# CROP YIELD PREDICTION USING TIME SERIES MODELS

**Askar Choudhury, Illinois State University**
**James Jones, Illinois State University**

## ABSTRACT

*Climate and other environmental changes in the developing world and the African continent has become a major threat to their agricultural economy. Traditional insurance for financial risk management is impractical in developing countries because of high transaction costs, adverse selection, information asymmetry, poor distribution and other challenges which hinder the availability of protection. Area-based index insurance is viewed as a promising financial risk management solution for smallholder farmers in developing countries, such as, Ghana. However, estimating the yield (i.e., yield prediction) is a critical part of pricing the premium for this insurance instrument. Because of the importance of predicting crop yield, the purpose of this study is to apply several forecasting methods for evaluating crop yield estimates in Ghana. Crop yield forecasting, which provides information for decision makers, is important in many ways to Ghana's economy. We compare yield forecasts using Simple Exponential Smoothing, Double Exponential Smoothing, Damped-Trend Linear Exponential Smoothing, and ARMA models applied separately to each district. The ARMA models proved to be more robust time-series models than the smoothing techniques for predicting crop yield in this study. This predictive power of ARMA models even with the presence of crop yield "cycle" does not depend on the length of cycle. Therefore, the results of this study indicate that the ARMA model is preferable over other time series models considered in this paper. The implication of the findings in this study is significant for insurance underwriters responsible for constructing area-based yield insurance that can benefit the Microinsurance market of smallholder farmers and for institutions that rely on those forecasts in providing capital.*

## INTRODUCTION

Farming is a major source of income for many people in developing countries. In Ghana farming represents 36 percent of the country's GDP and is the main source of income for 60 percent of the population (http://earthtrends.wri.org, 2003 p. 1). In addition, agricultural growth in Ghana has been more rapid than growth in the non-agricultural sectors in recent years, expanding by an average annual rate of 5.5 percent, compared to 5.2 percent for the economy as a whole (Bogetic et al., 2007). As with other parts of the developing world and the African continent, climate and other environmental changes in Ghana has become a major threat to their agricultural economy (Etwire et al., 2013). Direct losses to farming include destruction of their assets (such as, crop, livestock) which push poor farmers into poverty traps from which they have little means of recovery. Indirect impacts include sub-optimal management of this financial risk exposure, for example by selecting low-risk, low-return asset and activity portfolios that reduce the risk of greater suffering, but limit growth potential and investment incentives, selling assets (at inopportune times), reducing nutrient intake, and withdrawing kids from school and hiring them out to work. The problem is exacerbated by the reaction of financial institutions, which may restrict lending to farmers to minimize exposure to agricultural risk. These indirect consequences hinder economic growth (Barnett et al., 2008).

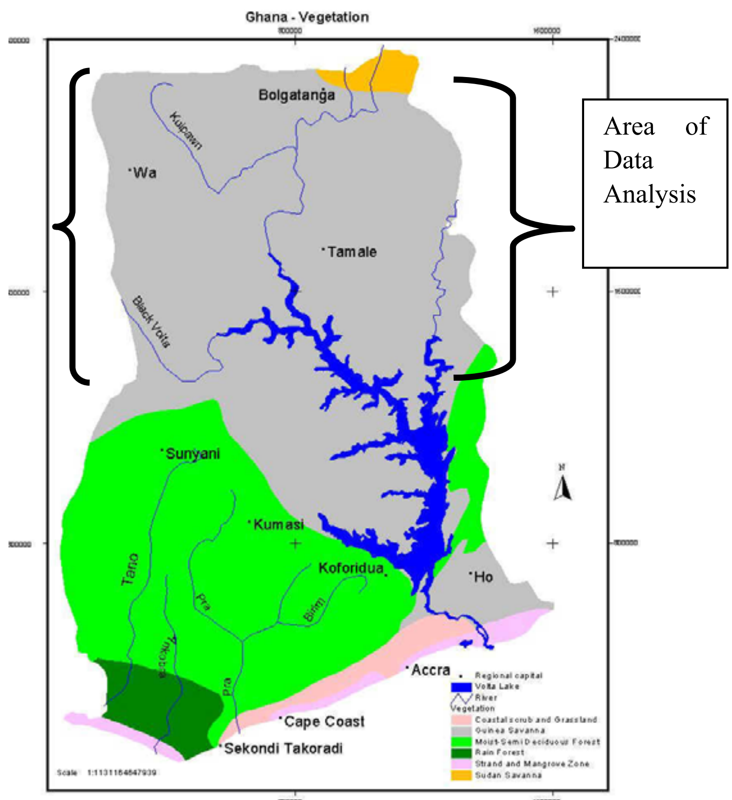
Traditional insurance is impractical in developing countries because of high transaction costs, adverse selection, information asymmetry, poor distribution, and other challenges which hinder the availability of protection (Skees, 2008). Furthermore, post-event response in the form of emergency aid, debt forgiveness, and grants are at risk following recent economic crises, and such public capital does not usually help create independent private solutions and can be inequitable and untimely. In recent years, index based insurance instruments have been piloted as a way for smallholder farmers to hedge their losses. Unlike traditional indemnity insurance, the payout on index insurance products is not based on actual farm level yield and/or revenue losses. It is rather based on realizations of an index which assumes correlations with actual farm yield (or revenue) losses. Since the indexes are based on objective and transparent sources of data, it is unlikely that informational asymmetries exist that can be exploited by index insurance contract purchasers. Thus, the inherent insurance problems of adverse selection and moral hazard, additionally the high transaction costs of implementation can be largely avoided (Deng et al. 2006). Index insurance may also have the benefit of crowding-in capital, and allow farmers to get loans for needed inputs, as the risk for agricultural losses and thus financial risk becomes more manageable (Carter et al. 2007).

The two types of index products are parametric and sample-based. Examples of parametric indices in insurance include weather (with triggers based on variables such as rainfall, temperature, humidity, wind speed, etc.), flooding (water levels and durations triggers), wind speed (velocity and duration triggers) and seismic activity (Richter scale triggers). Sample based indices include area based yield insurance and sample based livestock index insurance. Area yield insurance is essentially a put option on the average yield for a production in a region/area. Payouts are triggered by shortfalls in that area average yield rather than farm level yield. For this reason, area yield insurance requires no farm-level risk underwriting or loss assessment. If the area is sufficiently large, area yield insurance is not susceptible to moral hazard problems, since the actions of an individual farmer will have no noticeable impact on the area average yield. Area yield insurance also has relatively low transaction costs since there is no need to establish and verify specific farm yields for each insured unit nor is there any need to conduct on-farm loss adjustment.

Crop yield (in Africa and many other countries) is defined as metric tons of production per hectare or area cropped. For an area, such as a district in Ghana, the calculation requires a sampling of production for a crop for the entire district and dividing by the area cropped in that district for that given crop. Ghana's Ministry of Food and Agriculture (MoFA) conducts sample crop-cutting at a district level for maize, rice and other food crops throughout Ghana and reports their results to the Statistics, Research and Information Directorate (SRID) in Ghana (Stutley. 2010). The reliability of crop cutting, in developing countries, is sometimes questioned because of variations in resources and expertise available. More reliable resources and more accurate sampling techniques, video recording crop cutting experiments with GPS-enabled cell phones and remote sensing via satellite imagery, are new ways being piloted to help inform area yield estimations and make them more reliable.

**Selection of Crop Region**

Ghana produces a variety of crops in various climatic zones which range from dry savanna to wet forest. Agricultural crops including yams, grains, cocoa, oil palms, groundnuts and timber form the base of Ghana's economy. This research is focused mainly on the northern part of Ghana where there is substantial farming activity. The northern region of Ghana is considered the major bread basket of the country, and is also the most susceptible to the vagaries of the weather, especially the lack of rainfall. The northern part of Ghana is made up of three main regions; the Upper West Region, the Upper East Region and the Northern Region. The largest of these is the Northern Region which incidentally is the largest region in Ghana, covering a land area of about 70,383 square kilometers. However, it has the lowest population density of all the ten regions in the country (PPMED, Ghana, 1991) with 80% of its people dependent on farming. The major food crops grown here are yam, millet, rice, maize, sorghum, soybeans, groundnut and cassava.




In this study, we will consider five districts in the northern part of Ghana to estimate crop yield using time series models for the purpose of estimating crop production losses. Crops in this area are almost 100 percent rain fed (Stutley, 2008). Ghana is a country that is politically stable, has relatively good data and favorable regulation. A well designed financial risk management system in the agricultural sector could allow Ghana to act as a gateway to Africa for insurance underwriters who are not currently participating in Africa. As foreign donors have become increasingly diligent in assessing the need for loans and emergency relief, a credible index tied to true economic loss could be used by Ghana in justifying the need for emergency aid, loans and debt relief.

Accurate knowledge of crop yield behavior of the region is critical for devising such type of crop insurance product. Knowledge of the likelihood of yield and severity of yield shortfalls of the area are necessary components to create appropriate crop insurance. However, crop yield can be extremely dispersed from year to year and create complex scenario for predictability. Although understanding the stochastic nature of crop yield is important, characterizing yield behavior can be quite difficult. In general, historical yield distributions are used to set crop insurance premiums based upon the assumption that the following year's realization is drawn from the same distribution.

Because of the importance of crop yield prediction, the purpose of this study is to apply several forecasting methods for evaluating crop yield forecasting models. Crop yield prediction, which provides information to decision makers, is important in many ways to the economy. Because of its importance, researchers have proposed many forecasting methods to improve accuracy of yield estimates. However, obtaining accuracy is not an easy task, as many factors have impacts on crop production and thus crop yield. Many methods have been used in yield forecasts and different models have generated different results. The most widely used is the Box-Jenkins ARMA (autoregressive moving average) models. ARMA models have been used to forecast maize production in Nigeria (Badmus and Ariyo, 2011), wheat production in Pakistan (Najeeb et al., 2005), rice production in Ghana (Suleman and Sarpong, 2012) and rainfall in Ethiopia (Gerretsadikan and Sharma, 2011). As accuracy and simplicity is a big concern in projection, researchers have begun to explore other methods in their forecasting. These include Simple Exponential Smoothing (Boken 2000; Pal et al. 2007), Double Exponential Smoothing (Boken 2000; Pal et al. 2007), and Damped-Trend Linear Exponential Smoothing. These predictive models can be ranked by R-square and other model performance criteria. This method of model evaluation is then applied to five widely used time series models implemented in this paper. We find ARMA (autoregressive moving average) method outperform the competing methods in predicting crop yields in all five districts considered in this study.

Variations from the predicted farm-level yields are largely a function of systemic risk such as the pervasive drought or excessive rain (Halcrow, 1949). An area yield policy has an associated basis risk when farmers' experience farm-level yield losses while the area yield shortfalls are not sufficient to trigger a payout under an area yield policy. This occurs when shock losses are idiosyncratic. Area yield insurance provides more effective risk management where yield risks are largely systemic. Lowering the chances of such an event (i.e., lowering the basis risk) is an important objective when designing an area yield insurance policy. The magnitude of the basis risk is affected primarily by two elements of the contract design: (a) the area to be used for the yield index and (b) the procedures for forecasting the yields for the area (Skees et al., 1997). Crop yield distribution primarily consists of average yield and standard deviation of yields. We expect average yield to stay same over time if the factors that influence the yield also move in tandem. Similarly, variations in yield would be similar also if the factors themselves affecting the yield stay same. However, extreme changes in those factors, such as, weather (e.g., drought, flood, hail, etc.) can influence the crop yield adversely and widen the yield variance. Therefore, the purpose of crop insurance is to provide protection against yield shortfalls due to these natural hazards. Thus, a prediction model to estimate the crop yield that accounts for higher percentage of yield variations is a preferable estimation model.

## DATA AND RESEARCH METHODOLOGY

Data was collected from The Ministry of Food & Agriculture, which is the main government organization responsible for formulating and implementing agricultural policy in Ghana. The Statistics, Research and Information Directorate (SRID) and Policy Planning

Monitoring and Evaluation Division (PPMED) are two of the five directorates through which the ministry carries out its functions. According to information on the Ministry's website, the SRID has as some of its objectives "to initiate and formulate relevant policies/programs for creation of timely, accurate and relevant agricultural statistical database to support decision making" and "to conduct agricultural surveys and censuses covering major agricultural commodities". The PPMED, on the other hand, is responsible for undertaking, monitoring and evaluation of programs and projects under the Ministry. The statistical service department is an independent government department that is responsible for the collection, compilation, analysis, publication and dissemination of official statistics in Ghana for general and administrative purposes.

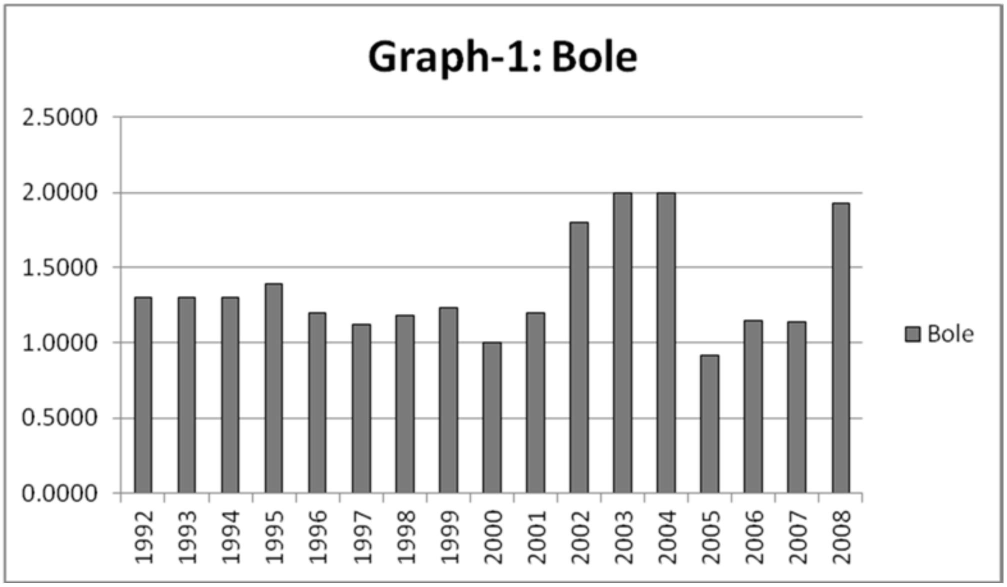
Crops which are likely to be suitable for Area-Yield Index Insurance include rain-fed maize and rice, and possibly millet, sorghum and groundnuts. This paper attempts to estimate the area "yield" of one crop, maize, for the purpose of creating an area-based index insurance instrument. Crop yield forecasting is primarily done with crop simulation models and empirical statistical regression equations relating yield with relevant predictor variables. These associative models require future data on the predictor variables. Crop forecasts are typically needed between the time of planting and the time of harvest. These associative models use past data to estimate the models and "future" data for prediction. Future data can be implicit or explicit. In general, forecasting methods can be subdivided into two categories: qualitative and quantitative (Makridakis et al., 1998; Armstrong, 2001) methods. Some of them are subjective, based on stakeholders' intentions or on the forecaster's or other experts' opinions or intentions, and others are objective/statistical, including univariate (extrapolation method), multivariate (associative method) and theory based methods. Other types may include expert systems or neural net, basically a variant of extrapolation with some subjective expert opinion.

Limitations of soil, weather and other relevant data cause a considerable uncertainty in the large area yield forecasting models (Hoogenboom, 2000; Russel and Gardingen, 1997). It is often unclear how these uncertainties transmit through the system given the non-linear behavior of crop yield models and the aggregation errors that may creep in when aggregating crop yields to larger regions (Hansen and Jones, 2000). Considerable amount of research to understand the effects of uncertainty in weather and other relevant factors on crop yield has been carried out by researchers. Crop yield modeling researchers primarily focused on local scale analyses in order to assess uncertainty in yield management (Bouman, 1994), condition of the soil (Pachepsky and Acock, 1998; Launay and Guérif, 2003), and weather components that affect crop yield (Fodor and Kovacs, 2005; Nonhebel, 1994; Soltani et al., 2004). In general, these studies demonstrate that the uncertainty in the modeling process is primarily a result of uncertainties in soil conditions and/or weather components. However, the local scale representation of these studies make the results less representative of regional scale crop yield forecast. Much of the research by climate researchers has been devoted to quantifying the climate variation effect on crop yield and studying the response of crop models to the climate change scenarios that are derived from general circulation models (GCMs). These research studies reveal that crop yield models are sensitive to the inconsistency of precipitation and temperature (Mearns et al., 2001; Semenov and Porter, 1995) and that the spatial scale of weather variables are also critical (Carbone et al., 2003; Mearns et al., 1999) to the crop yield prediction. In addition, when aggregating the yield at the regional scale, weather usually becomes the primary uncertainty factor compared to the soil (Easterling et al., 1998; Mearns et al., 2001).

| Table-1: Summary Statistics of Maize Yield | | | | | | |
|---|---|---|---|---|---|---|
| **Variable** | **N** | **Mean** | **Median** | **Std Dev** | **Maximum** | **Minimum** |
| Bole | 17 | 1.3625 | 1.2308 | 0.3475 | 2.0000 | 0.9134 |
| Damango | 17 | 1.2744 | 1.2000 | 0.5060 | 2.2898 | 0.1200 |
| Salaga | 17 | 1.1897 | 1.2006 | 0.3821 | 2.0000 | 0.4433 |
| Tamale | 17 | 1.1386 | 1.0000 | 0.3683 | 1.9000 | 0.6000 |
| Yendi | 17 | 1.1678 | 1.1000 | 0.2270 | 1.5785 | 0.7000 |
| Note: Crop yield was measured in Metric Tons per Hectare (Mt/Ha) in Ghana. Where, 1 hectare = 2.471 acres. | | | | | | |

Thus, to avoid these complexities we apply univariate time series methods to achieve simplicity in the model construction. In this paper, crop yield forecasting refers to univariate regional yield forecasts, i.e. forecasting of crop yield (metric tons of crop production per hectare) over large areas. The areas are administrative units called districts, as this is the scale at which most socioeconomic data and crop statistics are available to decision makers.

Table-1presents summary statistics of crop yield for five different districts in Ghana. Univariate time series methods were applied to predict the crop yield (Maize) using seventeen years of data. Average maize yields are more or less similar between districts. However, much variation exists in the maize yield between districts. Even though there are some similar trends observed in the yield plot over time (see, Graphs 1-5), the pattern is not systematic among the districts. As for example, “Damango” district has nine years of downtrend of crop yield that ended in 2003 (see, Graph-2). Similar down trend also exists with other districts that has ended in earlier years and thus makes these patterns non-systematic. To overcome these complex trend movements, we developed time series forecasting models that are applied separately to each district individually to capture the data pattern for that specific region. The following describes the concepts of different time series models briefly, which we have implemented in this research.




**Simple Exponential Smoothing:**

This technique is based on a series of averaging data in a decreasing (exponential) manner. The weights $\alpha$ is termed the smoothing constant which ranges from 0 to 1. The value of $\alpha$ determines the extent to which the most current observation influences the forecast. The simple exponential smoothing equation is expressed as,

$$L_t = \alpha y_t + (1-\alpha) L_{t-1} \ ,$$

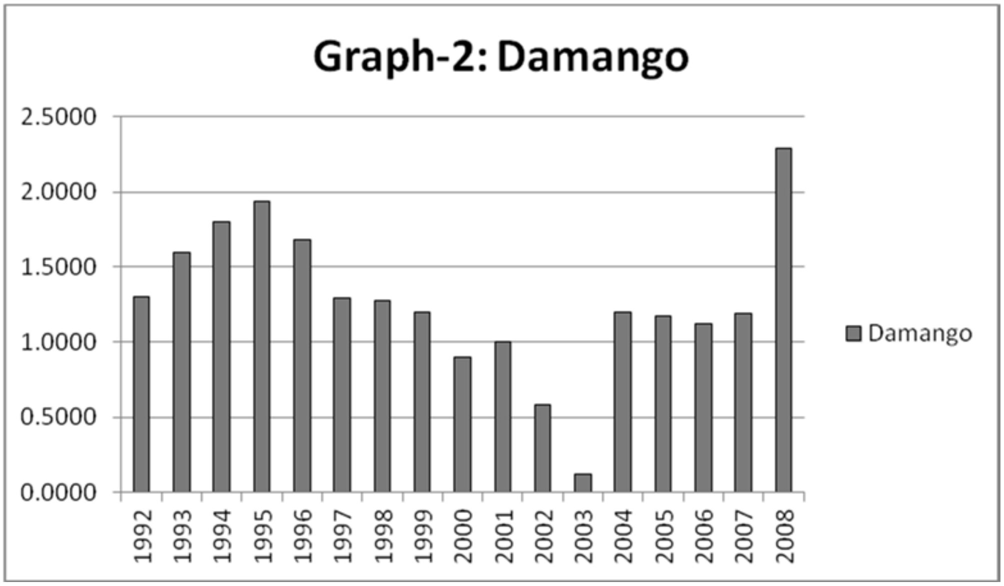
where $L_t$ the smoothed value for year t becomes the forecasted value for year t+1.

| Table-2: Maize in Bole | | | | | |
|---|---|---|---|---|---|
| Model | DF | Error Variance (MSE) | AIC | SBC | R-Square |
| AR(4) | 12 | 0.0787606 | -39.12403 | -34.95797 | 0.51 |
| Simple Exponential Smoothing | 15 | 0.1384718 | -30.66604 | -29.89345 | -0.1 |
| Double (Brown) Exponential Smoothing | 14 | 0.1278583 | -29.88738 | -29.17933 | -0.3 |
| Linear (Holt) Exponential Smoothing | 13 | 0.1848473 | -23.46989 | -22.05379 | -0.36 |
| Damped-Trend Linear Exponential Smoothing | 13 | 0.1485906 | -27.82719 | -25.50943 | -0.09 |

**Double Exponential Smoothing (Brown):**

A double smoothing technique is used when a series has a trend component. With this technique, each observation in a series is assumed to be consisted of two components, level or smoothing component and trend component. This controls any trend or nonstationary component that may exist in the data series.

$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$ and $T_t = \alpha(L_t - L_{t-1}) + (1-\alpha)T_{t-1}$




**Double Exponential Smoothing (Holt):**

Holt smoothing technique is different from Brown's technique in a sense that it uses different parameter value for estimating the trend component.

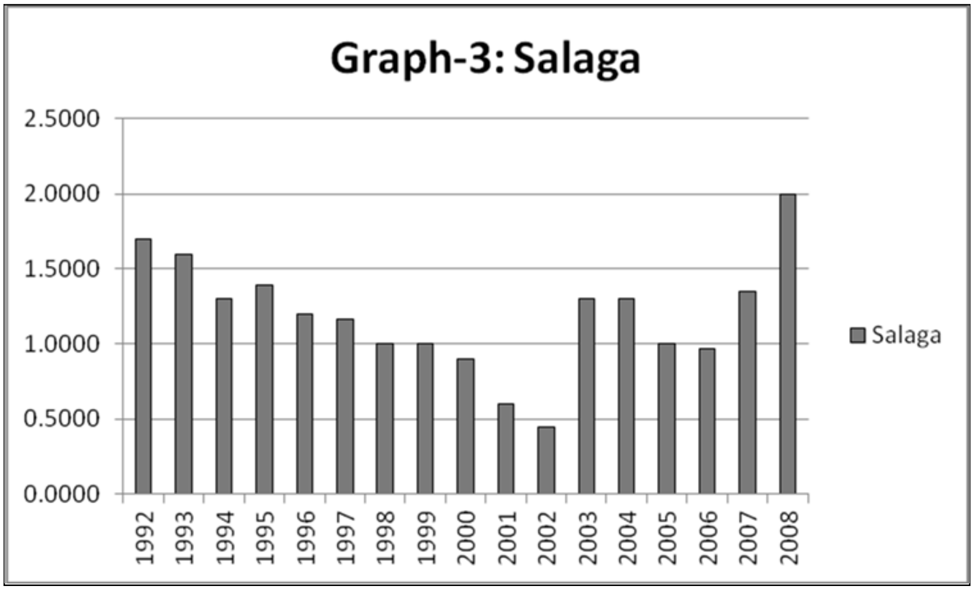
$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$

| Table-3: Maize in Damango | | | | | |
|---|---|---|---|---|---|
| Model | DF | Error Variance (MSE) | AIC | SBC | R-Square |
| AR(5) | 11 | 0.1509526 | -27.54382 | -22.54454 | 0.556 |
| Simple Exponential Smoothing | 15 | 0.2150581 | -23.62217 | -22.84958 | 0.212 |
| Double (Brown) Exponential Smoothing | 14 | 0.250109 | -19.82277 | -19.11472 | 0.112 |
| Linear (Holt) Exponential Smoothing | 13 | 0.2428938 | -19.37348 | -17.95738 | 0.188 |
| Damped-Trend Linear Exponential Smoothing | 13 | 0.2481439 | -19.62217 | -17.3044 | 0.212 |

**Damped-Trend Linear Exponential Smoothing:**

This smoothing technique is a variation of Holt smoothing technique that introduces a third parameter value to dampen the trend magnitude to align with a subdued trend data series. This works better with a data series that has weaker trend component.

$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + \varphi T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)\varphi T_{t-1}$



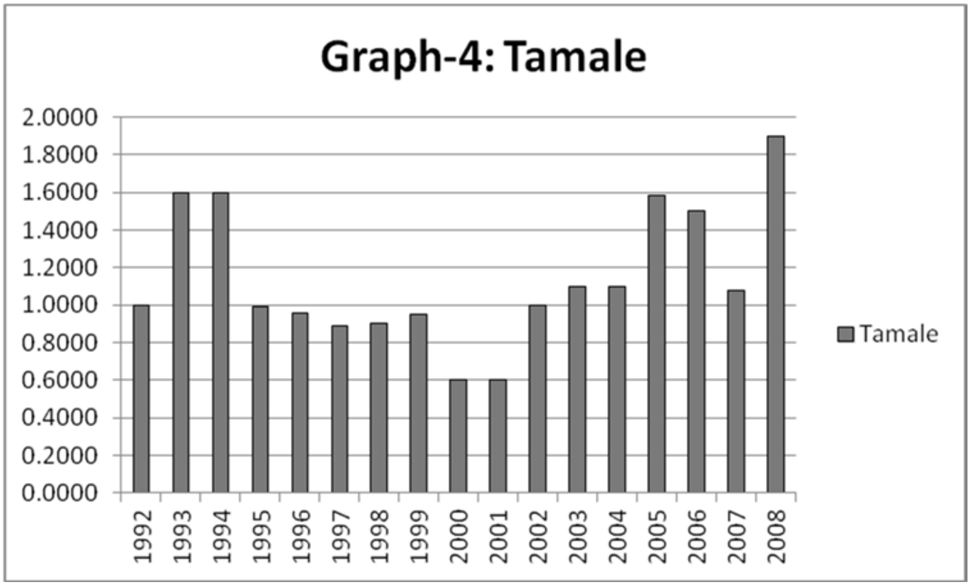

**Autoregressive Model – AR (P):**
Autoregressive (AR) model is a special case of ARMA model of Box-Jenkins (Box and Jenkins, 1976) approach with a stationary data series.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + .... + \phi_p y_{t-p} + \nu_t$$

| Table-4: Maize in Salaga | | | | | |
|---|---|---|---|---|---|
| Model | DF | Error Variance (MSE) | AIC | SBC | R-Square |
| AR(6) | 10 | 0.0814776 | -37.64694 | -31.81445 | 0.624 |
| Simple Exponential Smoothing | 15 | 0.112847 | -33.94017 | -33.16758 | 0.178 |
| Double (Brown) Exponential Smoothing | 14 | 0.1339997 | -29.18366 | -28.47561 | -0.04 |
| Linear (Holt) Exponential Smoothing | 13 | 0.1392374 | -27.72014 | -26.30404 | 0.002 |
| Damped-Trend Linear Exponential Smoothing | 13 | 0.1302081 | -29.94017 | -27.6224 | 0.178 |

To identify the order of the autoregressive model, we have evaluated the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the yield series using SAS procedure PROC ARIMA (see, SAS/ETS User's Guide, 1993). This allowed the observance of the degree of autocorrelation and the identification of the order of the model that sufficiently described the autocorrelation. After evaluating the ACF and PACF, the models are identified as fourth order to sixth order autoregressive models for various districts and a sixth order model is expressed as:

$\left(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 + \phi_4 B^4 + \phi_5 B^5 + \phi_6 B^6\right) \nu_t = y_t$ , (see, Box, Jenkins, & Reinsel, 1994) .




Maximum likelihood estimation method is used instead of nonlinear least squares to estimate the parameters of the models. Maximum likelihood estimation is preferable over nonlinear least squares, because maximum likelihood estimation accounts for the determinant of the variance-covariance matrix in its objective function (likelihood function). Further discussion on different estimation methods and the likelihood functions can be found in Choudhury, Hubata, & St. Louis, 1999 and also see SAS/ETS User's Guide, 1993 for the expression of the likelihood functions.

**Table-5: Maize in Tamale**

| Model | DF | Error Variance (MSE) | AIC | SBC | R-Square |
|---|---|---|---|---|---|
| AR(6) | 10 | 0.0562229 | -43.95401 | -38.12152 | 0.582 |
| Simple Exponential Smoothing | 15 | 0.1277286 | -31.95818 | -31.18559 | 0.084 |
| Double (Brown) Exponential Smoothing | 14 | 0.1477733 | -27.71603 | -27.00798 | -0.4 |
| Linear (Holt) Exponential Smoothing | 13 | 0.1255981 | -29.26653 | -27.85043 | -0.16 |
| Damped-Trend Linear Exponential Smoothing | 13 | 0.1456325 | -28.14894 | -25.83117 | 0.083 |

We have used the following model selection criterion:

Akaike Information Criterion: $AIC = n \ln \frac{SSE}{n} + 2k$ ,

Schwartz's Bayesian Criterion: $SBC = n \ln \frac{SSE}{n} + k \ln(n)$ ,

and R-Square = $1 - \frac{SSE}{SST}$ . Note that in this construct of $R^2$ the value of $R^2$ can be negative when the fitted model's performance is very poor. This means that the total squared deviation of predicted yield from actual yield is larger than the total squared deviation of average yield from actual yield.

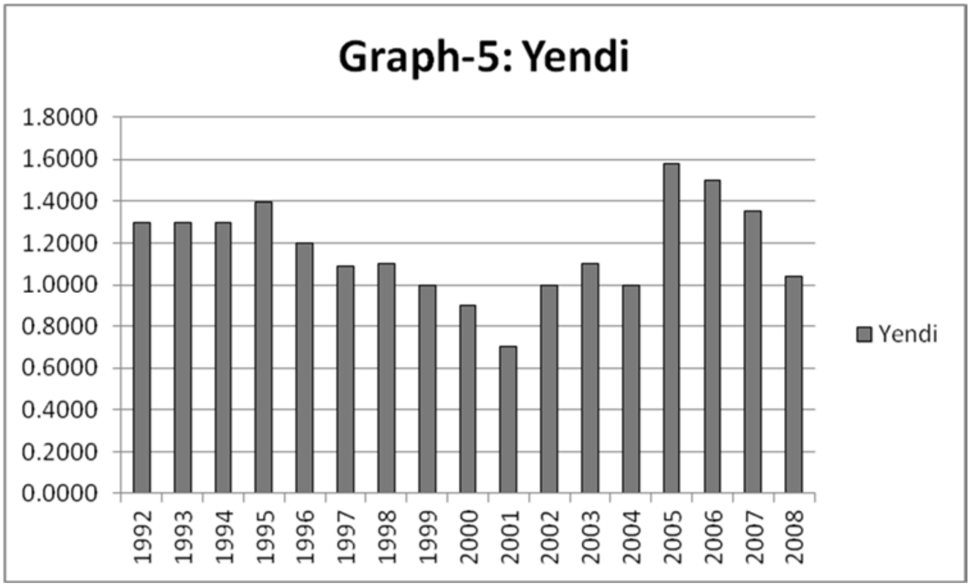
## EMPIRICAL RESULTS

Larger standard deviation 0.5060 of Maize yield in "Damango" district with the highest "maximum" yield of 2.2898 and lowest "minimum" yield of 0.1200 (see Table-1) does indicate much fluctuation in the Maize yields among the districts and thus introduces a challenge in model building strategy. Average "maximum" yield is about 2.0 Mt/Ht among these five districts, whereas the "minimum" yield varies quite a bit with a range of 0.9134 Mt/Ht to 0.1200 Mt/Ht. There appears to be a declining trend in Maize yield till 2001/2002, followed by an increasing trend in yield for a period of six/seven years (see, Graphs 1-5). This may be one of the reason why trend adjusted forecasting (or smoothing) technique is performing so poorly and thus producing negative $R^2$, which is essentially zero. Since, theoretically coefficient of determination ranges from zero to one and cannot be negative. Thus, it appears that there are possibly two opposite crop yield trends which create a cycle that split up around the year 2001/2002. It is possible that this may be due to weather cycle occurrence or management intervention or some other unobservable phenomena of similar nature. This cycle may be country specific and may also be region specific and therefore, needs to be explored further in the future research.

In a similar context, there are also visible differences in declining trend segment ending in a different time period for different districts and thus exhibiting differences in external factors' influence on the crop yield differently. This suggests that due to some unobservable factor(s) crop yield may differ in different time periods for different districts. Thus, the idea of this exploratory analysis is to obtain a best fit forecasting model of crop yield such that the association effect of unobserved external factors with crop yield is best reflected through models' performance criterion. The following results address our research studies of building the forecasting model of crop yields for different districts.

Among all different time series models estimated, ARMA models performed best with higher coefficient of determinations for all five districts considered in this paper. AR (6) model fitted for district "Salaga" has the highest $R^2$=0.624 that accounts for 62.40% variation (see, Table-3) in the Maize yield. The model that produced the lowest $R^2$, among these five districts is "Yendi". District "Yendi" fitted a model that accounts for 48.70 % variation (see, Table-5) in the crop yield. In addition, ARMA models also performed best when considering other performance criterion, such as, MSE or Akaike Information Criterion (AIC). Therefore, our research results show that

ARMA model provides better estimate of crop yield using historical data at a district level compared to other models considered in this paper.




It appears that in addition to the plant characteristics, external factors may also affect the crop yield differently given that which time period they are planted. Specifically, we observe that there exists a crop yield cycle in most of our data sets, which starts with the downtrend that ended around 2001/2002 and then an uptrend for next several years that creates a crop yield cycle. In general, any type of time series data has a cyclical component whether it is visible or subdued. A number of possible explanations can be explored for this time dependent yield cycle. However, considering that most of the time series has some serial correlation properties inherent in them direct comparison may be complicated and difficult to separate.

## CONCLUSION

This paper makes a number of significant contributions to the literature. It provides additional evidence of crop yield cycle component of a time series in most of the districts. In addition, it also suggests evidence of unobserved external factors' effect on crop yield that creates the crop yield cycle. However, any associations of crop yield that may exist with the unobserved external factors' are not explored in this study. These results while important are not unexpected given the dynamic changes that come from external factors, such as, weather (rainfall, temperatures, etc.), land management (that include re-division of districts), pests and diseases. The unexpected finding is the initial continuous decline of crop yield that went on for several years in most of the districts without any management intervention.

| Table-6: Maize in Yendi | | | | | |
|---|---|---|---|---|---|
| Model | DF | Error Variance (MSE) | AIC | SBC | R-Square |
| AR(4) | 12 | 0.0345774 | -53.11867 | -48.9526 | 0.487 |
| Simple Exponential Smoothing | 15 | 0.0455613 | -48.45175 | -47.67916 | 0.152 |
| Double (Brown) Exponential Smoothing | 14 | 0.0624674 | -40.63155 | -39.9235 | -0.12 |
| Linear (Holt) Exponential Smoothing | 13 | 0.0525918 | -42.32445 | -40.90835 | 0.035 |
| Damped-Trend Linear Exponential Smoothing | 13 | 0.0525708 | -44.45175 | -42.13399 | 0.152 |

Considering crop yield trend and crop yield cycle separately from other factors (external or internal) and purely from the historical point of view, illustrates how policy makers can benefit from using the results of this study. It is also well known that most of the time series has an inherent cycle component that may or may not be significant. However, understanding the mechanism of up-cycle and down-cycle with crop yield will provide an advantageous position to the policy makers to prepare an appropriate policy design for yield management.

Therefore, a successful operation of an Area-Yield Index insurance policy to work the crop grown in the Insured Unit (District) needs to be relatively homogeneous in terms of the varieties grown by farmers, sowing dates, crop husbandry practices and input utilization and finally the average yields of the crop obtained by the farmers in the defined unit. To date no work has been conducted on individual crop-cut yields to assess the degree of variability in crop yields obtained by farmers in the same district. Additional research development is needed, particularly with regard to the linkage between these factors and crop yield dynamics. To determine the length of downtrend or uptrend and therefore the total cycle of crop yield, future research could examine these phenomena over different periods of time.

The ARMA models, which are univariate models that use primarily autocorrelations from its past, proved more robust time-series models than the smoothing technique models for predicting crop yield in this study. This is consistent with the findings of Pal et al., 2007. They found that an ARMA model for forecasting Milk production resulted in much better estimates than the other time series approaches considered. The ARMA methodology avoided the problem of highly variable crop yields within the district over time, which has led to low performance on the prediction of crop yield by averaging or smoothing models. This predictive power of ARMA models does not depend on whether and how long the crop yield cycle persists. These findings are consistent with the objective that an efficient prediction modeling process is very much interrelated with the yield data itself. Therefore, the results of this study indicate that the performance of a prediction model is dependent on the dynamic nature of the crop yield data and this may be region specific. Thus, the districts with wider yield spreads may like to use different time series models than those districts with more homogenized yield.

## REFERENCES


Armstrong, J.S. (Ed), 2001. *Principles of forecasting, a handbook for researchers and practitioners*. Kluwer Acad. Press, Boston, Doordrecht, London.

Badmus, M.A., and Ariyo, O.S. (2011) 'Forecasting Cultivated Areas and Production of Maize in Nigeria Using ARIMA model', *Asian Journal of Agricultural Sciences* 3(3), 171- 176.

Barnett, B. J., Barrett, C. B. and Skees, J. R. (2008) 'Poverty traps and index-based risk transfer products', *World Development* 36(10), 1766-1785.

Bogetic, Y., Bussolo, M., Ye, X., Medvedev, D., Wodon, Q., and Boakye, D. (2007) 'Ghana's growth story: How to accelerate growth and achieve MDGs', *Background paper for Ghana CEM.*

Boken, V.K., 2000. Forecasting spring wheat yield using time series analysis: a case study for the Canadian Prairies, *Agronomy Journal*, 92(6),1047-1053.

Box,G. E. P.and Jenkins, G.M. (1976) *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day, Inc.

Box, G.E.P., Jenkins G.M., and Reinsel G.C. (1994) *Time Series Analysis: Forecasting and Control*, Englewood Cliffs, NJ: Prentice-Hall.

Carbone, G. J., Kiechle, W., Locke, C., Mearns, L. O., McDaniel, L. and Downton, M. W. (2003) 'Response of soybean and sorghum to varying spatial scales of climate change scenarios in the southeastern United States', *Climatic Change* 60(1 - 2), 73–98.

Carter, M. R., Galarza, F., and Boucher, S. (2007) 'Underwriting area-based yield insurance to crowd-in credit supply and demand', *Savings and Development* 31(3), 335-362.

Choudhury, A., Hubata, R. and St. Louis, R. (1999) 'Understanding Time-Series Regression Estimators', *The American Statistician*, 53(4), 342-348.

Deng, X., Barnett, B. J., Hoogenboom, G., Yu, Y., and Garcia, A. (2006) Evaluating the efficiency of crop index insurance products, Paper presented at the *Southern Agricultural Economics Association annual meetings*, Orlanda, Florida.

Easterling, W. E., Weiss, A., Hays, C. J. and Mearns, L. O. (1998) 'Spatial scales of climate information for simulating wheat and maize productivity: the case of the US Great Plains', *Agricultural and Forest Meteorology* 90(1-2), 51–63.

Etwire, P.M., Al-Hassan, R.M., Kuwornu, J.K.M. and Osei-Owusu, Y. (2013) 'Application of Livelihood Vulnerability Index in Assessing Vulnerability to Climate Change and Variability in Northern Ghana', *Journal of Environment and Earth Science* 3(2),157-170.

Fodor, N. and Kovacs, G. J. (2005) 'Sensitivity of crop models to the inaccuracy of meteorological observations', *Physics and Chemistry of the Earth*, Parts A/B/C 30(1-3), 53–57.

Gerretsadikan, A and Sharma, M.K. (2011) 'Modeling and forecasting of rainfall data of mekele for Tigray region (Ethiopia)', *Statistics and Applications* 9(1-2), 31-53.

Hansen, J.W., Challinor, A., Ines, A., Wheeler T. and Moron, V. (2006) 'Translating climate forecasts into agricultural terms: advances and challenges', *Climate research* 33,27-41.

Hansen, J. W. and Jones, J. W. (2000) 'Scaling-up crop models for climate variability applications', *Agricultural Systems* 65(1), 43–72.

Hoogenboom, G. (2000) 'Contribution of agrometeorology to the simulation of crop production and its applications, *Agricultural and Forest Meteorology,* 103(1-2), 137–157.

Launay, M. and Guérif, M. (2003) 'Ability for a model to predict crop production variability at the regional scale: an evaluation for sugar beet', *Agronomie* 23, 135–146.

Lawless, C. and M.A. Semenov, 2005. Assessing lead-time for predicting wheat growth using a crop simulation model. *Agricultural and forest meteorology* 135(1-4),302-313.

Makridakis, S., Wheelwright S.C. and Hyndman, R.J. (1998) *Forecasting, Methods and Applications*, New York: John Wiley & sons, Inc.

Mearns, L. O., Mavromatis, T., Tsvetsinskaya, E., Hays, C. and Easterling, W. (1999) 'Comparative responses of EPIC and CERES crop models to high and low spatial resolution climate change scenarios', *Journal of Geophysical Research-Atmospheres* 104(D6), 6623–6646.

Mearns, L. O., Easterling, W., Hays, C. and Marx, D. (2001) 'Comparison of agricultural impacts of climate change calculated from high and low resolution climate change scenarios: Part I. the uncertainty due to spatial scale', *Climatic Change* 51(2), 131–172.

Najeeb, I., Khuda, B., Asif, M., and Abid, S.A. (2005) 'Use of ARIMA Model for Forecasting Wheat Area and Production in Pakistan', *Journal of Agricultural and Social Sciences* 1(2), 120- 122.

Nonhebel, S. (1994) 'The effects of use of average instead of daily weather data in crop growth simulation models', *Agricultural Systems* 44(4), 377–396.

Pachepsky, Y. and Acock, B. (1998) 'Stochastic imaging of soil parameters to assess variability and uncertainty of crop yield estimates', *Geoderma* 85(2-3), 213–229.

Pal, S., Ramasubramanian, V. and Mehta, S.C. (2007) 'Statistical models for forecasting milk production in India', *Journal of Indian Society of Agricultural Statistics* 61(2), 80–83.

Russell, G. and Gardingen, P.R.v. (1997) 'Problems with using models to predict regional crop production', in Gardingen, P.R. v., Foody, M. and Curran P. (eds), Scaling up: from cell to landscape, Cambridge University Press, Cambridge, pp. 273–294.

Semenov, M. A. and Porter, J. R. (1995) 'Non-linearity in climate change impact assessments', *Journal of Biogeography* 22(4-5), 597–600.

Skees,J.R., Black, J.R. and Barnett, B.J. (1997) 'Designing and rating an area yield crop insurance contract', *American Journal of Agricultural Economics* 79, 430-438.

Skees, J. R. (2008) 'Innovations in index insurance for the poor in lower income countries', *Agricultural and Resource Economics Review* 37(1),1-15.

Soltani, A., Meinke, H. and de Voil, P. (2004) 'Assessing linear interpolation to generate daily radiation and temperature data for use in crop simulations', *European Journal of Agronomy* 1(2), 133–148.

Stutley, C. (2010) Innovative Insurance Products for the Adaptation to Climate Change Project Ghana, Innovative Insurance Products for the Adaptation to Climate Change Project Ghana.

Suleman, N. and Sarpong, S. (2012) 'Forecasting Milled Rice Production in Ghana Using Box-Jenkins Approach', *International Journal of Agricultural Management and Development*, 2(2), 79-84.