International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 4 (2015) pp. 10665-10670

http://www.ripublication.com

# Prediction of Weather Condition Using Probability Function

**S.Dhamodaran**[1]
*Assistant Professor ,*
*Faculty Of Computing,*
*Department of Computer Science and Engineering,*
*Sathyabama University.India.*
*s.dhamodaran07@gmail.com*
**C.AllwynRaja**[2]
*M.E Student,Faculty Of Computing,*
*Department of Computer Science and Engineering,*
*Sathyabama University.India.*
*allwynraja555@gmail.com.*

## Abstract

Flood prediction for planning purposes differs from flood warnings in respect of the need to estimate the long-term probability of flood impacts so that appropriate planning decisions can be made to minimize the impact.Which ensemble the based techniques for getting extremely popular in flood modeling. In this paper, we are utilizing the Probability function and weights-of-evidence (WoE) model, to assess the impact of classes of each conditioning factor on flooding through bivariate statistical analysis (BSA). Then, these factors were reclassified using the acquired weights and entered into the support vector machine (SVM) model to evaluate the correlation between flood occurrence Through this integration, the weak point of WoE which can be solved by the performance of the SVM The spatial database included flood inventory, slope, topographic wetness index (TWI), altitude, curvature, distance from the rainfall, and soil type. Is used to predict by SVM those radial basis function kernel (RBF), and were used to investigate the performance of each kernel type. Which measures the prediction and success rates.. The best results were obtained by mapping when compared with the other kernel types. The proposed ensemble flood susceptibility mapping method could assist researchers and local governments in flood mitigation strategies

**Keywords:** weights-of-evidence (WoE), bivariate statistical analysis (BSA), support vector machine (SVM), topographic wetness index (TWI), radial basis function kernel (RBF)

## Introduction

Natural disasters are the main cause of irrecoverable damages worldwide. Chennai (INDIA) had been affected by the T-SUNAMI in 2004. These flooding has caused considerable damage to highways, settlement, agriculture and livelihood. These consequences can be decreased through an appropriate and accurate susceptibility analysis. Therefore, generating accurate flood modeling is one of the prime goals of scientists and governments. Malaysia embraces flood events annually, mostly during the monsoon seasons 2012. Over the years, many hydrological approaches have been used in the literature. Traditional hydrological methods such as physical based rainfall–runoff modeling techniques and data-driven techniques are not capable of comprehensive analysis of rivers and inundation areas. The reason being the hydrological methods follows one-dimensional procedure while the morphology of the river is not stable and it has a dynamic characteristic due to the high erosive potential. Moreover, these methods require fieldwork and huge budget for data collection. Due to the aforementioned drawbacks, scientists have started to use the empirical and data driven methods such as Artificial Neural Networks (ANNs), Genetic Programming (GP), Evolutionary Polynomial Regression (EPR), M5 Model Trees (M5), and K-Nearest Neighbors (K-NN) techniques. Which makes a significant contribution in flood modeling and prediction. Variety of data sources, high data quality, day and night data collection capabilities, and rapid analysis were offered by RS and GIS techniques for hydrological studies.

**Weight of Evidence**

Measure of evidence on one side as compared with the evidence on the other side of the issue

**Support Vector Machine:**

It is a learning Model with, associated with a recognized pattern that analysis data with classification and regression analysis for a set of trained data and examples Note that if becomes small as grows further away from each term in the sum measures the degree of closeness of the test point to the corresponding database point.

$$\sum \quad i \propto i\, K(xi\, x) = constant$$

**Bivariate Statical Analysis:**

It is one of the simplest form of quantitative analysis .It involves analysis of two variables denoted as (x, y) it is commonly used to measure two variables and also used in creating percentage tables

**Topographic Wetness Index**

Which combines local upslope computing area and slope in commonly used quantity topographic control or topographic process methods of computing index differ primarily in the way where the upslope computing area is calculated

$$TWI = \ln(A_s/\tan\beta)$$

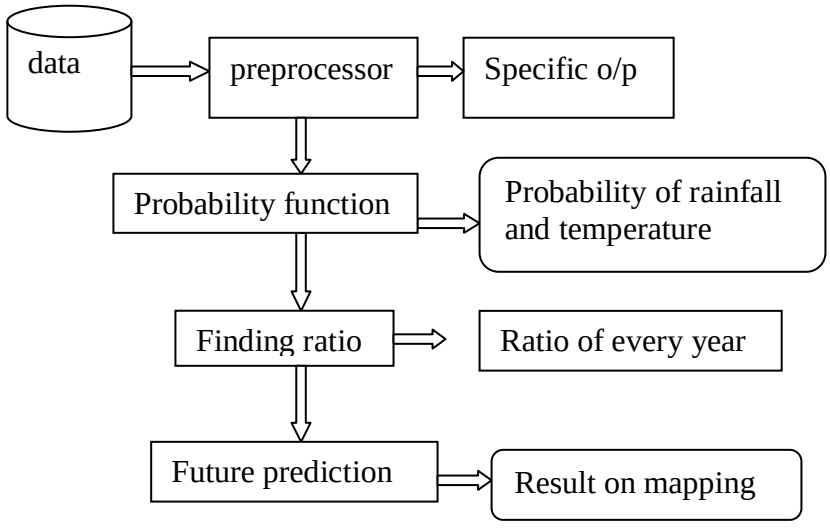
## Review of Existing Work

The most popular methods in natural hazard modeling are ANN analytic hierarchy process (AHP) frequency ratio (FR) logistic regressing (LR). The drawback of AHP is related to its dependency on the expert's knowledge which is the main source of uncertainty. Statistical methods of LR and FR utilize simple and understandable concepts which made them popular some drawbacks can be seen in their performance, although both have valuable advantages. LR is capable to implement multivariate statistical analysis (MSA) but it has some weak points to analyze the classes of each flood conditioning factor. The application of machine learning methods in flood studies is which is more robust that other statistical and deterministic methods were proved by many researches. ANN which is one of the machine learning methods that has been used in many hydrological and water resource engineering applications due to its computational efficiency. The length of the dataset can cause errors in the process of ANN modeling. Furthermore, poor predictions can be achieved when the validation dataset contains values outside of the range of those used for training. Although the ANN has some drawbacks, many researchers utilized this method in their hydrological analysis

## Proposed System

In our proposed system, we are presenting the Probability function and weights-of-evidence (WoE) model, to assess the impact of classes of each conditioning factor on flooding through bivariate statistical analysis (BSA). The current research demonstrates the application of the ensemble SVM and WoE method for prediction of rainfall and temperature on year wise and finding the similarity and the ratio on year wise for future prediction on one of the popular machine learning algorithms on SVM which is a supervised in learning binary classifier and which is based on the structural risk minimization principle Separating hyper-plane formation from training dataset is the basis of this method

**Architecture Diagram:**



**Modules**

The main objective of this project is:

- To predicate the temperature of any specific area for upcoming years.
- To predicate the rainfall of any specific area for upcoming years.
- Provide information about the weather condition with the help of a map.
- Make alert the locality for the any worst condition

**Prediction of Rainfall And Temperature**

The amount of rain forecast for each day is presented on a national color coded map, with options to zoom into states and then district level. Rainfall totals less than 1mm are not shown. The colors represent ranges of forecast rainfall, for example the lightest orange represents a forecast range from 1mm to 5mm

## Ratio Year Wise on Mapping

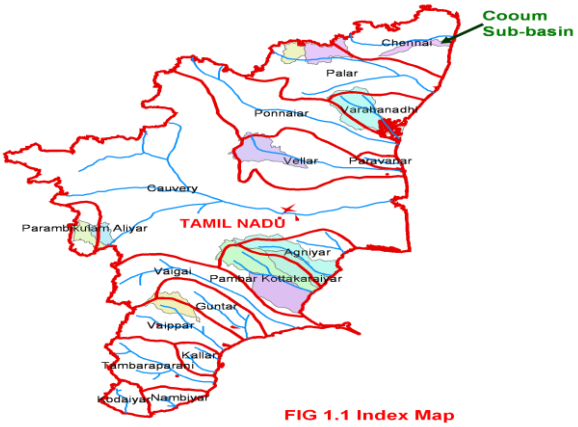



FIG 1.1 Index Map

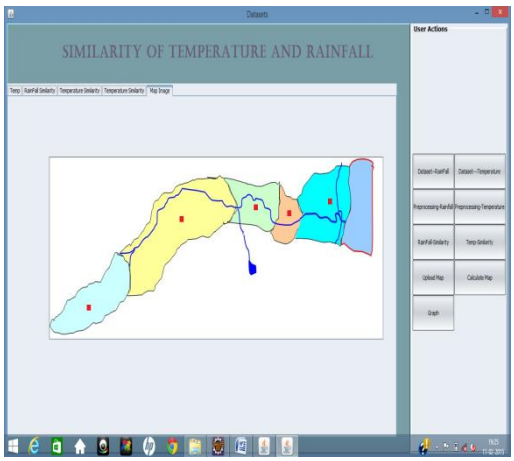
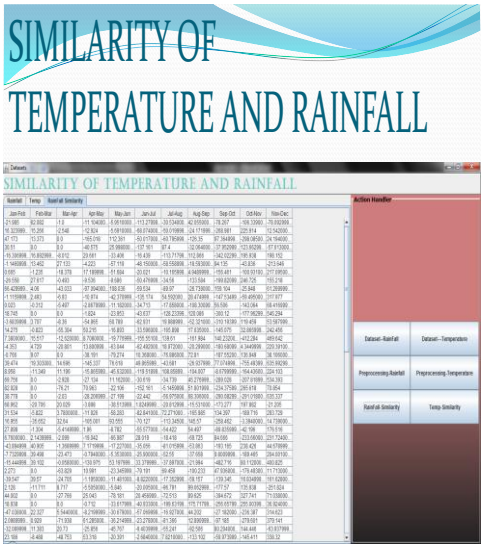



**Figure 1.2:** chennai area

**Rainfall And Temperature In Probabbilty**




**Weather Condition**

Weather is the state of the atmosphere, to the degree that it is hot or cold, wet or dry, calm or stormy, clear or cloudy Weather generally refers to day-to-day temperature and precipition activity, whereas climate is the term for the statistics of atmospheric conditions over longer periods of time For validation both success and prediction rate curves were performed. The validation results showed that, area under the curve for the results of DT and integrated method of FR and LR was 87% and 90% for success rate and 82% and 83% for prediction rate respectively.

## Result and Future Work

Hence by presenting the Probability function and weights-of-evidence (WoE) model, to assess the impact of classes of each conditioning factor on flooding through bivariate statistical analysis (BSA) ,the method for flood susceptibility mapping using GIS, the support vector machine(SVM) which is a supervised learning binary classifier and is based on the structural risk minimization principle Separating hyper-plane formation from training dataset is the basis of this method. Able to analyze the condition of Rainfall and temperature in specific region Satisfactory Prediction. Can assist governments, planners, and researchers to perform proper actions in order to prevent and mitigate this disaster in the future

## Conclusion

Effective modeling has to include topographic characteristics of the area. Many topographic characteristics are described by the digital terrain model. It helps to predict the response of the area The model makes satisfactory decisions for short forecasting horizons in all the above mentioned scenarios, although the variance of

the predictions is different in alternate scenarios. The area should be detected to predict its spatial distribution for future events

## References


[1]. Abarbanel, H.D.I., 1996. Analysis of Observed Chaotic Data. Institute for Nonlinear Science.

[2]. Ayewah, N., 2003. Prediction of spatial, temporal events using a Hidden Markov model. Department of Computer Science and Engineering, Southern Methodist University, Dallas, TX

[3]. Ali, S., Smith, K.A., 2003. Automatic parameter selection for polynomial IEEE International Conference on Information Reuse and Integration, Australia pp243-249

[4]. Tehrany, M.S., Pradhan, B., Jebu, M.N., 2013a. A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping SPOT 5 imagery. Geocarto In

[5]. Ali, S., Smith, K.A., 2003. Automatic parameter selection for polynomial kernel. In IEEE International Conference on Information Reuse and Integration, Australia, Pp. 243–249.

[6]. Povinelli, R.J., Feng, X., 2003. A new temporal pattern identification Method for characterization and prediction of complex time series events. IEEE Transactions on Knowledge and Data Engineering 15 (2), 339–352.

[7]. Xiong, L., Shamseldin, A.Y., O'Connor, K.M., 2001. A non-linear Combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system. Journal of Hydrology 245, 196–217.

[8]. Ohlmacher, G.C., Davis, J.C., 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. Eng. Geol. 69. 331–343.

[9]. S.M. Monzurur Rahman and Xinghuo Yu (2011) , An unsupervised neural network approach to predictive data mining, Int. J. Data Mining, Modelling and Management, Vol. 3, No. 1,pp 18-41.

[10]. Yaquan Xu and Haibo Wang (2011), A new feature selection method based on support vector machines for text categorization, Int. J. Data Analysis Techniques and Strategies, Vol. 3, No. 1,pp 1-20

[11]. Kerry Poitier and Sohyung Cho(2011), Estimation of true efficient frontier of organizational performance using data envelopment analysis and support vector machine learning, International Journal of Information and Decision Sciences,Vol. 3, No.2 pp 148 – 172.

**[12].** Chih-Chung Chang and Chih-Jen Lin (2011), LIBSVM: A Library for Support Vector Machines,ACM Transaction on Intelligent System and Technology, Vol.2,No.3,Article 27.