Semiconductor Optoelectronics, *Vol. 41 No. 12 (2022), 1288-1296*
https://bdtgd.cn/

Semiconductor Optoelectronics
ISSN: 1001-5868

RAINFALL MONITORING USING AN ARTIFICIAL NEURAL NETWORK AND ARIMA-BASED GROUNDWATER LEVEL PREDICTION

K Ajaykumar
Student,Department Of Computer Science And Engineering, K.S.Rangasamy College Of Technology, Tiruchengode-637215

Dr S.Madhavi
M.E Ph.D, Head Of Department, Department Of Computer Science And Engineering, K.S.Rangasamy College Of Technology, Tiruchengode-637215

Abstract

Wireless communication networks have arisen in recent years as a possible replacement for rain gauges, which have been a part of the monitoring infrastructure for decades. Empirical approaches to computing are losing their ability to accurately depict the real world as a result of enormous databases of precipitation and telecommunication networks. As a result, deep learning models for analysing large amounts of data have been proposed, as they provide more precise assessments of the real world. The results of several types of rainfall monitoring were investigated in this study. The main goal of this study is to create a neural network-based categorization system for rainfall data based on historical rainfall data production information. Farmers can benefit from a classification based on the quantity of rain that fell in prior years to help them make the required preparations to gauge crop productivity for the coming growing season. Food security and climate change mitigation can both benefit from a better understanding of and assessment of future crop growth. Auto regression analysis (AR) and moving average (MA) are both elements of the ARIMA forecasting model when applied to well-behaved time series data. Combining these two approaches results in the ARIMA model. In ARIMA, the time series is thought to be stable, with changes occurring at the same time each time. It's also thought that when it does change, it does so at around the same time every time. The ARMA (Auto Regressive Moving Average) technique was used to finish the job in this study. The precipitation and groundwater levels in our country are calculated using a decade's worth of data. The proposed research will use an ARIMA model to classify the ground water level data set records in order to predict the model for future test record data sets. It will assist us in determining how much water there was in the past and how much there will be in the future.

Keyword : Auto Regressive Moving Average, Rainfall, Accuracy, KNN, crop yield

1.Introduction

One of India's largest economic contributors is the agricultural sector. It is essential to the growth and sustainability of rural communities. A decrease in agricultural output may be caused by unpredictable precipitation, climate change, overuse of pesticides, and other related factors. With the help of past climatic and output data, this project aims to create a system for producing agricultural yields. Farmers can increase crop production for the future season by using a crop yield estimate based on the temperatures and precipitation totals of the previous year. Crop yields can be better understood to secure food supplies and lessen the impact of climate change.

The researchers made an effort to develop a method for accurately forecasting crop yields in the future using past data on temperature and precipitation. We set out to gain as much knowledge about crop output from only two variables—temperature and precipitation—in order to maximise our efforts. Precipitation is difficult to estimate accurately due to a number of factors, such as cloud cover, evapotranspiration, and many others. However, we wanted to complete all of this with just weather data.

The Auto Regressive Moving Average (ARMA) and Seasonal ARIMA models can be used to predict temperatures. These models are excellent at forecasting future temperatures using past data. Using a time series created from the dataset, the model forecasts temperatures. Precipitation forecasts can be made using both the ARMA model and the ARMA model enhanced by exogenous variables (ARMAX). The ARMAX model is used to account for many factors, including cloud cover, temperature, and evapotranspiration, in the case of precipitation. To produce precise yield projections, we used a framework based on fuzzy logic. The season's yield is calculated using a fuzzy model that takes into consideration the expected values with the fewest possible errors.

The creation of an intelligent agricultural system for crop production is demonstrated by Luis Omar Colombo-Mendoza et al (2022). The components of this system are low-cost Internet of Things sensors and well-known cloud services for data analysis and storage. Machine learning is heavily utilised in the Sangeeta and Shruthi G (2020) project, and methods including Random Forest, Polynomial Regression, and Decision Tree are used to assess performance. The random forest algorithm offers the supplied model's most precise yield forecast when compared to the other two methods. A data warehouse solution for crop farming was covered in a 2018 publication by Vuong M. Ngo and his colleagues. Large amounts of data can be stored and analysed using the data warehouse (DW). reviewing and evaluating the most current uses of precision agricultural systems and the most urgent problems that need to be solved in order to build a successful precision agriculture data warehouse. The use of the Internet of Things (IoT) in a poly house was proposed in 2018 by Rahul Dagar, Subhranil Som, and Sunil Kumar Khatri. A poly home is protected from the effects of the elements because it is completely enclosed. For instance, since insects cannot enter a poly house, the harvest is unaffected. It follows that you won't need to apply as many insecticides.

Alahi, M. E. E., Xie, L., Mukhopadhyay, S., and Burkitt, L. discuss their extensive research into the design and creation of a sensitive nitrate sensor for monitoring nitrate levels in surface and groundwater in this paper. A planar interdigital sensor, additional hardware, tools, and an electrochemical impedance spectroscopy-based test make up the developed flexible detecting system. Edge computing can be used for artificial intelligence to exchange cloud worker and capacity models, analyse information, and preprocess data. Edge registration is a technique for dealing with problems that appear when connected terminal devices to a blockchain do not have enough computing power or energy.

3. ARIMA-BASED GROUNDWATER LEVEL PREDICTION

For time series that can be made "stationary" using differencing (if necessary) and nonlinear transformations like logging and deflating, ARIMA models are the most versatile type of forecasting model. The most common type of time series prediction model is the ARIMA model, which has been used as a training model (if necessary). While a random variable with a stationary distribution changes with time, its statistical characteristics remain constant. A stationary series lacks a trend, fluctuates in amplitude about its mean on a regular basis, and moves in a known direction. This implies that the series' short-term random patterns of time appear to be identical from a statistical perspective. The stability of the power spectrum or the consistency of the variable's own correlations across time are the second requirements. Relationships between a variable's own historical standard deviations from the mean are known as autocorrelations. This kind of random variable can be thought of as a combination of noise and signal. If a signal is present, it can be interpreted as a seasonal pattern of rapid sign alternation, sinusoidal oscillation, or quick mean reversion.

The model is pure autoregressive, also referred to as a "self-regressed" model, if the only predictors are the lagging values of Y. It is essentially an extra regression model that may be fitted using the same tools as other regression models. A simple regression model with Y advanced by one period as the independent variable can be used to demonstrate the first-order autoregressive (AR(1)) model for Y. This model is shown in Statgraphics as LAG(Y,1) and in RegressIt as Y LAG1. An ARIMA model is NOT comparable to a linear regression model if any of the predictors have error delays. This is due to the fact that an ARIMA model cannot use the "error from the prior period" as an independent variable. The errors for each period must be estimated before the model can be fitted to the data. While the model's predictions are not linear functions of the coefficients, trailing errors are linear functions of the data before them. It poses a technological challenge when attempting to use this as a predictor. Therefore, nonlinear optimization methods must be used in place of solving a series of equations to get coefficients for ARIMA models with lagged errors ("hill-climbing").

Auto-Regressive Integrated Moving Average is referred to by its acronym. Moving Average That Is Auto-Regressive (ARIMA) A "integrated" variant of a stationary series is a time series that needs to be distinguished in order to become stationary. ARIMA models come in many different variations. Examples include autoregressive models, random-walk and random-trend models, and exponential smoothing methods.

A nonseasonal ARIMA model is a "ARIMA(p,d,q)" model, in which:

p denotes the number of autoregressive terms.

d is the number of nonseasonal deviations required for stationarity, and

The number of lagged forecast errors is q in the prediction equation.The forecasting equation is constructed in the following manner. Let's start by referring to y as the dth difference of Y, which equals:

If d=0, yt equals Yt.

If d=1, yt = Yt - Yt-1.

If d is equal to 2, then yt = (Yt - Yt-1) 2Yt-1 + Yt-2 = Yt - 2Yt-1 + Yt-2 = (Yt-1 - Yt-2)

Data from a Time Series

When working with time series data, each observation has a timestamp associated with it. We are unable to shuffle the data during training because it is organised in a sequential manner. The observations in the data are autocorrelated, which means they are closely related to one another and to the observations that came before them. Time-series data must be treated in a certain way due to these constraints.

Using time series data to predict rainfall

We can't utilise standard rain fall detection algorithms to find abnormalities in time series data since the data won't allow it. Detecting fake time series news is a challenging task that must be done separately from other tasks.

To see if time-series data indicates rain, follow these steps:

• Keep track of whether or not the rainfall data changes. This status must be applied to the rain fall data if it is not already immovable.

Using the study's conclusions as a reference, fit a time series model to the preprocessed data.

• For each data set observation, calculate the total squared error.

• Decide how much data error you're willing to put up with. We can identify rain fall when the number of errors within an observation surpasses a certain threshold.

Based on what we've learned so far, time-series false news data follows a strict pattern and is distributed similarly. Time-series models would be developed and used to learn the general behaviour of the fictitious data and to forecast the actual data using the fictitious data. To accomplish this, the models would be taught to work with fake data before being used. If an observation is correct, the prediction will be as close as possible to the actual value. If an observation is incorrect, on the other hand, the forecast will be as far away from the actual value as feasible. As a result, if we look at the predicted errors, we might be able to find the rain in the data..

4.Result and analysis

The information came from the Kaggle source. The science of teaching machines to learn and develop models that can predict the future is widely used, and for good reason. Agriculture is extremely important to the world economy. It is critical to assess how much food is produced globally in view of the growing global population in order to address food security issues and reduce the effects of climate change. Table 1 represents the data set information

Table 1 Data Set Information

S.no	Feature name
[1].	Area
[2].	Year
[3].	average_rain_fall_mm_per_year
[4].	avg_temp
[5].	Domain Code
[6].	Domain
[7].	Area Code
[8].	Area
[9].	Element Code
[10].	Element
[11].	Item Code
[12].	Item

[13].	Year Code
[14].	Year
[15].	Unit
[16].	Value

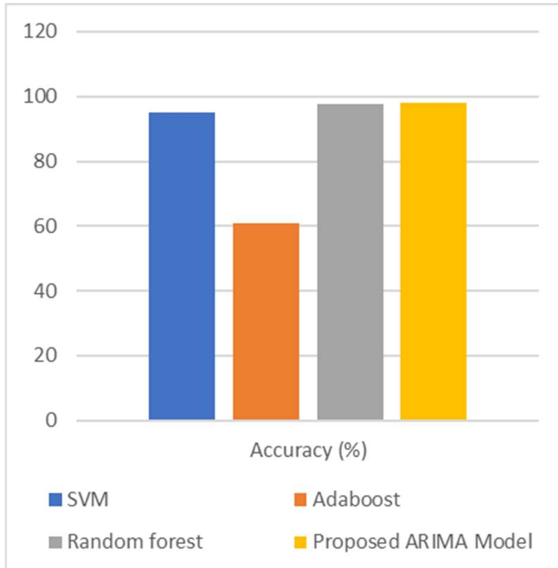

Figure 1 Accuracy Comparison Results

Crop yield prediction is an important part of agriculture. Climate (rainfall, temperature, etc.), pesticides, and previous crop yield are the key elements that determine agricultural yield. Figure 1 describes the accuracy comparison with related work. Table 2 represents the Accuracy comparison results of Adaboost, SVM, deep forest and modified deep forest

Table 2 Accuracy comparison

Methods	**Accuracy (%)**
SVM	95.2
Adaboost	60.9
Random forest	97.8
Proposed ARIMA Model	98

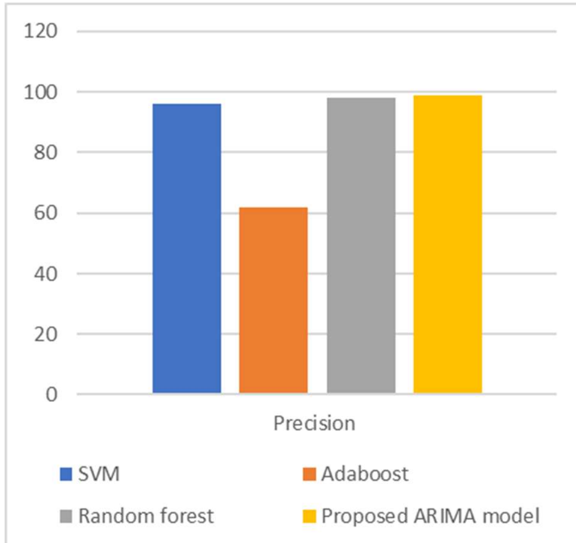

Figure 2 Precision Comparison Results

This is necessary for making decisions about agricultural risk management and forecasting the future. Figure 2 describes the precision comparition with related work. Table 3 representas the precision comparison results of Adaboost, SVM, deep forest and modified deep forest

Table 3 Precision comparison

Number of images	Precision
SVM	96
Adaboost	61.9
Random forest	98
Proposed ARIMA model	99

Figure 3 describes the recal comparition with related work. Table 4 representas the precal comparison results of Adaboost, SVM, deep forest and modified deep forest

Table 4 Recall comparison

Methods	Recall
SVM	96.7
Adaboost	77.4
Random forest	98
Proposed ARIMA model	99

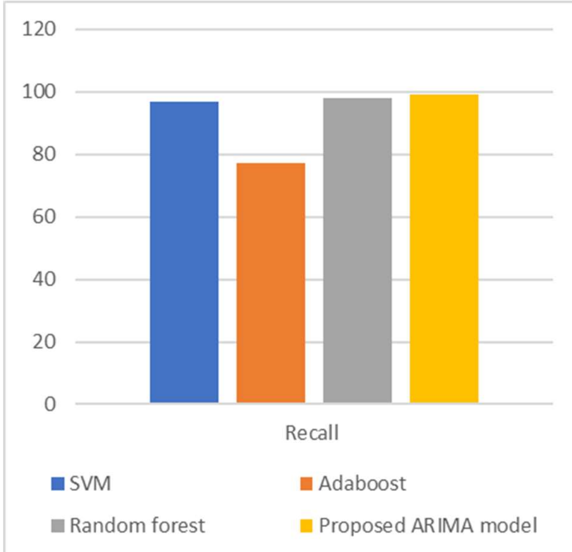

Figure 3 F-measure comparison Results

Figure 4 describes the f-measure comparition with related work. Table 5 representas the precision comparison results of Adaboost, SVM, deep forest and modified deep forest

Table 5. F-measure comparison

Methods	F-Measure
SVM	97.1
Adaboost	68.79
Random forest	98.3
Proposed ARIMA model	99

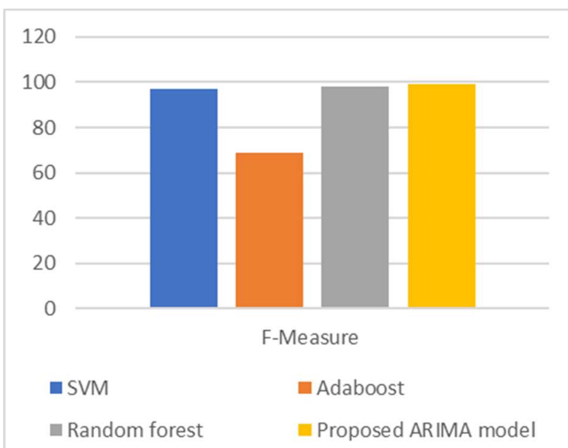

Figure 5 F-measure

5. Conclusion

The results show that the ARIMA model outperforms the ARMA model in terms of predicting temperature and precipitation. Rainfall predictions, despite their importance in predicting how much food a crop will produce, are difficult to get right. There are several more variables that can affect the accuracy of rainfall forecasts.. If the previous year's ground water level is known, the model can accurately predict the current year's ground water level. Furthermore, KNN is used in this study to classify the ground water level data set records so that a model prediction for future record test data sets may be made. It will assist us in determining how much water there was in the past and how much there will be in the future. In the future, a technique known as "logistic regression" could be used to further categorise the data. One of the simplest examples of machine learning is the K-Nearest Neighbor algorithm, which employs the Supervised Learning technique. The KNN approach is based on the assumption that the current case or data is analogous to previous cases. The new example is then assigned to the category that is most similar to the current categories. The KNN algorithm keeps track of all available data and determines how a new data point should be classified based on how similar it is to previous data. Because of the KNN algorithm, it will be simple and quick to place new data into a category that is a good fit in the future.

References

[1]. Aziz, A.A., Sekercioglu, Y.A., Fitzpatrick, P., Ivanovich, M., 2013. A survey on distributedtopology control techniques for extending the lifetime of battery powered wirelesssensor networks. IEEE Commun. Surv. Tutorials 15 (1), 121–144. https://doi.org/10.1109/SURV.2012.031612.00124.

[2]. Baltej Kaur , Danish Inamdar, Vishal Raut,Akash Patil,Nayan Patil, A Survey On Smart Drip Irrigation System International Research Journal of Engineering and Technology (IRJET) Volume: 03 Issue: 02 | Feb-2016.

[3]. Bhangu, B.S., Bentley, P., Stone, D.A., Bingham, C.M., 2005. Nonlinear observers forpredicting state-of-charge and state-of-health of lead-acid batteries for hybrid-electric

[4]. Bouzekri Amel, Chabane Mohamed, Benahmed Tarek, smart irrigation system using Internet of Things The Fourth International Conference on Future Generation Communication Technologies (FGCT 2015).

[5]. Burrell, J., Brooke, T., Beckwith, R., 2004. Vineyard computing: sensor networks inagricultural production. IEEE Pervasive Comput. 3 (1), 38–45. https://doi.org/10.1109/MPRV.2004.1269130.

[6]. C.Thulasi Priya, K.Praveen, A.Srividya , proposed a system for monitoring Pest Insect Traps Using Image Sensors & Dspic, International Journal Of Engineering Trends And Technology- Volume 4 Issue 9-September2013

[7]. Chen, M., Rincon-Mora, G.A., 2006. Accurate electrical battery model capable of predictingruntime and i–v performance. IEEE Trans. Energy Convers. 21 (2), 504–511.https://doi.org/10.1109/TEC.2006.874229.

[8]. Chen, W., Lin, Y., Lin, Y., Chen, R., Liao, J., Ng, F., Chan, Y., Liu, Y., Wang, C., Chiu, C.,Yen, T., 2019. Agritalk: Iot for precision soil farming of turmeric cultivation. IEEEInternet Things J. 6 (3), 5209–5223. https://doi.org/10.1109/JIOT.2019.2899128.

[9]. Daskalakis, S.N., Goussetis, G., Assimonis, S.D., Tentzeris, M.M., Georgiadis, A., 2018. Auw backscatter-morse-leaf sensor for low-power agricultural wireless sensor networks.IEEE Sens. J. 18 (19), 7889–7898. https://doi.org/10.1109/JSEN.2018.2861431.

[10]. Fiona Edwards Murphy_, Michele Magnoz, P´adraig Whelany, and Emanuel Popovici, b+WSN: Smart Beehive for Agriculture, Environmental, and Honey Bee Health Monitoring -Preliminary Results and Analysis , 978-1-4799-6117-7/15/$31.00 ©2015 IEEE.

[11]. G. Parameswaran, K.Sivaprasath, Arduino Based Smart Drip Irrigation System

Using Internet of Things DOI 10.4010/2016.1348 ,ISSN 2321 3361 © 2016 IJESC.
[12]. Gao, L., Liu, S., Dougal, R.A., 2002. Dynamic lithium-ion battery model for system simulation.IEEE Trans. Compon. Packag. Technol. 25 (3), 495–505. https://doi.org/10.1109/TCAPT.2002.803653.
[13]. Gayathri.R, Saranya.B, Binu., Lavanya, Devi, Optimized Equipment for measurement of Soil Parameters and Conservation of Water in Agricultural Fields,.International Journal of Innovative Research in Computer and communication Engineering,Vol.4,Issue 6,June 2016.
[14]. Grove - moisture sensor. http://wiki.seeedstudio.com/Grove-Moisture_Sensor/.Han, D., Lim, J., 2010. Design and implementation of smart home energy managementsystems based on zigbee. IEEE Trans. Consum. Electron. 56 (3), 1417–1425. https://doi.org/10.1109/TCE.2010.5606278.
[15]. Ivanov, S., Bhargava, K., Donnelly, W., 2015. Precision farming: Sensor analytics. IEEEIntell. Syst. 30 (4), 76–80. https://doi.org/10.1109/MIS.2015.67
[16]. Joaquín Gutiérrez, Juan Francisco Villa-Medina, Aracely López- Guzmán, and Miguel Ángel Porta-Gándara Smartphone Irrigation Sensor 1530-437X (c) 2015 IEEE.
[17]. Keerthi.v , Dr.G.N.Kodandaramaiah, cloud IoT Based greenhouse Monitoring System Int. Journal of Engineering Research and Applications ,ISSN: 2248-9622, Vol. 5, Issue 10, (Part - 3) October 2015, pp.35-41.