Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 67(3) 2013 371-379

Markov Chain based Crop Forecast Modeling Software*

Ram Manohar Patel, R.C. Goyal, V. Ramasubramanian and Sudeep Marwaha
Indian Agricultural Statistics Research Institute, New Delhi

Received 05 July 2012; Revised 27 November 2012; Accepted 30 November 2012

SUMMARY

Crop yield forecasts are quite useful in formulation of policies regarding stock and distribution of agricultural produce to different areas of any country. One among the various statistical approaches in vogue includes models based on Markov chains for providing objective forecasts of crop yields well in advance before harvest for taking timely decisions. A situation which takes the form of a chain of stages with a limited number of possible states (plant condition classes) within each stage is called a Markov chain, if there is a case of simple dependence that any state of a particular stage depends directly on any of the states of the preceding stage. However, for dealing with the key features of Markov chains like estimation of transition probability matrices, predicted yield distributions etc. to get final forecasts, the computational efforts are tedious. One has to either take recourse to writing programs or use statistical packages. Many standard statistical software packages cater to analyze data and obtain forecasts either by using the regression or time series approaches. No single software has tailor-made and customized module to get forecast using the stochastic approach viz. Markov chain modeling. Hence a userfriendly software has been developed based on Markov chain model. It can be used in any platform having Java Virtual Machine (JVM), Java being a platform independent language, programming has been done in Core Java (as back end) and Java Swing (as front end). For testing the software, two years data on biometrical characters and crop yield collected by Indian Agricultural Statistics Research Institute (IASRI), New Delhi under the pilot study on pre-harvest forecasting of sugarcane in Meerut district were utilized. The salient features of the software are highlighted briefly. The software builds first order finite Markov chain (hereinafter referred to as FOMC) model. The software allows up to twenty stages (excluding the harvest stage for first year) depending upon the crop, in the Markov chain model, up to 16 states within each stage and up to four variables can be considered within each stage. Minimum ten records are required for performing analysis through this software. The software is having full description of Markov chain based forecasting model used under the help menu. It also has online help at each screen, which enables the user to decide what to do next and about the details of data files. The software shows output in terms of expected crop yield (forecast) at various stages.

Keywords: First order Markov Chain, Java swing, Predicted yield distribution, Transition probability matrix.

1. INTRODUCTION

Crop forecasts are pre-harvest estimates about produce of a crop provided in advance while the crop is still standing in the field. Crop forecasts indicate in advance the availability, quantity etc. of the raw material, which forms the basis of planning of production processes and trade operations. These forecasts are extremely useful in framing policy regarding stock, distribution and supply of agricultural produce to different areas in the country. However, the forecast of production should be given well before the harvest for taking timely decisions. Various approaches have been used to forecast crop yields at intermediate times in the growing season. Prominent among these methods of forecasting are based on models that utilize data on crop biometrical characters, weather parameters, farmers eye estimates, agrometeorological conditions,

Corresponding author: V. Ramasubramanian
E-mail address: ram.vaidhyanathan@gmail.com
*Work carried out as part of M.Sc. Dissertation of Post Graduate School, IARI, New Delhi – 110012

remotely sensed crop reflectance observations etc., utilized either separately or in an integrated approach.

Most of the earlier studies carried out in this respect utilize the traditionally used regression models. All the models developed require fitting of regression using least square technique for estimating the parameters. The optimality properties of these estimates are described in an ideal setting that is not always realized in practice. An alternative approach to crop yield forecasting is the probability model based on Markov chain theory. This method overcomes some of the drawbacks of the regression model. Markov chain method has the advantage of providing nonparametric interval estimates and is robust against outliers and extreme values. The only assumption under a FOMC *i.e.* first order Markov chain (or simply a Markov chain) set up is that the past (crop) conditions are statistically uninformative for predicting the future (yield forecasts), after the present (crop) conditions are known. A situation that takes the form of a chain of stages with a limited number of possible states (plant condition classes) within each stage is called a Markov chain, if there is a case of simple dependence that any state of a particular stage depends directly on any of the states of the preceding stage.

Markov chain modeling approach is increasingly growing and recognized in recent times. However, for dealing with the key features of Markov chains like construction of stages and states within stages, estimation of transition probability matrices, multiplication of these matrices etc. to get final forecasts, the computational efforts are tedious. One has to take recourse to writing programs or use statistical packages. Statistical software packages have indeed revolutionized the way we analyze data. Many standard statistical software packages (SAS, SPSS etc.) cater to analyze data and obtain forecasts either by using regression or time series approaches. The problem with SAS is that, though it is statistical software package, it requires programming knowledge for writing the program syntax codes to analyze data for forecasting purposes. In the context of Markov chain model based forecasting, earlier workers of course have utilized this software, but have to write many programs in bits and pieces. Thus SAS programming is not a good way for the end users in this sense to get the forecast results. The statistical package to arrive at forecasts, SPSS, though is a menu driven software but it require good basic skills to operate this software.

Attempts were made to search for possible utilities written for Markov chain modeling in other software packages like FORECAST PRO (2012), EVIEWS (2011) etc. but no single software has tailor-made and customized module to get forecast using Markov chain modeling. Hence user-friendly forecasting software has been developed which will be based on the Markov chain model under Windows operating system environment.

2. REVIEW OF LITERATURE

Matis *et al.* (1985) proposed a statistical methodology for forecasting crop yields at intermediate times of the growing season of the crop using Markov chain theory. Jain and Agrawal (1992) developed Markov chain models for forecasting sugarcane yields and found that yield forecast can be provided at about 2-3 months before harvest. Singh and Ibrahim (1996) have attempted the use of remotely sensed spectral data in Markov chain model for obtaining pre-harvest wheat yield forecasts. Jain and Ramasubramanian (1998), Ramasubramanian and Jain (1999) and Ramasubramanian *et al.* (2004) have developed Markov chain models for forecasting sugarcane yields by refining the procedure. Sonis and Dendrinos (2009) have employed Markov chain models in the field of discrete socio-spatial dynamics by focusing on the redistribution of different statistical populations relatively distributed between different locations by utilizing the transfer from any time period to the subsequent time period for constructing a probabilistic Markov chain. Ramasubramanian *et al.* (2010) have developed yield forecast models for sugarcane crop using higher order (multiple) Markov chains by considering different possible combinations of various aspects viz., orders of Markov chain used, number of biometrical characters used and percentile definitions of plant condition states of biometrical characters within stages. They have found that when the order of Markov chain increases and/or the definition of states become finer, the mean yield forecasts approach the actual yield rapidly justifying the development of multiple Markov chain models with finer definitions of states of plant conditions. They have also found that by using transformation of data such as principal components or growth indices the order of the chain could be increased without having problems in rapid increase in the number of states so that the chain becomes more manageable and efficient. It is emphasized here that higher order Markov chain models can also be fitted using the developed software as every such Markov chain of higher order can be developed as if it were a first order Markov chain by resorting to appropriate data arrangement itself. Buchman (2012) has attempted predictive land use modeling based on higher order Markov chains. They have used spatio-temporal data to determine the cropping patterns of the Big Creek watershed of USA. They have also implemented a post classification per polygon technique to determine the Markov transition matrices to compare first, second and third order Markov models. In this study a significant amount of change occurring in the Big Creek watershed has been observed with crop rotation being a large portion of these changes. All the aforesaid workers have developed the model by writing the programs in computer languages viz. FORTRAN and C or utilized some computational aspects of packages like SAS, SPSS etc. by writing package specific syntax codes. However, such piecemeal programming approach is obviously not user-friendly. So the need for a user-friendly, cost effective and interactive package need hardly be emphasized.

3. MATERIALS AND METHODS

The Markov chain crop forecast modeling software has been developed on the Windows platform. Nevertheless it can be used in any platform having Java Virtual Machine (JVM), Java being a platform independent language, programming has been done in Core Java (as back end) and Java Swing (as front end). Core Java and Swing provide a complete set of tools for the creation of powerful features including some advanced features like JTable, JTree, JEditorPane etc. for all operating systems (having JVM).

The software has been developed using Java Swing language (Horstmann and Cornell 2004. Loy *et al.* 2003 and Pantham 1999). Swing is a set of customizable graphical components whose look and feel (L&F) features can be dictated at run time. Swing is the next generation GUI toolkit that Sun Microsystems' Java group (Javasoft) created to enable enterprise development in Java. Its first official release was in March 1998 which revolutionalised Java user interface environment. It was evolved as an improvement upon Abstract Window Toolkit (AWT) and Java Foundation Classes (JFC). It has a flexible and portable user interface. The special features of Java Swing over and above Java are:

- **PLAF feature – Pluggable Look And Feel:** Swing is in complete control of its components. It is in control of the way components look on the screen and gives you more control over how your applications work and represent the same on the screen.
- **Lightweight components:** Swing components are not dependent on native peers to render themselves. It solves most of the portability problems since it creates its own components and does not rely on the runtime platform's native environment.
- **Model and View separation:** Swing makes a very clear distinction between the data a component displays and the actual view. It is flexible to adapt components to display new kinds of data that was not originally anticipated.

The software has been designed as GUI (Graphical User Interface), in Core Java and Swing mainly using modular programming and Object Oriented methodology (*i.e.* using class modules). The user

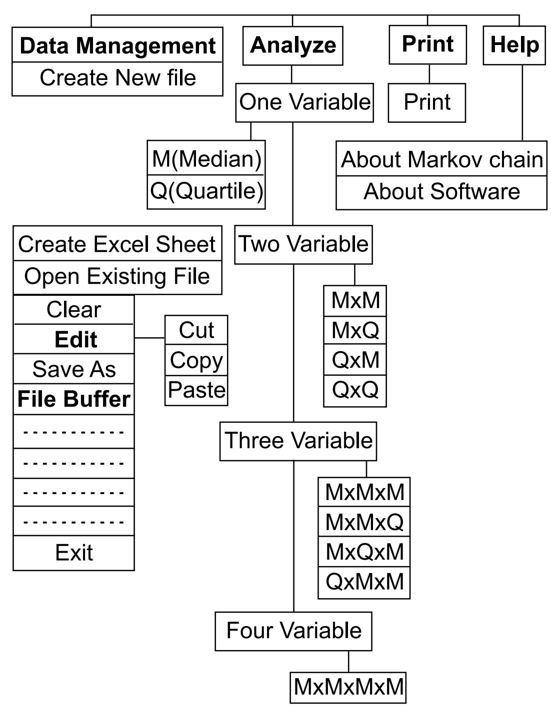

Fig. 1. User Interface Design in Markov Chain software

interface of Markov chain crop forecast modeling software is designed using Java Swing with standard menus and dialog boxes as in all basic Windows applications so that the user will feel comfortable in using the software. The various Menus in the user interface design are Data Management, Analyze, Print and Help as shown in Fig 1. Various algorithms were written for calculation of percentiles, formation of states within stages, generation of transition probability matrices, predicted yield distributions and mean yield forecasts.

For testing the software, two years data on biometrical characters and yield collected by IASRI, New Delhi under the pilot study on pre-harvest forecasting of sugarcane in Meerut district were utilized. The first year data has been used for model development and forecasts obtained for second year. In Markov chain modeling set up, growth process of the crop can be divided into s stages based on calendar dates. For the test data that has been used in the software the number of stages were $s = 6$. Let these stages be denoted by $S1$, $S2$, $S3$, $S4$ and $S5$ with the final (harvest) stage denoted by $S6$. Thus the basic input for the program is the first and second year data upon the variables say X_{ij} where i represent stages, j represent different biometrical characters at various stages and yield data at harvest stage of first year. Let m_i denote the number of states defined on the basis of percentiles of observations of the selected biometrical characters within stage i ($i = 1, 2, \ldots, s$). The different summary statistics calculated upon these variables are maximum and minimum values and percentiles such as median, quartiles, and deciles. The percentiles (of each data column *i.e.* each variable) are utilised for definition of Markov chain states. For the sake of getting insight, let us consider only two variables (for simplicity) viz. $X1$ and $X2$ upon which data points at various stages $S1$ through $S5$ for first year are available apart from information upon yield Y at harvest stage $S6$. The different states within the stage, for stage, say $S1$ can be formed by the combination of following conditions (based on M × M approach *i.e.* Median of $X1$ and Median of $X2$). It is noted here that we could have used other percentiles like quartiles as well to form the states.

At Stage $S1$, $X11$ is classified on the basis of medians ($M11$, say) and thus we get two classes viz. $X11 <= M11$ and $X11 > M11$. Similarly, $X12$ is classified as $X12 <= M21$ and $X12 > M21$. In all, we get four states within the stage $S1$:

(i) $X11 <= M11, X12 <= M21$

(ii) $X11 <= M11, X12 > M21$

(iii) $X11 > M11, X12 <= M21$

(iv) $X11 > M11, X12 > M21$

Similarly, within the stage $S2$, such states can be defined. Likewise for other stages in the FOMC model these states can be defined accordingly. The first year data in different stages have been classified into different plant condition classes (states) for the various selected variables within stages of the Markov Chain model.

Let $\mathbf{A}_{i,i+1} = ((p_{kl}))$ ($i = 1, 2, \ldots, s - 1$; $k = 1, 2, \ldots, m_i$; $l = 1, 2, \ldots, m_{i+1}$) denote the ($m_i \times m_{i+1}$) transition probability matrices (t. p. m.'s) which give the transition probabilities p_{kl} of a group of plants moving from any possible state k of stage i to any possible state l of stage (i+1), each row summing to one. The observed frequencies of the plant moving from one plant condition class(state) of a stage, say, $S[i]$(i = 1, 2, 3, 4, 5) to different condition classes(states) of the next stage $S[i+1]$(i = 1, 2, 3, 4) can be calculated. For instance, the frequency matrix of transition from $S1$ to $S2$ can be obtained as (with two variables within each stage classified on the basis of medians for forming states)

		Stage $S2$			
	States	(i)	(ii)	(iii)	(iv)
	(i)	$c11$	$c12$	$c13$	$c14$
Stage $S1$	(ii)	$c21$	$c22$	$c23$	$c24$
	(iii)	$c31$	$c32$	$c33$	$c34$
	(iv)	$c41$	$c42$	$c43$	$c44$

These frequencies are then divided by their corresponding row sums to compute the elements of the transition probability matrix (t.p.m.). Hence row sum of t.p.m. is unity. This will give rise to t.p.m.s' $\mathbf{A}[i][i+1]$ ($i = S1, S2, S3, S4, S5$). For the illustration under discussion, each matrix will be of order 4 × 4 except for the last t.p.m. whose order is 4 × 10 because in the final stage, only information upon yield Y is available and it has been classified on the basis of deciles. For instance, from the above frequency matrix, the t.p.m. of transition from $S1$ to $S2$ is given by

		Stage $S2$			
	States	(i)	(ii)	(iii)	(iv)
	(i)	$P11$	$P12$	$P13$	$P14$
	(ii)	$P21$	$P22$	$P23$	$P24$
Stage $S1$	(iii)	$P31$	$P32$	$P33$	$P34$
	(iv)	$P41$	$P42$	$P43$	$P44$

Similarly, other t.p.m.s' can be computed.

For a given stage i, the predicted yield distributions (p.y.d.s') have been obtained as the product $\prod_{a=i}^{s-1} \mathbf{A}_{a,a+1}$ which is of order ($m_i \times m_s$). The midpoints of these yield classes can be formed as a summary mean vector $\underline{y}_m$ of order ($m_s \times 1$).

Means of predicted yield distributions for each of the states of a stage for the first year has been calculated by simply multiplying the p.y.d.s' with $\underline{y}_m$. Thus at each stage i, means of p.y.d.s' can be obtained as

$$\underline{y}^{(i)} = \left(\prod_{a=i}^{s-1} \mathbf{A}_{a,a+1} \right) \underline{y}_m$$

of order ($m_i \times 1$). The vector $y^{(i)}$ will contain elements as y_{ij}'s with j ranging from 1 to m_i.

To forecast yield of second year, the second year data can be classified as per the states of a stage in first year. This will result in number of observations of second year, say f_{ij}, falling in various states 1, 2, …, m_i of a particular stage i in first year. These values at the various stages for the different states for the FOMC model can be written as

States	(i)	(ii)	(iii)	(iv)
Stage $S1$	$f11$	$f12$	$f13$	$f14$
Stage $S2$	$f21$	$f22$	$f23$	$f24$
Stage $S3$	$f31$	$f32$	$f33$	$f34$
Stage $S4$	$f41$	$f42$	$f43$	$f44$

Note that the row sum here should be the number of data points in second year at each stage. Weighted mean of means of predicted yield distributions $\underline{y}^{(i)}$ for each of the states of a stage i can be worked out, weights being the number of observations of second year falling in different states/stages of the first year data, which would give mean yield forecast Y_{Fi} at each stage i for the second year. This (weighted) mean yield forecast at stage i can be mathematically written as

$$y_{F_i} = \frac{\sum_{j=1}^{m_i} f_{ij} y_{ij}}{\sum_{j=1}^{m_i} f_{ij}} = \frac{1}{n} \sum f_{ij} y_{ij} \text{ (say), where } n = \sum_{j=1}^{m_i} f_{ij}$$

is the total number of observations available in the second year at each stage i, however, since irrespective of any stage same number of observations are used hence 'n' not suffixed by i.

Some points are to be taken care of while using the software. The number of columns except yield column should be equal in both year data and compatible to the model. Minimum ten records are required for performing the analysis. Number of stages should not be greater than 20 excluding the harvest (yield) stage. The control variables are number of columns, stages, states per stage. The data that are read into the program must be ordered by the stages so that the yield data (last stage for first year) should be in last column. The data can be entered in free format *i.e*. the data value of each variable is separated by comma/a blank space/ a tab and each line should be terminated by Enter/ Semi-colon/new line(\n). The data should have column headers in the first row. Data can be read from text files or notepad files. The software is having full description of Markov chain based forecasting model used in the software under the help menu. It also has online help at each screen, which enables the user to decide what to do next and about the details of data files.

4. SALIENT FEATURES OF THE SOFTWARE

The salient features of the software are highlighted briefly. It is user friendly, Windows based (but can be installed on any platform due to platform independence), and simple to be utilized in crop forecast modeling. Designing with Java Swing gives a lot of flexibility in organizing user interface and also it gives a pluggable look and feel of basic windows applications. It allows up to 20 stages (excluding the harvest (yield) stage for first year) depending upon the crop, in the Markov chain model. It allows up to 16 states within each stage (depending upon the crop considered) in the Markov chain model. It also allows up to four number of variables that can be considered within each stage. The software is having full description of Markov chain based forecasting model used in the software. The software has online help at each screen, which enables the user to decide what to do next and about the details of data files. The software shows output in terms of the expected crop yield (forecasts) at different stages.

5. INSTALLATION AND OPERATIONS WITH SOFTWARE

Install Java Development Kit **(JDK)** in any drive (say C :\), and execute the following commands at the command prompt:

C:\JDK1.4.0bin>javac className.java (for compilation)

C:\JDK\bin\>java className (for executing the program)

where className.java is the name of the java file & className is the name of the class file. For running the Java programs path variable has to be changed in the **Autoexec.bat** file as **Set PATH = <rootdirectory>:\jdk1.4.0\bin** where <rootdirectory> is the directory where JDK is installed.

6. EXECUTION OF THE SOFTWARE

User has to start the computer and come to the default desktop of windows. He has to double click on **My Computer** icon. Insert the disk of software in computer then double click on that disk drive icon. It will show the content of the disk. All the files have been

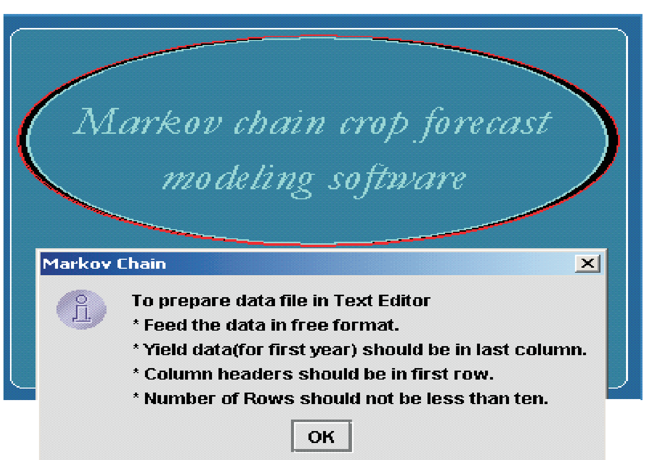

Fig. 2. Splash window cum Creation of data file

categorized under MC-FOMOS folder. The user can copy this folder/files in any directory.

After clicking the MC-FOMOS a splash window will appear suggesting that it is a software for analysis of crop forecast data using Markov chain models as shown in Fig. 2. Next, after clicking the OK button of confirmation dialog box a message dialog box will appear as superimposed in Fig. 2 suggesting about data entry. After clicking OK, the default blank Text Editor screen is displayed. First year data can be entered in this screen or to create a new (Text Editor) file, click on **DataManagement** menu and then go to **Create New File** option. A new data file (in text Editor Format) will be opened and user can enter observations based on the plant biometrical characters (variables) with each column representing one variable.

To create new (Spreadsheet) file, click on **DataManagement** menu and then go to **New Table Sheet** option. An input dialog box asking for the number of first year data columns will popup. After entering this value, a new sheet (in Spreadsheet format) will be opened as shown in Fig. 3 and user can enter

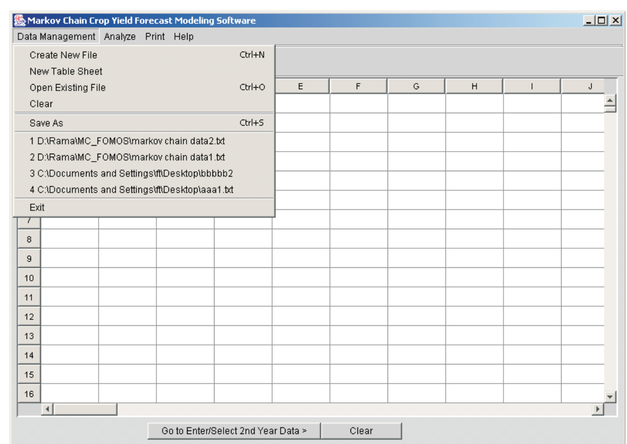

Fig. 3. Opening the data source

observations based on the plant biometrical characters. To open an existing data file, click on **File** menu and then go to **Open Existing File.** Thus the data can be entered in plain text editor or in Worksheet editor or user can import an existing file in worksheet format. After feeding first year data if user clicks on 'Analyze' menu, then different models in the Analyse menu will appear in disabled mode. Hence, after taking first year data, user have to click on 'Go to Enter/Select 2[nd] year data >' button. Before second year data screen appears a confirmation dialog box for 1st year file will appear to check whether first year data has been correctly fed and complete as shown in Fig. 4. If yes, then second year screen will appear for feeding second year data. After feeding second year data, user can select any one option from among "One Variable/ Two Variable/ Three Variable/ Four Variable" and then within that model given below select any one of the available sub-analyses from Analyze menu as shown in Fig. 5. If the model selected is incompatible with the data for first year then an error message will appear. Now, user can click on 'Go to Result' button, a confirmation dialog box for data file and model selection will popup before opening Result screen as shown in Figs. numbered 6 through 11. The software will give output about model specifications and summary statistics (Fig. 6), frequency distribution table of yield (Fig. 7), Transitional Probability Matrices (Fig. 8), Predicted Yield Distributions (Fig. 9), frequency (weights) table of second year based on the first year percentiles (Fig. 10) and the Forecast Yield (Fig. 11). Output/result can be

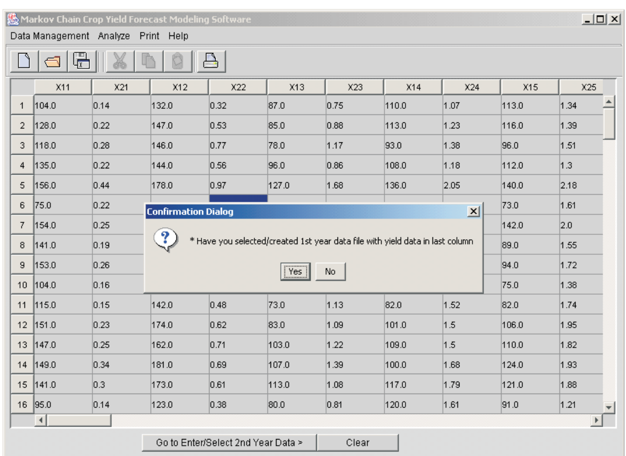

Fig. 4. First year confirmation dialog box for file selection

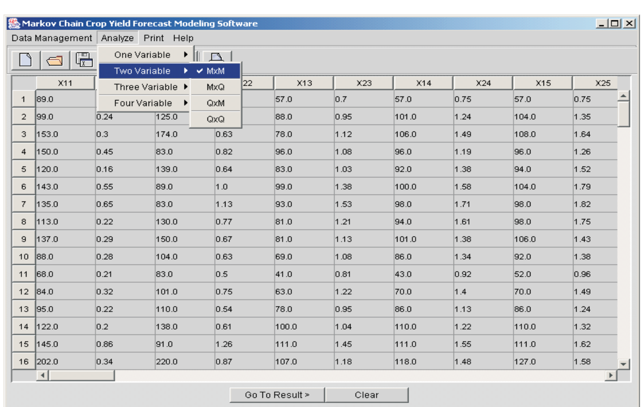

Fig. 5. Selecting model from the Menu

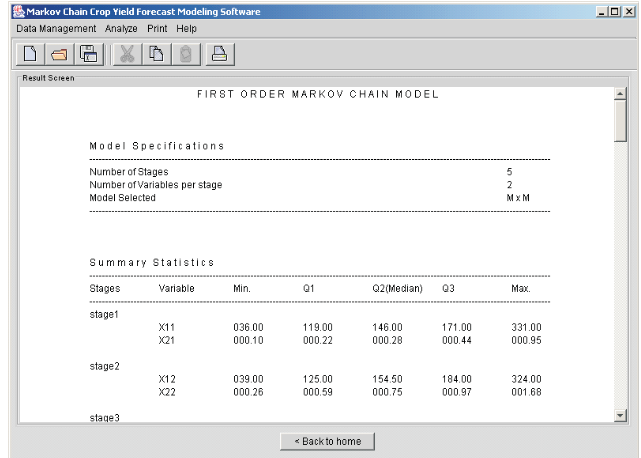

Fig. 6. Output showing model specifications and summary statistics

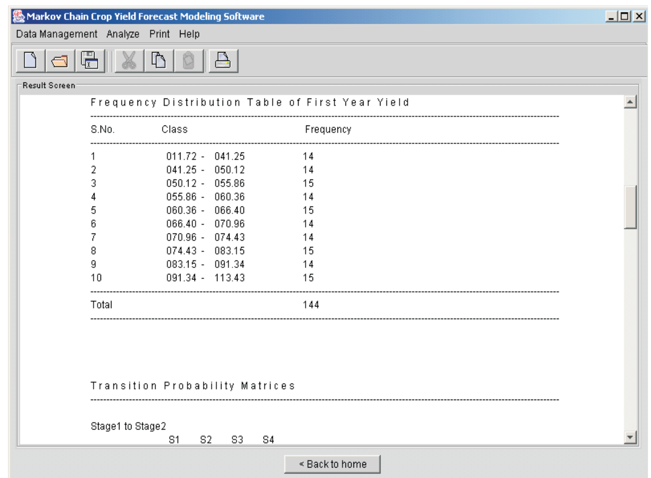

Fig. 7. Output showing frequency distribution table for first year

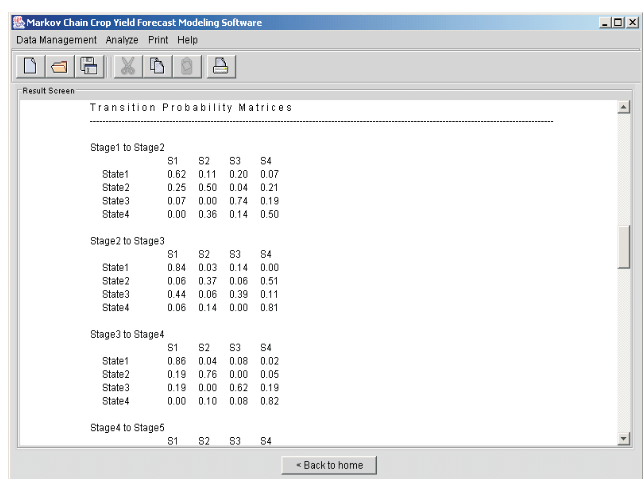

Fig. 8. Output showing transition probability matrices

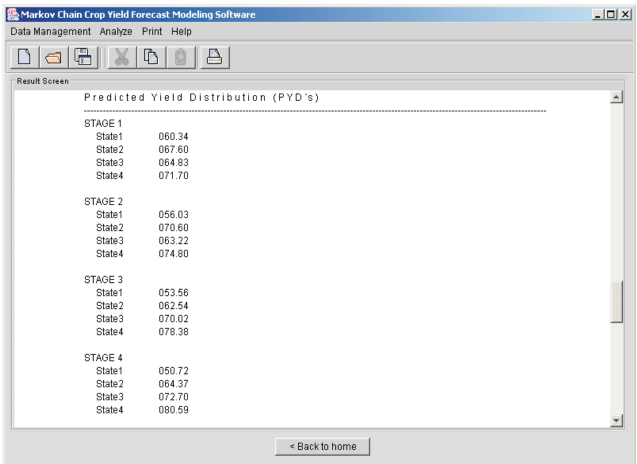

Fig. 9. Output showing predicted yield distributions

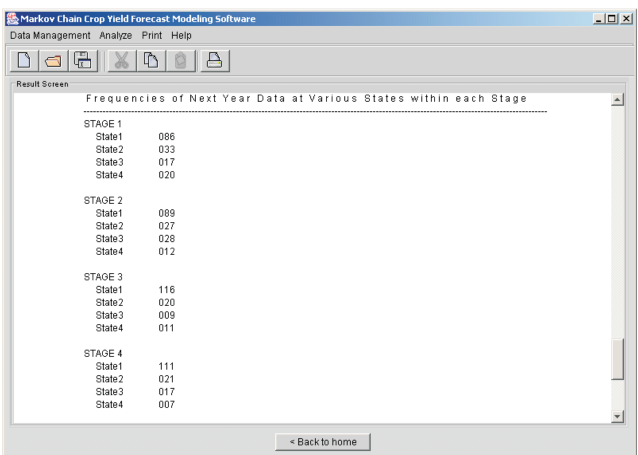

Fig. 10. Output showing frequencies of second year data in each stage

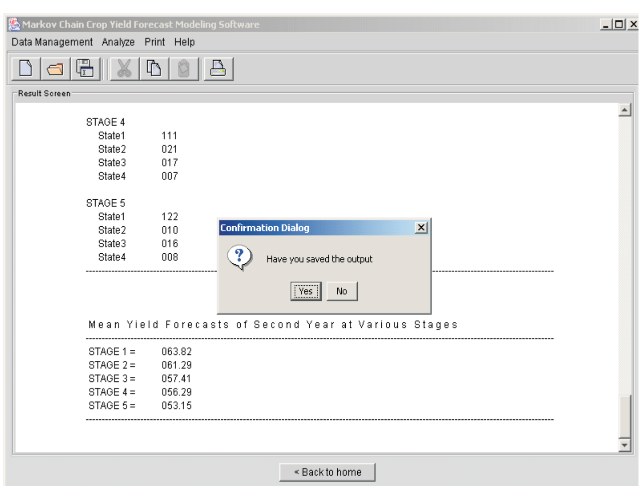

Fig. 11. Output showing mean yield forecasts of second year at various stages

saved and printed. User can go to first year default (Text Editor) screen by clicking on '< Back to home' button in result screen to do analysis based on a different set of data or for the same data set but using a different model.

7. FUTURE SCOPE OF INVESTIGATION

Software can never be complete. The scope for future investigation is summarized here. The Markov chain software accepts the data given by the user as such. Attempts can be made to allow for the software to transform the raw data into principal components or growth indices. Upon these transformed data, the Markov chain modeling could be developed for getting improved forecasts. Graph/bar/charts can be generated to have visual representations and comparison of the forecasts obtained by various types of Markov chain models. The Markov chain methodology requires sizable dataset to estimate the transition probability parameters. When the data set is small, the estimated transition probabilities will not be precise and many zero probabilities may occur in the transition probability matrices. To this end, the software could be extended to develop yield forecast models for crop based on simulation based Markov chains. For any two years data set, the first year data set could be blown up into a larger data set. Upon this larger data set Markov chain model could be build and forecasts can be made for second year. This software has been developed for only first order Markov chain forecast models. Attempts could be made to develop Markov chain models based on higher order Markov chains. Standard error of the Markov chain based forecasts can also be incorporated in the software.

ACKNOWLEDGEMENTS

Authors are grateful to the Associate Editor and referees for their valuable comments.

REFERENCES

Buchman, M. (2012). *Land Use Modeling Using Higher Order Markov Chains*. ProQuest LLC, Michigan.

Eviews ver. 7.2 (2011). *EViews (Econometric Views) - Statistical Package for Windows*. IHS Inc., USA.

Forecast Pro ver. 7.2 (2012). *Software Solution for Business Forecasting*. Business Forecast Systems, Inc., Massachusetts, USA.

Horstmann, C.S. and Cornell, G. (2004). *Core Java 2, Volume II-Advanced Features*. The Sun Microsystems Press JAVA SERIES, Pearson Education, New Delhi.

Jain, R.C. and Ramasubramanian, V. (1998). Forecasting of crop yields using second order Markov chains. *J. Ind. Soc. Agril. Statist.*, **51**, 61-72.

Jain, R.C. and Agrawal, R. (1992). Probability model for crop yield forecasting. *Biom. J.*, **34**, 501-511.

Loy, M., Eckstein, R., Wood, D., Elliott, J. and Cole, B. (2003). *Java Swing, Developing GUIs in Java*. O'REILLY, New Delhi.

Matis, J.H., Saito, T., Grant, W.F., Iwig, W.C. and Ritchie, J.T. (1985). A Markov chain approach to crop yield forecasting. *Agricultural Systems*, **18**, 171-187.

Pantham, S. (1999). *PURE JFC Swing*. Techmedia, New Delhi.

Ramasubramanian, V. and Jain, R.C. (1999). Use of growth indices in Markov chain models for crop yield forecasting. *Biom. J.*, **41**, 99-109.

Ramasubramanian, V., Agrawal, R. and Bhar, L.M. (2004). Forecasting sugarcane yields using multiple Markov chains. Project Report, Indian Agricultural Statistics Research Institute, New Delhi.

Ramasubramanian, V., Agrawal, R. and Bhar, L. (2010). Crop forecasting using multiple Markov chains. *Assam Statist. Rev.,* **24(1)**, 37-56.

Singh, R. and Ibrahim, A.E.I. (1996). Use of spectral data in Markov chain model for crop yield forecasting. *Jour. Ind. Soc. Remote Sensing*, **24**, 145-152.

Sonis, M. and Dendrinos, D.S. (2009). Socio-spatial dynamics and discrete non-linear probabilistic chains. In: M. Sonis and G.J.D. Hewings (eds.), *Tool Kits in Regional Science,* Springer-Verlag, Berlin Heidelberg, 177-197.