International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-9, July 2020

# Flood Prediction and Warning System using Dam Data Monitoring

**Meenu Thomas, Shuhaib Ummer, Sruthi Vijayan T K, Vivek T, Ashok S Kumar**

***Abstract:** Flood is one of the most devastating natural calamities affecting parts of the state from past few years. The recurring calamity necessitates an efficient early warning system since anticipation and preparedness play a key role in mitigating the impact. Though heavy and erratic rainfall has been marked as one of the main reasons for flood in several places, flood witnessed by various regions of Kerala was the result of sudden opening of reservoirs indicating poor dam management. The unforeseen flow of water often provided less time for evacuation. Prediction thus plays key role in avoiding loss of life and property, followed by such calamities. The vast benefits and potentials offered by Machine Learning makes it the most promising approach. The developed system is a model by taking Malampuzha Dam as reference. Support Vector Machine (SVM) is used as machine learning method for prediction and is programmed in python. The idea has been to create early flood prediction and warning system by monitoring different weather parameters and dam-related data. The feature vectors include current live storage, current reservoir level, rainfall and relative humidity from the period 2016-2019. Based on the analysis of these parameters, the open/closure of shutters of the dam is predicted. Release of shutters has varied impacts in the nearby regions and is measured by succeeding prediction, by mapping regions on grounds of level warning to be issued. Warning is issued through Flask-based server, by identifying vulnerable areas based on flood hazard reference for regions. The dam status prediction model delivered highest prediction accuracy of 99.14% and associated levels of warning has been generated in the development server, thus preventing unexpected release.*

***Keywords:** Flask, GridsearchCV, inundation, SVM*

## I. INTRODUCTION

Recently, Kerala witnessed excessive rains and the state went through one of the largest, most catastrophic floods, causing irreparable damage to people's livelihood, property and infrastructure. According to preliminary estimates, in 2018, about 480 people were dead, 140 went missing and economic damage exceeding Rs.400 billion was incurred.

**Revised Manuscript Received on July 30, 2020.**
* Correspondence Author
**Meenu Thomas***, Electronics and Communication, NSS College of Engineering, Palakkad, India. E-mail: 98meenuthomas@gmail.com
**Shuhaib Ummer**, Electronics and Communication, NSS College of Engineering, Palakkad, India. E-mail: shuhaib360@gmail.com
**Sruthi Vijayan T K**, Electronics and Communication, NSS College of Engineering, Palakkad, India. E-mail: sruthivijayan998@gmail.com
**Vivek T**, Electronics and Communication, NSS College of Engineering, Palakkad, India E-mail: vivekthottungal1998@gmail.com
**Ashok S Kumar**, Electronics and Communication, NSS College of Engineering, Palakkad, India. E-mail: ashoksk1812@gmail.com



Again, 2019 proved disastrous for the state at the cost of over 120 lives. The disaster aggravated by sudden opening of reservoirs across the state, accompanied with heavy rainfall. The outflow rate from the dams could not be predicted which resulted in huge loss of infrastructure. In the regions of Palakkad, the flood was due to the sudden opening of shutters of Malampuzha dam. The dam has a spillway of 4 shutters of 10.97m x 4.57m each. When water level rises above crest level, shutters must be opened. When this release happens, while downstream areas are subject to weather vulnerabilities, the dams end up increasing the proportions of downstream flood disaster. Flood occurring from water released by dams comes much more suddenly, leaving little time to retort. This points to the need of an efficient flood prediction system. Flood prediction models are of significant importance for such hazard evaluation and extreme event management. The focus of our work is on Support Vector Machine (SVM) model-based prediction. The model takes into consideration weather parameters such as rainfall and relative humidity and dam-related data for predicting the opening and closure of shutters and a webpage is created, into which the registered users can enter the input datas to view the predicted dam status and subsequent warning. The rest of paper is as below. Section II deals with some of the research papers which have been studied for the development of the proposed system. Section III is regarding system design which explains the prediction model and about its deployment. Section IV is about the result and accuracy of the model. The paper is concluded in Section V.

## II. RELATED WORKS

The developed model in [1] is implemented using SVM. The input parameters included 3 gauging stations and the dataset had a span of 2 years. The collection is done from prediction station and upstream station. The system provided prediction of flood 1 hour and 7 hours in advance. The model could be used in real time prediction systems and alerts can be issued as electronic billboards. A comparison between SVM and MLP is also done from which SVM was found to be more accurate.

The past rate of flow and rainfall are taken as inputs in system [2]. The lead time may also influence the data input combinations. A lead time is set which has its own time series. If short lead time is considered the latest flow is considered in the prediction whereas when the lead time is longer, rainfall data would play prominent role. A smart weather station that monitors and predicts weather data and generate instant alerts for dwellers of different areas, to warn them about the future hazard, making use of Internet of things and Machine Learning is described in [3].

It is deployed with different sensors that collect weather data from the environment, which are sent to cloud, where predictions are made, for which certain neural network models have been compared to find out which gives the most accurate results.
Prediction efficiency in terms of minimized errors are computed using 4 nonlinear auto regression models of which NARXNET (Nonlinear Auto-Regressive Exogenous Neural Network) provided maximum efficiency. 24/7 support by alert tweets generated are accessible to people.

The developed system in [4] has daily rainfall and dam's level as crucial variables. The system has Klang Gate Dam as its reference model. It involves a time series forecasting model and an SVM model that are based on attribution of missing values and choosing variables. Model is initially applied with various attribution methods and results are compared with delete approach to forecast missing values. After determining the main variables which affects water level, environment conditions are also taken into consideration. The immaterial factors are then successively isolated.

The system in [5] monitors water level in the dam and predicts risk of floods through the canal using back propagation mechanism. Rainfall data, run off, ground infiltration parameters of the watershed is measured. Batch learning model is used for training. It delivered visualization of predicted results before hand for a month.

## III. SYSTEM DESIGN AND DESCRIPTION

The system makes use of a dam shutter status prediction at the primary level succeeded by another prediction depending on status of the dam and directing to a webpage-based user interface to map regions as inundation zones. Fig.1 depicts the block diagram of the entire system.
Flood prediction is carried out by the system as 2 stages.
1. Malampuzha Dam shutter open/close prediction: Supervised machine learning approach is utilized. The feature vectors include Relative Humidity, Rainfall, Current Reservoir level, Current Live Storage and class target can be either 0 or 1 depending on closure/open status of the shutters of the dam respectively. SVM model accurately predicts the target depending upon the inputs. If target is 0, the second prediction need not be executed.
2. Determine possibility of inundation in different areas: Depending on the state of operation of dam and weather parameters, the raising of shutters can have different impacts in the nearby areas. Taking into consideration the rate of flow ($cm^3/s$) and shutter opening height (cm), the impact can be categorized into 4. 0 denoting the safest, 1 denoting a warning, 2 having mild chance of inundation and 3 denoting inundation prone area. Vulnerable areas can be divided into 4 zones being green, yellow, orange and red depending on possibility of inundation.

### A. Prediction Model

In the prediction model, the dataset consists of current live storage (BCM), current reservoir level (m), rainfall (mm) and relative humidity (%) as feature vectors and open/close status of dam shutters forming the label. The target variable has values of 0 and 1 where, 0 indicates the shutter closed condition and 1 represents shutter open condition. Data was collected from Malampuzha Irrigation Department, for a period from 2016-2019. During data preparation, the relationship between different variables was analyzed and data mining was carried out. On carrying out spotchecking on number of algorithms, SVM algorithm delivered the highest mean estimated accuracy. SVM model was thus chosen and the database inputted to SVM based machine learning model spans 1168×5 matrix. The labelled features were analyzed using scatter plots and heat maps. In our model, 80 percent data has been used for training and 20 percent data has been subject to testing. SVM model performs classification, regression and outlier detection. The train dataset is subject to training on Support Vector Classifier (SVC). Since the prediction model involves 4 labeled feature vectors, there are 4 classifiers corresponding to each of the combinations. The SVM algorithm, finds points closest to the line from both the classes, forming the support vector. SVM also performs nonlinear classification. The distance between line and the support vectors is found and goal is to maximize the margin. Optimal hyperplane is determined. After data has been trained, evaluation has been carried out. Test data has been predicted using generated classifiers. Cross validation has been used to test the effectiveness of model.

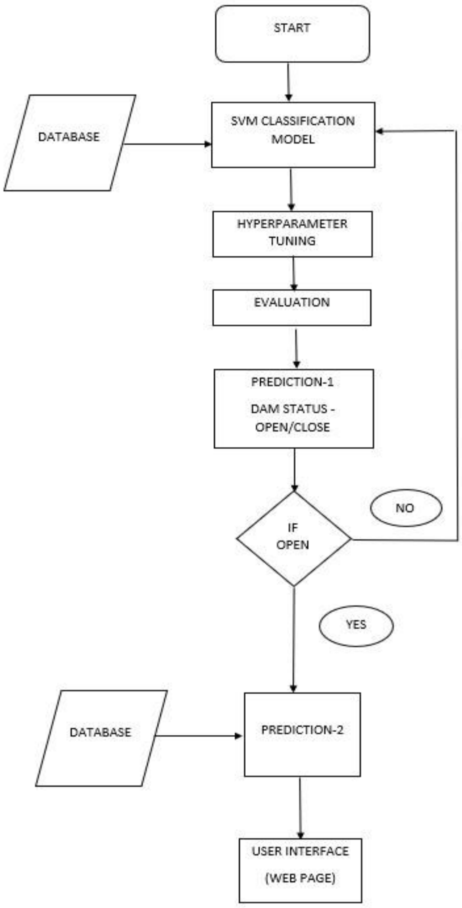


**Fig. 1. Block Diagram**

The prediction model after SVC classification, on evaluation produced a prediction accuracy of 94.8 percent.
Approach to improve the prediction accuracy was done by hyperparameter tuning using gridsearchcv. Grid search algorithm is directed by performance metric, and it builds and evaluates model for each combination of algorithm parameters specified in a grid.

Performance of entire model is dependent on the hyper parameter values specified. On our model, grid searching is carried out to find the optimized values for tuning parameter (C), and parameter for implementation in the most suitable kernel.

On fitting the model, radial kernel was found to be optimal and best grid parameters were found as- 'C': 1000, 'gamma': 0.01, 'kernel': 'rbf'. After grid fitting, on carrying out evaluation, the prediction model produced a prediction accuracy of 99.14 percent. Confusion matrix and accuracy report for the evaluation has been created.

A prediction model is designed at the secondary level to carry out the predictions for mapping the inundation zones if the shutter of dam is found to be open. The dataset for the second level of prediction spans 152×5 consisting of the 4 independent feature vectors and label-0, 1, 2, 3. By analyzing the feature vectors, classification is done into 4 levels (safe, mildly safe, warning and vulnerable). The labels have been created depending on the shutter height (cm) and flow rate(c/s) datas obtained from the data collected. Keeping in view of the limited dataset available as shutter open conditions, the entire dataset has been used to train the model. Training is done using svc model with hyperparameter tuning using gridsearchcv. Best grid parameters were found as- 'C': 1000, 'gamma': 0.001, 'kernel': 'rbf'. On training set, the model delivered a prediction accuracy of 99.3 percent.

### *B. Model Deployment*

Model deployment has been carried out using Flask server. Flask is a built-in development server and debugger. It offered integrated unit testing support. Fig 2 represents the basic block diagram for model deployment using Flask.

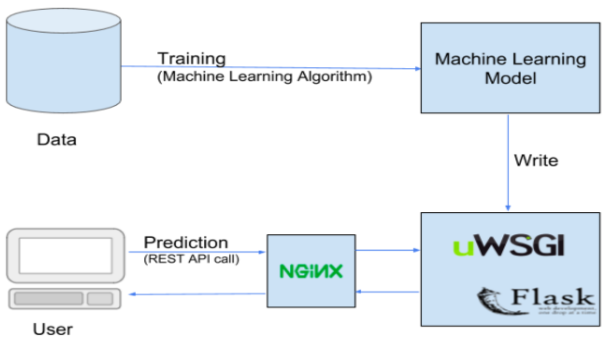


Fig. 2. **Model deployment to flask**

The developed webpage for our model is shown.

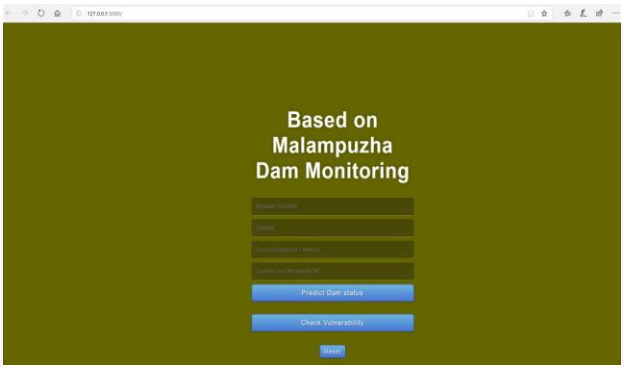


**Fig. 3. Webpage**

A registered user is able to access the webpage. The user has to input values which include parameters such as Current Reservoir Level, Current Live storage of dam, Rainfall and Relative humidity, which is carried over to the prediction model in the back end, and subject to processing. Thus, the Flask APIs receives these inputs through GUI or API calls and computes the prediction value based on our model and returns it. Prediction results are displayed in the webpage. If the dam shutter status is open, then the level of impact is displayed in terms of mapping of regions using different color streams. The webpage layout has been set using HTML template and CSS styling. The predicted status of dam being displayed in the webpage for different inputs is shown below (Fig.4(a), 4(b), 4(d)). When the predicted status of dam is open, the user is instructed to check the vulnerability. Upon clicking "Check Vulnerability' button, succeeding prediction is carried out on the same set of inputs and the prediction result is generated by mapping regions under different level-color alerts. The mapping of regions for inputs as in Fig. 4(b) and Fig. 4(d) has been shown in Fig. 4(c) and Fig. 4(e) respectively.

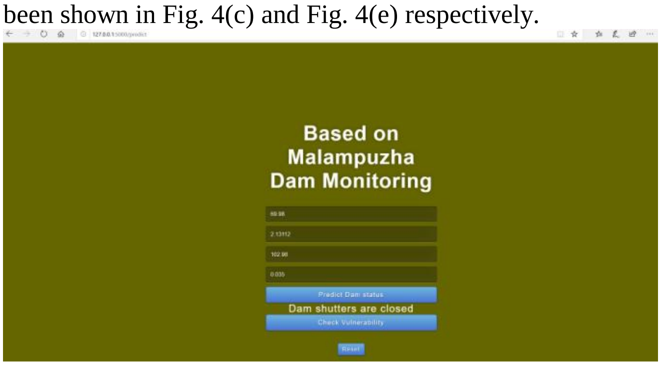


**Fig. 4(a). Predicted dam status**

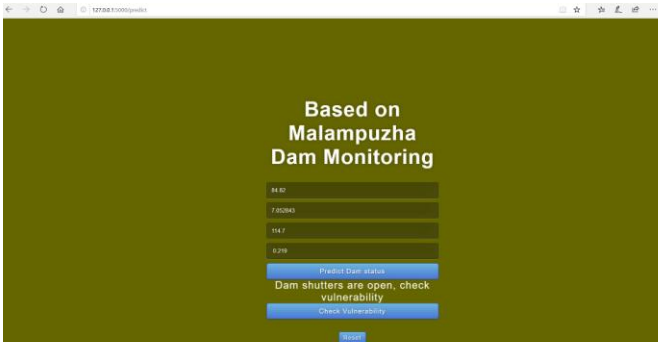


**Fig. 4(b). Predicted dam status**

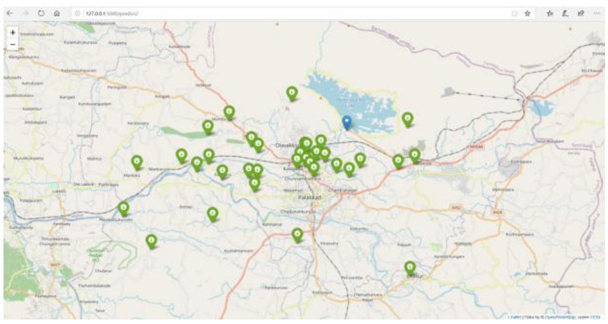

**Fig. 4(c). Mapping of regions**

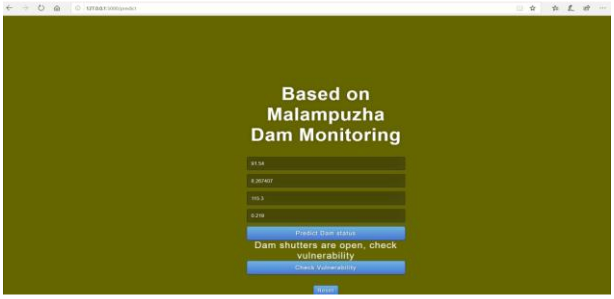


Fig. 4(d). Predicted dam status

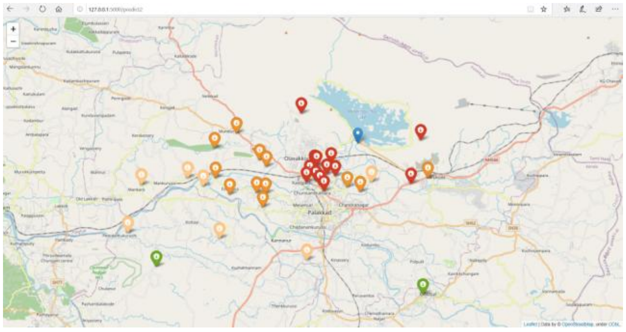

Fig. 4(e). Mapping of regions

## IV. RESULTS AND DISCUSSION

The confusion matrix in Fig. 5 is created after performing hyperparameter tuning on the SVM model for primary prediction. The entries in the confusion matrix is corresponds to test dataset. All cells in the matrix depict unique conditions. The first cell denotes true positive, second denotes false negative, third denotes false positive and last denotes true negative. Here, positive means dam shutter status is closed and negative means dam shutter status is open. The first cell in the matrix shows the correctly predicted dam closed status and the fourth cell shows the correctly predicted dam open status. The second and third cells show the incorrectly predicted cases. From the matrix it is clear that only 2 cases out of 234 test cases have been predicted incorrectly, which gives the primary prediction model a prediction accuracy of 99.14 % after hyperparameter tuning. A comparative study between accuracy report generated after SVM classification and accuracy report of model after using hyperparameter tuning is shown in Table- I.

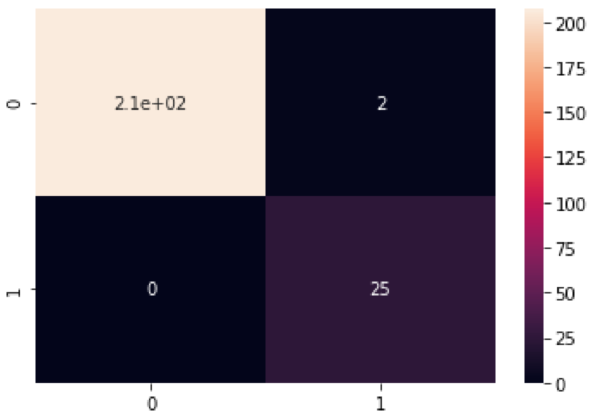


Fig. 5. Confusion matrix

**Table-I: Comparative study between accuracy report of SVM model before and after using hyperparameter tuning**

| Performance measures | | SVM | SVM with hyperparameter tuning | SVM | | SVM with hyperparameter tuning | |
|---|---|---|---|---|---|---|---|
| | | | | Macro Average | Weighted Average | Macro Average | Weighted Average |
| Precision | 0 | 0.95 | 1.00 | 0.94 | 0.95 | 0.96 | 0.99 |
| | 1 | 0.93 | 0.93 | | | | |
| Recall | 0 | 1.00 | 0.99 | 0.78 | 0.95 | 1.00 | 0.99 |
| | 1 | 0.56 | 1.00 | | | | |
| f1 score | 0 | 0.97 | 1.00 | 0.84 | 0.94 | 0.98 | 0.99 |
| | 1 | 0.70 | 0.96 | | | | |
| Support | 0 | 209 | 209 | 234 | 234 | 234 | 234 |
| | 1 | 25 | 25 | | | | |

## V.CONCLUSION

Natural disasters like floods are becoming more and more devastating every year due to the increase in rainfall and variations in other climatic factors. Floods in Kerala, being greatly influenced by emergency outflow from dams point to the need of an efficient system to monitor water level in dam. The losses can be minimized to an extend by effective early prediction and warning systems. The idea has been to develop a system to monitor dam parameters along with weather factors including rainfall and relative humidity and predict the possibility of inundation, specific to regions and may initiate evacuation whenever necessary. The model implemented using SVM after hyperparameter tuning has promised to provide an accuracy of 99.14% which is higher when compared to other models. Flask based web server is capable of generated mapping of regions depending on levels of warning corresponding to the set of inputs as inputted by the user. The system thus makes it possible to foresee possibility of inundation allowing authorities and people to undertake necessary measures to handle or prevent the crisis.

Future work may include addition of parameters which influence the target and thereby improving prediction accuracy.

## ACKNOWLEDGMENT

The authors take privilege in thanking the Malampuzha Irrigation Department for providing the necessary data required.

## REFERENCES

1. Nipon Theera -Umpon, "River Basin Flood Prediction using Support Vector Prediction," *IEEE* 2008
2. D Han, K Chan, N. Zhu "Flood forecasting using support vector machine", *IWA* publishing, 2007.
3. Areeha Durrani, Dr. Muhammad Khurram, Dr. HR Khan "Smart Weather Alert System for dwellers of different Areas", *IEEE* Conference 2019.
4. Wong Jee Khai, Moath Alraih, "Daily forecasting of dam level using Machine learning", *IJCIET* 2019
5. Azfina Putri Anindita, Puja Laksono, I Gusti Bagus Baskara Nugaraha, "Dam water Level Prediction system utilizing Artificial Neural Network Back Propagation", *IEEE* 2016

## AUTHORS PROFILE

**Meenu Thomas** currently a final year B. Tech student Of Electronics and Communication Engineering at NSS College of Engineering, Palakkad affiliated to APJ Abdul Kalam Technological University, she will be completing her undergraduation by 2020. She holds a student membership of ISTE and has a competent academic career. She has completed NPTEL online certification course on "Introduction to Programming in C" by IIT Kanpur. She has attended various workshops and internship program conducted by reputed organizations like Srishti Robotics. She has participated in a National Level Practical Workshop on IoT with Drone maneuvering organized by aeroengineers. She also attended lecture program on "Machine learning" organized by IEEE IAS Student Branch Chapter, NSSCE.

**Shuhaib Ummer** currently a final year B. Tech student Of Electronics and Communication Engineering at NSS College of Engineering, Palakkad affiliated to APJ Abdul Kalam Technological University, he will be completing his undergraduation by 2020. He is familiar with different programming languages because of his educational background in Computer Science. He has shown great interest in coding and has completed NPTEL online certification course for Data Science using Python. He holds an ISTE membership and has attended workshops conducted by reputed organizations like Srishti Robotics. He has participated in a National Level Practical Workshop on IoT with Drone maneuvering organized by aeroengineers.

**Sruthi Vijayan TK** currently a final year B. Tech student Of Electronics and Communication Engineering at NSS College of Engineering, affiliated to APJ Abdul Kalam Technological University, she will be completing her under graduation by 2020. Her educational background in Computer Science has given her a base in coding skills. She holds an ISTE membership and has a competent academic career. She has participated in an industrial training program conducted by BSNL. She has attended various workshops and internship program conducted by reputed organizations like Srishti Robotics. She has participated in a National Level Practical Workshop on IoT with Drone maneuvering organized by aeroengineers. She also attended a lecture program on "Machine learning "organized by IEEE IAS Student Branch Chapter, NSSCE.

**Vivek T** currently a final year B. Tech student of Electronics and Communication Engineering at NSS College of Engineering, Palakkad affiliated to APJ Abdul Kalam Technological University, he will be completing his under graduation by 2020. He shows immense interest in updating information about the emerging technologies and has shown good academic growth. He has attended a workshop on Android Controlled Robotics from the esteemed Indian Institute of Science (IISC), Bangalore. He has successfully completed an internship program conducted by Srishti Robotics from Cochin which included basic hands on training for various fields like microcontrollers, embedded system programming and raspberry pi.

**Ashok S Kumar** is currently working as an Assistant professor in Electronics and Communication Engineering at NSS College of Engineering, Palakkad affiliated to APJ Abdul Kalam Technological University. He has four years of experience as a college lecturer. He is the coordinator of Industry Institute Partnership cell (IIPC) of Electronics and Communication Department of NSS College of Engineering and was also member in various technical programs. The programs he coordinated also includes workshops with hands on training. He has published more than 5 research papers in various categories under journals and conferences. His area of interest includes cognitive radio networks, wireless communication systems and electromagnetics.