ISSN 2347-4289

Intelligent System For Predicting Agricultural Drought For Maize Crop

Solomon Mwanjele Mwagha, Peter W. Waiganjo, Christopher A. Moturi, E. Muthoni Masinde

Taita Taveta University Colege, Department of Mathematics and Informatics, Voi, Kenya; University of Nairobi, School of computing and Informatics, Nairobi, Kenya.
Email: soproltd@gmail.com

ABSTRACT: There has been little information in regard to agricultural drought prediction. This paper aimed at coming up with an efficient and intelligent agricultural drought prediction system. By using a case study approach and knowledge discovery data mining process this study was preceded by drought literature review, followed by analysis of daily 1978-2008 meteorological and annual 1976-2006 maize produce data both from Voi Taita-Taveta (Coast Province in Kenya). The design and implementation of an agricultural drought prediction system, was made possible by computer science programming for meteorological data preprocessing, classification algorithms for training and testing as well as prediction and post processing of predictions to various agricultural drought aspects. The study was evaluated by comparison of predicted with actual 2009 data as well as the Kenya Meteorological Department (KMD) 2009 records. The evaluation of this study results indicated consistency with the KMD 2009 outlook. The results showed that the application of classification algorithms on past meteorological data can lead to accurate predictions of future agricultural drought. The recommendation is that future work can be based on designing a solution for multiple regions with multiple crops.

Keywords : Agricultural drought, intelligent system, Knowledge discovery, nearest neighbor classification, Drought prediction

INTRODUCTION

Report by Apollo (2002) stated that agriculture is very important in Kenya as 75% of the country's population is dependent on agriculture for food and income; however, only about one third of the total land area of Kenya is agriculturally productive. Two thirds of the Kenya land is semi-arid to arid, and characterized by low, unreliable and poorly distributed rainfall. Report by Patrick, O. & Rosemary A. (2006) indicated that over 80% of the Kenyan population live in the rural areas and derive their livelihoods, directly or indirectly from agriculture. The development of agriculture is important for poverty reduction since most of the vulnerable groups like pastoralists, the landless, and subsistence farmers, also depend on agriculture as their main source of livelihoods. Report by Southtravels (2010) shows Kenya has climatic and ecological extremes with altitude varying from sea level to over 5000 meters in the highlands. Rainfall occurs seasonally throughout most parts of Kenya. Most parts of Kenya are subject to periodic droughts or delays in the start of the rainy seasons. Rainfall ranges from mean annual of less than 250 mm in arid and semi-arid lands to mean annual of greater than 2000 mm in high potential areas. Catholic Relief Services (2011) reported that failure of seasonal rains in Kenya is coupled with increased food prices leading to emergency food assistance from United Nations and Government of Kenya. Recurring agricultural drought leaves millions with little or nothing to eat. Agricultural drought is the major constraint of Kenyan agriculture sector severely affecting seasonal crops. As Kenya relies heavily on rain-fed agriculture, agricultural drought prediction can play a vital role, as it can provide necessary parameters to use in planning for agricultural drought mitigation measures. There is lack of regular information, education and social mobilization in strategic sectors to mitigate the agricultural drought related shocks. ICT tools continue to produce significant transformations in several sectors of the economy including agriculture. This research purpose was to take these transformations a notch higher; by developing an intelligent agricultural drought prediction system that integrates historical knowledge on droughts, maize production and climate data. The first step towards this research was based on investigations to ensure that there are relevant truths regarding rainfall patterns and agricultural droughts prospects. The design was preceded by the analysis of two sets of historical data; 1) maize production and 2) weather data. The aim was to identify agricultural drought patterns on historical data and use the results for the prediction of future agricultural droughts. The prediction was to provide information on agricultural drought occurrences onset/offset, intensity magnitude and the impact on crop.

Literature Review

Report by Apollo, B. (2002) indicated that two thirds of the Kenya land area is arid and semi arid (ASALs), and characterized by low, unreliable and poorly distributed rainfall. Patrick, O. and Rosemary, A. (2006) acknowledged that the development of agriculture is also important for poverty reduction since most of the vulnerable groups like pastoralists, the landless, and subsistence farmers, also depend on agriculture as their main source of livelihoods. Drought is a serious problems that significantly affect millions of people in the ASALs and it occurs when the rainfall and soil moisture are inadequate to meet the water requirements of crops. In a study by London school of hygiene & tropical medicine, (1986) on predicting famine researchers observed rain and crop data as well as human behavioral patterns with regard to famine. The researchers regarded human responses to drought such as migration, livestock sales, loans, and increase in grain prices as useful famine indicators. In Tanzania Ladislaus B. et al (2010), studied rainfall prediction using environmental indicators through appraisals, interviews, focus groups used to collect data while SPSS was used for analysis. Their study reported that local environmental indicators and astronomical factors pathology are widely used in the region to forecast rainfall. In China Gong Z, (2010), studied agricultural drought prediction. By analyzing the occurrence trend of agricultural drought by using grey catastrophic forecast models the study reported that serious drought can occur in 2012. The study offered decision basis for disaster prevention and risk reduction. Staff of Cook island department of Water Works, (2003) study aimed at monitoring of evolving drought conditions. By using 70 years daily rain data they developed a drought index that compared current condition and previous drought. The study allowed monitory of evolving drought condition as well as de-

ISSN 2347-4289

velopment of drought management plans. In China, Lin Zhu, (2008) studied monitoring drought losses and drought influence on agriculture. The researcher used soil moisture and daily metrological data were used as input to Borel Ecosystem productivity simulator (BEPS) to assess agricultural drought. The findings were that assimilated remotely sensed soil moisture in BEPS model improved the way of monitoring drought losses and drought influence on agriculture. In Kenya the European Commission, (2009) predicted drought using Food sec crop yield model. In their study they established crop yield, calculated FAO crop evapotranspiration. They also incorporated Land cover weighed normalized difference vegetative index (WNDVI). In San Francisco Celso A, (2009), studied drought forecasting. The study analyzed rainfall frequencies using data from 248 rain gauges (1938-2005). SPI was determined using ANN feed feedfoward/back propagation algorithm. The findings showed that the result ANN suitable for drought forecast. In Iran Dostrani, M.(2010) study aimed at comparing ANN and ANFIS in precipitation prediction. Dostrani study realized ANN efficient in rain prediction. Xin H. (2010) in Puyang studied predicting agricultural drought. Xin study used 1880-2005 rain data to analyze agricultural drought. By applying fuzzy sets analysis on the condition of crops and valid rain history, result of fuzzy clustering obtained. Drought years extracted from fuzzy clustering results. Time series used to predict next drought year. M Ashock, 2006 study in Kenya and Zambia aimed at translating seasonal forecast to agricultural terms. Ashock study used crop simulation model to translate seasonal forecast to agricultural terms. The results offered support to farmer's climate risk management. Z. Bob, 2007, in China studied rainfall prediction by direct determination of surface soil moisture using microwave observation. In Bob study data was acquired and analyzed over several test sites. The study was validated by conducted large field experiments. Niu S. (2006), Predicted agricultural drought in paddy fields using remotely sensed data. Niu study used and found NDVI to be reasonable in detecting agricultural drought. The study was limited by insufficient data as fuzzy was done in non cropping time. Kozyra, J. (2009) study evaluate metrological conditions causing drought using the differentiate between precipitation & evapotranspiration to evaluate metrological conditions causing drought. Tsegaye T. (2007) study in USA identified historical patterns for drought using VegOut Model that integrated Climate Ocean, satellite indicators; used regression trees to identify historical patterns for drought intensely and vegetation. SPI and PDSI were used to represent climate vulnerability. Tadesse study was evaluated using 2006 drought year. Unlike previous studies this paper contributes on prior work by considering drought literature, crop produce and weather data together with classification algorithms. Apart from providing historical agricultural drought analysis our work provides future projections with limit of twelve months agricultural drought predictions. Unlike previous studies this research emphasis is on prediction of agricultural drought using both historical and projected meteorological conditions. Our study provides user friendly output concerning the impact of agricultural drought on maize crop.

Methodology

Study Area

Using Taita-Taveta district as a case study area we applied the Knowledge Discovery and Data mining (KDD) process steps. There being no much study done on agricultural droughts prediction in Kenya and further agricultural drought prediction problem being rather complicated to analytically explain, the case study approach was the best to yield a rich picture of the situation, which may well be further subjected to comparative analysis. The study area chosen is classified as an arid and semi arid (ASAL) district in Kenya. Case study method enabled close examination of the data within the agricultural drought context. As indicated in the Taita-Taveta District profile, (2010) out of the total area of 17,128.3 Km2 covered by the district 24 per cent is range land suitable for ranching and dry land farming, while only 12 per cent is available for rain-fed agriculture. Of the 2,055.4 km2 arable land, 74 per cent is low potential agriculture land, receiving an annual mean rainfall of 650mm. The district lies between 2° 46' north to 4° 10' north and longitudes 37° 36' east to 30°14' east. The average temperature in the district is 23˚C. The district is divided into three major topographical zones. These are the upper zone, lower zone and volcanic foothills. The District experiences two rain seasons the long rains between the months of March and May and the short rains between November and December. The rainfall distribution is uneven in the district, with the highlands receiving higher rainfall than the lowland areas. The lowland areas, which are mainly ASAL, are only suitable for planting crops with short maturing period like sorghum, cowpeas, green grams, cashew nuts, sunflower, millet and dry land hybrid maize varieties. According to Kenya Seed Company, (2010) Maize crop is the most planted crop during the rain seasons. Pwani hybrids maize (PH1 and PH4) which are resistant to moisture stress, are considered as the most common varieties of maize that are grown in the area. The Pwani hybrids are more tolerant to moisture stress.

Data

To understand the application domain, a period of 1979 to 2008 daily Voi KMD historical dataset on minimum/maximum temperatures and precipitations and annual Kenyan Ministry of Agriculture 1976 to 2006 maize production dataset from Taita-Taveta district was obtained. The selection of samples of the datasets to use in analysis was done based on picking of data range without much of missing data. For the maize production dataset yearly production in tons and area cultivated in hectares were used as independent variables to determine production per hectare. In the meteorological dataset monthly average temperatures and monthly total precipitation were used as study variables to establish monthly, seasonal and annual weather conditions during the various years in the dataset. Missing data or gaps in climate were estimated from stations with similar climatic conditions. Data related to maize crop water requirements was determined using existing FAO crop growth models. Various analysis were done using spreadsheet.

Results

Table 1 shows analysis of precipitation during drought years. During the drought years the ratio of actual precipitation to normal went below 0.5 in one or both the seasons. Table 2 shows analysis of precipitation during non drought years. During the non drought years the ratio of actual precipitation to normal was above 0.5 in one or both the seasons. Climatologically precipitation prediction analysis results are shown in graph 1 with realization that 10 year sampling step is the most appropriate. Based on the analysis results the solution was

ISSN 2347-4289

designed using process logic and data flow diagrams and implemented using Java programming language and Weka knowledge discovery software. The implementation involved preprocessing output consisting of two sets of files as follows; Training file with attributes year, month, scaled precipitation values and class index (1979-2008) and Prediction file with attributes year, month, scaled precipitation values (2009). The Weka knowledge flow processing module produced output for Year, Month, Scaled precipitation values (range: 0 to 1), Precipitation class values, Index class (range: -2 to 2). Three classifiers (Isotonic Regression, K-nearest neighbourclassifier, and RegressionByDiscretization) were considered appropriate in working with the preprocessed training and testing sets since the classifiers could come up with desired output classes on processing. Classifiers comparisons were done by performing 10 fold cross validation and comparing the performance in relation to the actual as shown in table 3.

Discussions

The results on data preprocessing with comparison of 2009 prediction sets on various sample steps are depicted in graph 2. The outputs of classifiers were evaluated with actual data for year 2009 as shown in graph 3. The output of the Weka knowledge flow formed the input to post processing module that manipulated processed data to user understandable form. The end results of 2009 agricultural drought predictions are evaluated through comparison to actual drought situation in 2009. The consistency of solution was evaluated to Kenya Meteorological Department 2009 seasonal forecasts.

Conclusion

The results of this research show an improvement from previous researchers as the solution contributes to agricultural drought prediction with an emphasis specifically on agricultural drought for maize crop. The use of case study in analysis of data allowed design of a solution that is possible to test using data from other regions as well as investigations different seasonal crops. This paper demonstrated that the technique of combining meteorological data, crop production data, literature review on drought, FAO crop models, together with classification algorithms can be a feasible way of predicting agricultural drought and evaluate its impact on agriculture. The agricultural drought prediction output results obtained showed that the nearest neighbor classifier is a suitable tool for training meteorological data for agricultural drought output classes. As part of machine learning the IBk classifier results accomplished intelligence through the knowledge discovery and data mining process as aimed in the study major objective. Evaluation of results indicated a close relation in agricultural drought predictions with the outlook provided by KMD. The recommended future work is designing a solution that can cater for predictions in multiple regions for multiple crops; in Kenya for instance all districts can be represented and in each agricultural drought impacts evaluated for various seasonal crops. Possible adjustments of our solution output parameters socioeconomic measures on agricultural droughts anticipations can be suggested e.g. early warning allowing for government provision on importing calculated quantities of produce for crops believed to experience agricultural drought

References

[1]. Kenya Meteorological Department, Agrometeorological bulletin, Issue No. 27/2009.

[2]. Gong, Z. et al, 2010. Risk Prediction of Agricultural Drought in China. 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010).

[3]. Kozyra, J. et al 2009. Institute Of Soil Science and Plant Cultivation National Research Institute, International Symposium, Climate change and Adaptation Options in Agriculture, Viena, June, 22-23 2009.

[4]. Lin Zhu, Jing M. Chen, Qiming Qin, Mei Huang, Lianxi Wang, Jianping Li, Bao Cao: Assimilating Remote Sensing based Soil Moisture in an Ecosystem Model (BEPS) for Agricultural Drought Assessment. IGARSS (5) 2008: 437-440

[5]. Ladislaus B. et al, 2010. Indigenous knowledge in seasonal rainfall prediction in Tanzania: A case of the South-western Highland of Tanzania, Journal of Geography and Regional Planning Vol. 3(4), pp. 66-72, April 2010.

[6]. Ashok, M. et al, 2006. Linking Seasonal Climate Forecasts with Crop Simulation to Optimize Maize Management, CCSP Workshop: Climate Science in Support of Decision Making, 14-16 November 2005 Crystal Gateway Marriott Arlington, Virginia 14-16 November 2005

[7]. Dastorani, M. et al (2010). Application of ANN and ANFIS Models on Dryland Precipitation Prediction (Case Study: Yazd in Central Iran). Journal of Applied Sciences, 10: 2387-2394.

[8]. Niu Shulian; Susaki Junichi, 2006. Detection of Agricultural Drought in Paddy Fields Using NDVI from MODIS Data. A Case Study in Burirum Province, Thailand.

[9]. Patrick O. Alila and Rosemary Atieno, 2006. Agricultural Policy in Kenya: Issues and Processes, A paper for the Future Agricultures Consortium workshop, Institute of Development Studies, 20-22 March 2006.

[10]. Tsegaye T. & Brian W. 2007. The Vegetation Outlook (VegOut): A New Tool for Providing Outlooks of General Vegetation Conditions Using Data Mining Techniques. ICDM Workshops 2007: 667-672

[11]. Z. (Bob) Su, Y. Chen, M. Menenti, J. Sobrino, Z.-L. Li, W. Verhoef, L. Wang, Y. Ma, L. Wan, Y. He, Q.H. Liu, C. Li, J. WEN, R. van der Velde, M. van Helvoirt, W. Lin1, X. Shan, 2007. Drought Monitoring and Prediction over China, In: Proceedings of the 2008 Dragon symposium, Dragon programme, final results, 2004-2007, Beijing, China 21-25 April 2008.

[12]. Peter Reutemann, (2007). WEKA KnowledgeFlow Tutorial for Version 3-5-7, University of Waikato 2007

ISSN 2347-4289

Online References

[1]. Apollo Bwonya Orodho, 2002. Department of Range Management, Faculty of Agriculture and Animal Production, University of Nairobi, Kenya. http://www.fao.org/ag/AGP/AGPC/doc/counprof/Kenya.htm#1, accessed on 14th February 2011

[2]. Catholic Relief Services, 2011. http://crs.org/kenya/ongoing-hunger-crisis, accessed 14th February 2011

[3]. Kenya Seed Company, 2010. http://www.kenyaseed.com/maize.html, accessed on 14th February 2011.

[4]. Taita Taveta District profile, 2010. www.aridland.go.ke/semi_profiles/taita_taveta_profile.pdf, accessed on 25th January 2011, 25th January 2011

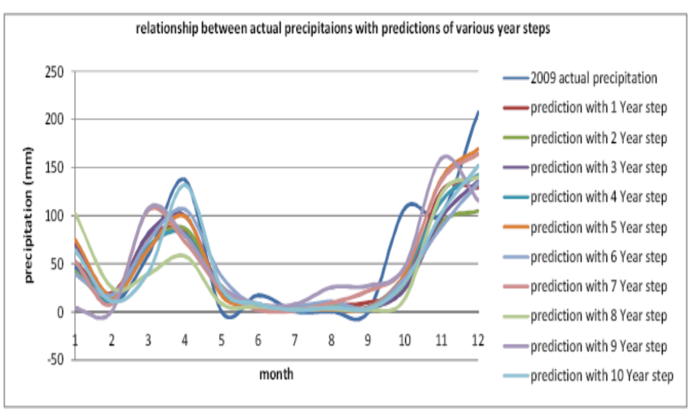
Appendix I: Tables and Graphs

Drought Year	Season 1 (mm)	Season 2 (mm)	Annual (mm)	Annual production (tones)	Area planted (hectares)	Production per hectare (tones)	Ratio of precipitation to normal (season 1)	Ratio of precipitation to normal (season 2)
1996	206.4	234	440.4	2538	4230	0.6	0.456637168	0.517699115
2005	158.7	88.1	246.8	3798	17464	0.217476	0.351106195	0.194911504

Table 1: Total precipitation during selected drought years.

Non Drought Year	Season 1 (mm)	Season 2 (mm)	Annual (mm)	Annual production (tones)	Area planted (hectares)	Production per hectare	Ratio of precipitation to normal	Ratio of precipitation to normal (season2)
1981	355.1	356.8	711.9	22615	16509	1.369859	0.785619469	0.789380531
1982	182.7	424.6	607.3	8946	8082	1.106904	0.40420354	0.939380531
1986	151.5	387.1	538.6	12556	7715	1.627479	0.335176991	0.856415929
1991	380.9	139.9	520.8	8146	4007	2.032942	0.842699115	0.309513274
1997	217.3	495	712.3	7765	6488	1.196825	0.480752212	1.095132743

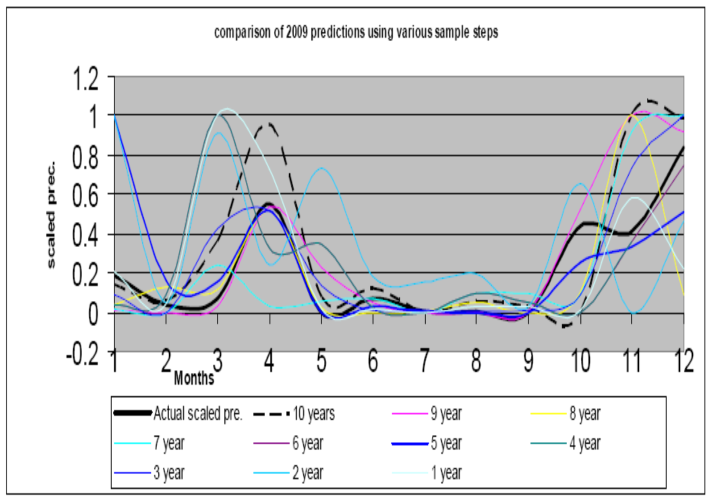
Table 2: Total precipitation during selected non drought years.



Graph 1: Climatologically 2009 precipitation predictions using various sample years.

Month	Actual class	Class predicted by classifiers					
		Ibk		Isotonic Regression		Regression By Discretization	
		5 Step	10 Step	5 Step	10 Step	5 Step	10 Step
Jan	-0.5	2	-1	2	-1	2	-1
Feb	-1.5	-1	-1.5	-0.5	-1.5	-0.5	-1.5
Mar	-1.5	-0.5	0	-0.5	0	-0.5	0
Apr	0.5	0	2	0.5	2	0.5	2
May	-2	-1.5	-1.5	-2	-1.5	-2	-1.5
Jun	-1.5	-1.5	-1.5	-1.5	-1	-1.5	-1
Jul	-1.5	-2	-2	-1.5	-1.5	-1.5	-1.5
Aug	-1.5	-1.5	-1.5	-2	-1.5	-2	-1.5
Sep	-2	-1.5	-1.5	-2	-1.5	-2	-1.5
Oct	0	0	-1.5	-0.5	-2	-0.5	-2
Nov	0	0	1.5	0	2	0	2
Dec	1.5	0.5	1.5	0.5	2	0.5	2
Error of the predicted values		0.0027		0.0027		0.0108	
Root relative squared error		3.7048%		3.7127%		7.4096%	
Time taken to train/build model		0 seconds		0.03 seconds		0.05 seconds	

Table 3: Comparison of actual class to classes predicted by various classifiers



Graph 2: Comparison of various predictions sets to 2009 actual.