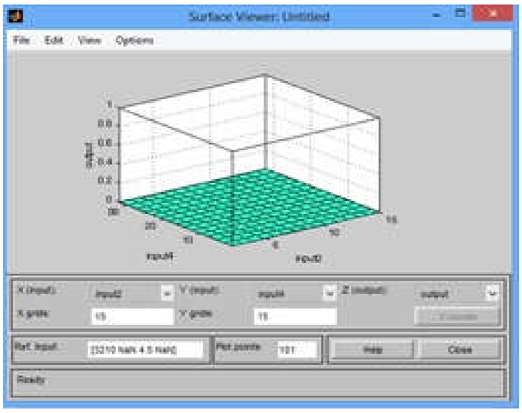

Figure 18 Surface Viewer for Flood Risk

EXPERIMENTAL RESULT AND DISCUSSION

Implementation of the proposed hybrid intelligent system was achieved with MatLab) Version 7.9.0.529 (R2009b) which served as the core programming tool, Microsoft Access 2013 Version which served as the Database for flood incidence records, Microsoft Excel 2013 Version which was to pre-process require data set into a format that could be exported to MatLab workspace. The flood incidence data set was partitioned into two parts:

a. Training (70%) equals to 642 instances
b. Testing (30%) equals to 276 instances

A customized MatLab code is used to train the ANFIS. ANFIS is a sophisticated tool to predict the flood risk level based account a set of data due to its control interpolation and adaptability. Hence, ANFIS suffered computational complexity restrictions. Therefore, there is need to optimize ANFIS parameters. In this study, Genetic Algorithms (GA) and Particle Swan Optimization (PSO) are used as optimization tools. The experimental results of ANFIS trained by GA and PSO for flood risk are presented in Table 10, Figure 19, Figure 20 and Figure 21. The performance measures for the ANFIS optimized by both GA and PSO were Mean Square Error, RMSE, Error Mean and Error Standard Deviation. Figure 7 and 8 gives the various plots obtained from the experiments.

Table 10 Results of ANFIS for flood risk prediction

ANFIS for flood risk prediction				
Optimization Tool	**MSE**	**RMSE**	**Error mean**	**Error Standard Deviation**
GA (Train Data)	0.10403	0.32254	-0.01565	0.4080
GA (Test Data)	0.11956	0.34577	-0.022962	0.43756
PSO (Train Data)	0.082918	0.28795	-0.0042385	0.3645
PSO (Test Data)	0.086681	0.29442	0.038969	0.37104

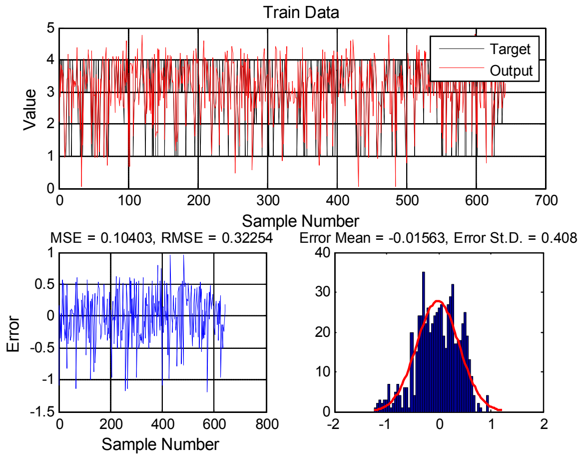

Figure 19 ANFIS Optimized by GA for Train Data

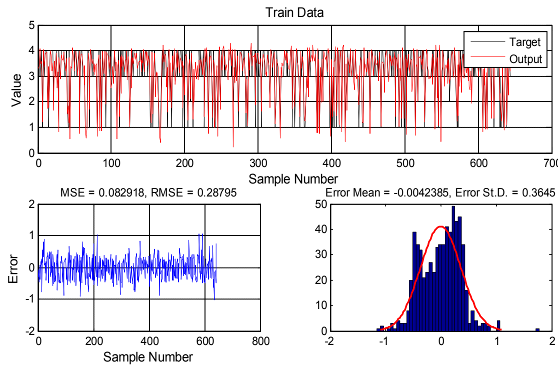

Figure 20 ANFIS Optimized by PSO for Train Data

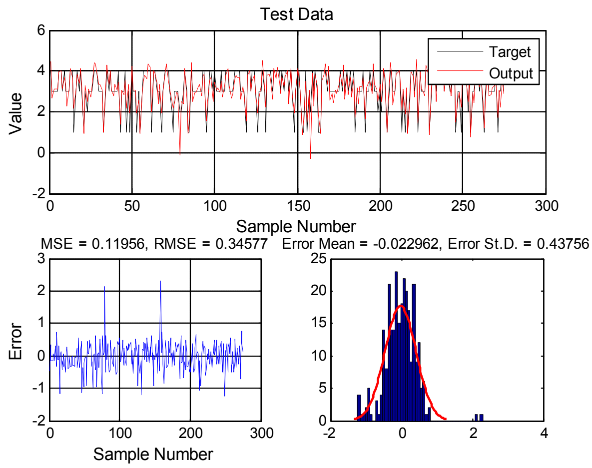

Figure 21 ANFIS Optimized by GA for Test Data

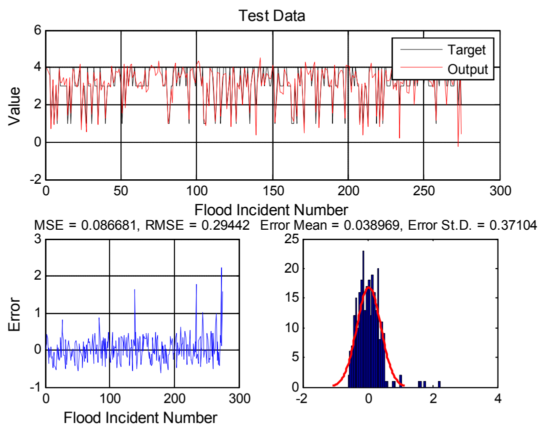

Figure 22 ANFIS optimized by PSO for Test Data

Figure 21 shows ANFIS for flood risk classification trained by GA, the top part diagram is a plot showing almost coincidence of target and output in which the red colour represents the output and the black represents the target. It has RMSE of 0.32254 and Error Standard Deviation of 0.4080. The second right figure is the histogram of the error and shows the mean and the standard deviation of error. Figure 22 shows ANFIS optimized by PSO which gives a better result than ANFIS trained by GA, it has RMSE of 0.28795 which is 11% better than Figure 7. This is significance proof that the performance by PSO is better. In terms of Standard Error Deviation, ANFIS trained by PSO is better than the ANFIS trained by GA. Comparing the test trained results of ANFIS trained by GA and PSO, ANFIS trained by PSO still provides better result with smaller value of RMSE. The mean squared error (MSE) and root mean square error (RMSE) are standard statistical metrics to measure model performance[50]. With low value of RMSE, it confirms that the ANFIS model is most fitted model and will generalize better if presented with unknown flood incidence dataset

Decision Support Engine for Flood Risk Management (DSEFRM)

The output of the ANFIS goes into the DSE which is made up the Cognitive and Emotional filters. The Cognitive filter for example enables the disaster managers to know the level of severity and the impact of flood, while the Emotional filter for example provides the disasters manager with information that enables to decide whether the flood incidence require mitigation measures. In short, the DSEFRM further enhances the performance of the flood risk management and disaster system administration.

System Evaluation

The degree of validity of any system is typically based on its evaluation's outcome. As part of effort to examine the efficiency of the proposed Hybrid Intelligent System, a comparative analysis of the flood incidence data of the flood risk data obtained from the proposed Hybrid Intelligent System (ANFIS) optimized by GA and ANFIS optimized by PSO is obtained. Performance measures are used to indicate how well a method performs its tasks. The performance of the model can be measured in three dimensions: accuracy, complexity and convergence. In accuracy, generalization is an important aspect that measures the ability of the network to correctly classify new unseen data. In supervised learning, the trained data should have the ability to generalize and not memorize the training in order to avoid over fitting. Over fitting can occur when a small error is gained on the training set, but the error becomes large when the network seen new data. There are alternative approaches to optimizing the generalization performance such as selection of the number of hidden nodes, early stopping, cross validation, weight decay, Bayesian technique and others as discussed in the literature. The Root Mean Square Error (RMSE), measures between the desired target output and the actual current output from an estimated model. The RMSE has no upper limit and zero (0) indicates a perfect it. In the ANFIS contexts, if the RMSE measure is not satisfactory, the adjustment of membership functions and the rule refinement procedure is activated towards better optimization of the model. The RMSE can be presented as:

$$RMSE = \sqrt{\frac{\sum(T_p - O_p)^2}{P}} \qquad (2)$$

where P is the number of pattern vectors; T_p is the desired output and O_p is the actual output from each training epoch. From the outcome of the preceding statistical values obtained, it was concluded that the proposed hybrid intelligent system for flood risk management (HISFRS) optimized by PSO provided better classification and prediction result than (HISFRS) optimized by GA.

CONCLUSION

Annualy, around the world, community suffers devasting affect and fiscal loss due to flood risk. The community keeps on experiencing fear of flood risk yearly due to climate change, severe rainfall, rapid population growth, urbanization and poor governance. Management of flood risk is limited due to inadequate information and awareness of hazard. Hence the import of exploring realsitic flood risk mitigation measure become very paramount. This research work proposes a hybrid intelligent system for flood risk management. It aims at developing unstructured approach for handling flood risk. The system adopts a neuro-fuzzy-genetic model to cluster flood incidence data, predict and classify the severity of risk as associated with flood risk in Nigeria. The system capitalizes merits of each tool thereby offsetting each weakness. The work exploits the potential of unsupervised and supervised learning methodologies to analyze and extract useful information from flood data set. Hybridized unsupervised tools (K-means and SOM) were utilized to discover the natural divisions in the flood dataset, visualize and assign class (target variable to the unlabeled flood dataset). The top three performing number of clusters was 2, 3 and 4. This result was used in conjunction with human experts judgment to rely on 4 clusters. In terms of distance measure, Squared Euclidean measure performed more than the other measures such as cityblock, cosine and correlation. The SOM results were severity level of flood magnitude, this was assigned to each records of the flood dataset while a visualization of the input attribute yielded four different groups with similar patterns were discovered. The supervised learning procedure using intelligent supervised machine learning tools (ANFIS) to predict based on resultant dataset with (flood dataset with target variable) for flood severity level was performed. ANFIS for flood risk classification trained by GA produced RMSE of 0.32254 and Error Standard Deviation of 0.4080.

The research examined the application of hybrid intelligent for descriptive and predicative analytic framework in the domain of flood risks management with good result in flood data clustering, visualization, classification and predication. The following recommendations are made for further research:

The result of this research is of significance to environmentalist and emergency managers in making good decision. The predictive analytic and machine learning algorithms developed in this study revealed improved accuracy in identifying low, moderate and high risk flood prone areas. NN design lacks generic methods to determine the optimal number and weights of hidden layers and nodes, which invariably increases the cost of computation. In this work, GA optimized NN configuration thereby producing superior result, however, the outcome can be further enhanced by optimizing NN model with another evolutionary optimization intelligent tools such as Particle Swan Optimization (PSO), Random Forest (RF) and Ant Colony Optimization (ACO).

b. This work could be deployed to computer network environment. There is a growing use of Internet of Thing (IoT) devices in both industrial and home. IoT allows use of embedded intelligence, wired and wireless communication to achieve effective connectivity. If IoT is integrated it would provide seamlessly and effective real time data analysis and decision making.

References

1. Eamon, O. (2008). The Role of Information Technology in Hospital Emergency Management. Emergency Readiness Solution. White paper. www.medworxx.com
2. Kirkos, E, Spathis, C. and Manolopoulos, Y. (2007). Data mining Techniques for the Detection of Fraudulent Financial Statements. Expert Systems with Application (32) pp995-1003. Elsevier www.elsevier.com/locate/eswa.
3. Yao, J. E.; Chang L., Qiyang, C. and June, L. (2008). Administering and Managing a Data Warehouse. Encyclopedia of data warehousing and mining. Information Science Reference. IGI Global Publishing, pp17-22
4. Kao, W, Hung, J. C. and Hsu, V. (2008): Using Data Mining in MURA Graphic Problems. Journal of Software. 3(8), pp73-99.
5. Jagielska, I. (1998): Linguistic Rule Extraction from Neural Networks for Descriptive Data Mining. The Proceedings of Second Conference on Knowledge Based Intelligent Electronic Systems, KES'98, (2), pp88-92
6. Steiger, D. M. (1998): Enhancing User Understanding in a Decision Support System: A Theoretical Basis and Framework, Journal Of Management Information Systems. Pp199-221
7. Devlin, B (1997): Data Warehouse: From Architecture to Implementation, Addison Wesley Longman, Inc: Menlo Park, CA.
8. Nonaka, I. and Takeuchi, H. (1995): The Knowledge-Creating Company, Oxford University Press, NY.
9. Foss, B. A. (1997). Knowledge Discovery – A Control Theory Perspective. In Kromorowski and Zykow J eds. Principles of Data Mining and Knowledge Discovery. Lecture Notes in Artificial Intelligence. Springer.
10. Turrof, M., Chumber, M., Walle, B. V. D. and Yao, X. (2004). The Design of a Dynamic Emergency Response Management Information Systems (DERMIS). The Journal of Information Technology Theory and Applications (JITTA). 5(4): pp1-35.
11. Han, J., Cai, Y. and Cercone, N. (1992). Knowledge Discovery in Databases: An Attribute-Oriented Approach. In Proceedings of the 18th Conference of Very Large Databases, Vancouver, British Columbia, Canada.
12. Miller, H.J. and Han, J. (2001). Geographic Data Mining and Knowledge Discovery-An Overview. Geographic Data Mining and Knowledge Discovery. In: Miller, H.J., Han, J.(eds),. Taylor and Francis, pp 3-32.
13. Agrawal, R. and Srikant, R. (1994). Fast Algorithms for Mining Association Rules. In Proceedings of the 20th International Conference on Very Large Data Bases (VLDB 1994), Santiago de Chile, Chile. Pp487 - 499
14. Obot, O. U. (2007): NeuroFuzzy Expert System Model for the Diagnosis and Therapy of Cardiovascular Diseases; A Case Study of Heart Failure, Ph.D. Thesis, Federal University of Technology, Akure, Nigeria.
15. Akinyokun, O. C., Obot, O. U. and Uzoka, F. M. E. (2009): Application of Neuro-Fuzzy Technology in Medical Diagnosis: Case study of Heart Failure, In proceeding of World Congress on Medical Physics and Biomedical Engineering, Munich, Germany, pp301 – 304.
16. Whitley, D., Starkweather, T., and Bogart, C. (1990). Genetic Algorithms and Neural Networks: Optimizing Connections and Connectivity. Parallel Computing, 14(1). Pp.347-361. http://dx.doi.org/10.1016/0167-8191(90)90086-O
17. Bodenhofer, U. and Herrera, F. Ten Lectures on Genetic Fuzzy Systems Slovak Technical. University, Bratislava. 1997, Pages 1-69.
18. Zadeh, L.A. Fuzzy Logic, IEEE Computer. 1988: 83-89.
19. Akinyokun, O. C. (2002). Neuro-Fuzzy Expert Systems for Evaluation of Human Resources Performance. First Bank of Nigeria PLC Endowment Fund Lecture Series. Delivered at Federal University of Technology, Akure,
20. Akinyokun, O. C., Obot, O. U. and Udoh, S. S. (2008): Application of Softcomputing Methodology to the Management of Hypertension. International Journal Computer Science and Its Applications, 15(2), pp137 – 14.
21. Jain, L. and Martin, N. M. (1998). Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications. CRC Press, CRC Press LLC
22. Akinyokun O.C and Inyang G.U, 2013: Experimental Study of Neuro-Fuzzy-Genetic Framework for Oil Spillage Risk Management. *International Journal of Artificial Intelligence Research. Vol. 2 No. 4, pages 13-36; Published by Artificial Intelligence Research, Canada*
23. EMC (2005). An Emergency Management Framework for Canada Emergency Management Policy Directorate. Public Safety and Emergency Preparedness Canada, Ottawa. www.PublicSafafty.gc.ca. Retrieved on September, 23, 2013.
24. WMO/NOAA (1988). "CLICOM System Reference Manual: Data Management Module, Version 2.1". World Climate Data Program, NOAA, USA.
25. Bhagat, M. (2013). Parametric Evaluation for Flood Management. International Journal of Science & Engineering Research, Vol. 4, pp1147 – 1153
26. Merz, B., Hall, J., Disse, M., and Schumann, A. (2010). Fluvial flood risk management in a changing world, Nat. Hazards Earth Syst. Sci., 10, pp509–527.
27. Messner, F., Penning-Rowsell, E. C., Green, C. and Veen, A. (2007). Evaluating flood damage: guidance and recommendations on principles and methods. Flood site Project Report. Deliverable D9.1 Contract No: GOCE-CT-2004-505420, retrieved on May 20, 2014 from http://www.fllodsite.net/html/partner_area/project.docs/
28. Singh, R., Jain, S. K.. and Tyagi, A. (2007). Risk and Relaibility Analysis: a handbook for civil and environmental engineers. ASCE Press, United State of America.
29. Akinyokun, O. C, Akpan, E. E and Inyang, G.U, (2018). Design of A Hybrid Intelligent System for Flood Risk Management. Submitted to *International Journal of Artificial Intelligence Research. Published by Artificial Intelligence Research, www.sciedu.ca/air, Canada*