

Investigating the Trend and Impact of Some Variables on HYV Boro Rice Production in Northern Region of Bangladesh



*A Dissertation Submitted to the University of Rajshahi in Partial
Fulfillment of the Requirements for the Degree of
Master of Philosophy in Statistics*

Supervised By



Dr. Provash Kumar Karmokar
Associate Professor
Department of Statistics
University of Rajshahi, Bangladesh

Submitted By



Ranjan Kumar Kundu
M. Phil. Research Fellow
Roll No.: 07313, Session: 2007-2008
Department of Statistics
University of Rajshahi, Bangladesh

DECLARATION

This dissertation entitled, “Investigating the Trend and Impact of Some Variables on HYV Boro Rice Production in Northern Region of Bangladesh” submitted by me in the Department of Statistics, University of Rajshahi, Bangladesh for the award of the degree of Master of Philosophy is based on my research work carried out under the supervision of Dr. Provash Kumar Karmokar, Associate Professor, Department of Statistics, University of Rajshahi.

To the best of my knowledge, this work neither in part nor in full has been submitted to any other university or institution for the award of any degree.



Dr. Provash Kumar Karmokar
Associate Professor
Department of Statistics
University of Rajshahi
Rajshahi-6205, Bangladesh



Ranjan Kumar Kandu
M. Phil. Research Fellow
Roll No.: 07313, Session: 2007-2008
Department of Statistics
University of Rajshahi, Bangladesh

ACKNOWLEDGEMENTS

First of all, I express my deep emotion to thank the Almighty God for giving me strength, courage and patience enabling me in accomplishing the study. I feel great pleasure to express my deep sense of gratitude, profound appreciation and indebtedness to my reverend supervisor Dr. Provash Kumar Karmokar, Associate Professor, Department of Statistics, University of Rajshahi for his supervision and helpful decisions throughout the research period. He has taken more than ordinary care in shaping the work.

I am thankful to my respected teachers Professor Samad Abedin, Professor M. A. Razzaque, Professor Dr. M. A. Basher Mian, Professor Dr. Sushanta Kumar Bhattacharjee, Professor Dr. M. Assaduzzaman Shah, Professor Dr. Nurul Islam, Department of Statistics, University of Rajshahi for their advice helped me in innumerable ways to improve the work.

I would like to express my gratitude to my honorable teachers Professor Dr. M. Nasser, Chairman, Department of Statistics, and Professor Dulal C. Roy, without whom this research could never been accomplished. I am grateful to them for their proper and necessary guidance and suggestions from the very inception of this effort. I also acknowledge with thanks all of teachers, Department of Statistics, University of Rajshahi and well wishers who offered many thoughtful suggestions.

Sincere gratitude and indebtedness go to Samar Paul, Ex. Secretary, Rural Development and Co-operative Division, Govt. of the People's Republic of Bangladesh for his continuous encouragement and suggestions to conduct the study.

I am highly obliged to Md. Ekramul Alam, Principal, Bonpara Degree College, Natore and other members of the Governing Body who sanctioned me study leave, without which it could not at all be possible for me to continue this study.

Finally, a deep appreciation is expressed to my wife, Kumkum Kundu for her patience and rendering constant inspirations during the study period. Heartfelt indebtedness goes to my parents and all other family members of mine for their sacrifice and encouragement.

The Author

March, 2011.

Abstract

Boro rice is playing a significant role in the total rice production of Bangladesh. High yielding variety (HYV) Boro rice is grown enormously in the Northern region of Bangladesh. On the basis of per hector production, three HYV Boro rice producing districts have been selected by Simple Random Sampling for the present study. The selected three districts are Rajshahi, Rangpur and Dinajpur.

To fulfill the objectives of the present study Trend Analysis and Factor Analysis have been anticipated. Suitability of commonly used trend models like Linear, Compound and Quadratics trend models for the selected districts were checked by Mann-Kendall and Spearman non parametric statistical tests and parameters of these models were estimated by OLS technique. The DW statistics are obtained by OLS technique shows the evidence of autocorrelation in the data sets. To overcome this problem we have used Cochrane-Orcutt technique to estimate of parameter of the trend models. The Cochrane-Orcutt technique minimizes the autocorrelation problem. The JB test shows the normality assumption of errors for all the models. Although the R^2 and adjusted R^2 have increased for all the models it is remarkable of about 17%-18% increasing of these values for Rajshahi districts in both the Linear and Compound models. The HYV Boro productions of the selected districts have been forecasted up to year 2021 for three trend models using the data from 1971 to 2003 by Cochrane-Orcutt technique and the results are given in detailed in Chapter Three. It is seen from the result that the Compound Trend model performs better maintaining the growth rate 1.009 for Rajshahi, 1.012 for Rangpur, 1.008 for Dinapur and its forecasting performance is the best among the trend models.

Production of HYV Boro rice may depend on several climatic and non climatic variables like, rainfall, temperature, wind speed, humidity, cloud coverage, CO₂ emission, fertilizer

use and tractor use. In this study, we are also interested to identify the factors whose impacts are influential to the HYV Boro production of the selected districts of northern region of Bangladesh. Factor analysis, a data reduction technique have been used to identify the impact such variables on HYV Boro rice production in northern region of Bangladesh. Our analysis has identified five influential factors for the HYV Boro rice production for each of the districts. Therefore it is shown that the selected factors account for 82.33% of the total variance for Rajshahi district, 75.71% of the total variance for Rangpur district and 71.57% of the total variance for Dinajpur districts.

The findings of this study convey a message to the agriculturist, concerned authorities and policy makers that keep watching of climatic factors and proper handling of other relevant factors in Boro rice production of northern region that would be significant in food supplying in this region and enhance the food security of Bangladesh.

CONTENTS

	Page No.
Acknowledgements	iv
Abstract	v
List of Table	x
List of Figures	xiii
CHAPTER ONE: Background Study	1-28
1.1 Introduction	1
1.1.1 Rice around the world	2
1.2 History of Rice	4
1.3 Cropping pattern of Bangladesh	4
1.4 Rice in Bangladesh	8
1.4.1 Present Status of Rice Production in Bangladesh	8
1.4.2 Boro Rice in Bangladesh	11
1.5 Factors Affecting the HYV Boro rice	13
1.5. 1 Rainfall	13
1.5 .2 Temperature	14
1.5. 3 Wind Speed	16
1.5.4 Relative Humidity	17
1.5. 5 Cloud Coverage	18
1.5. 6 Carbon Dioxide	19
1.5 .7 Fertilizer use	19
1.5 .8 Tractor use	21
1.6 Literature review	21
1.7 Aim and Objective	28
CHAPTER TWO: Material and Methods	29-59
2.1 Introduction	29
2.2 Definition of Time Series	29

2.2.1	Time Series Components	30
2.3	Time Series Analysis	31
2.3.1	Stationary Process	31
2.3.2	Test of Stationary	33
2.3.2.1	Autocorrelation Function and Correlogram	34
2.3.2.2	Unit Root	35
2.3.2.2.1	Dickey Fuller and Augmented Dickey Fuller Unit Root Test	36
2.4	Trend Analysis	37
2.4.1	Non parametric Statistical Tests	38
2.4.1.1	The Mann-Kendall Rank Test	38
2.4.1.2	Spearman's Rank Test	39
2.4.2	Estimation Methods	40
2.4.2.1	Ordinary Least Square Method	40
2.4.2.2	Cochran-Orucutt Method	42
2.5	Data Reduction Techniques	43
2.5.1	Factor Analysis	44
2.5.1.1	Principal Factor Method	45
2.5.1.2	Unweighted Least Squares Method	48
2.5.1.3	Generalized Least Squares Method	51
2.6	Akaike's Information Criterion	52
2.7	Normality Checking	53
2.7.1	Jarque-Bera Normality Test	54
2.8	Sampling Design	56
2.8.1	Simple Random Sampling	56

2.9	Study Area	57
2.10	Sources of Data and Construction of Variables	58
2.11	Statistical Software	59

CHAPTER THREE: Trend Analysis **60-78**

3.1	Introduction	60
3.2	Trend Models	60
3.3	Estimation of Parameters	61
3.4	Results of Trends Models	62

CHAPTER FOUR: Factor Analysis **79-106**

4.1	Introduction	79
4.2	The Model for Factor Analysis	79
4.3	Estimation of Factor Loadings	80
4.4	Statistical Test used in Factor analysis	91
4.5	Results of Factor Analysis	93

CHAPTER FIVE: Result Discussion and Conclusion **107-123**

5.1	Introduction	107
5.2	Discussion of Results	107
5.2.1	Results of Trend Analysis	107
5.2.2	Results of Factor Analysis	113
5.3	Conclusion	117
5.4	Policy Implication	123
	References	124-133

List of Tables

Table	Title	Page No.
Table 1.1	Rice Production by Continent	3
Table 1.2	Rough rice production for the period 2000 -2008	3
Table 1.3	Top ten world rice produce countries 2009-2010	4
Table 1.4	Cropping pattern of Bangladesh	7
Table 1.5	Environment and season wise cropping pattern in Bangladesh	8
Table 1.6	Import of Rice in Bangladesh, 1990-91 to 2008-09 Bangladesh	10
Table 1.7	Contribution of rice to growth in total rice production, 1951–2008	12
Table 1.8	Critical temperature of rice plant at different growth stages	15
Table 1.9	Recommendation dose of nutrients for Boro rice production	20
Table 3.1	Desc riptive Statistics for HYV Boro rice production	62
Table 3.2	Mann-Kendal test statistic values for the period 1971 to 2007	65
Table 3.3	Spearman test statistic values for the period 1971 to2007	65
Table 3.4	AIC values of variables with trend for selected districts	66
Table 3.5	AIC values of variables without trend for selected districts	66
Table 3.6	Test and Critical values of variables with trend for selected districts	67
Table 3.7	Test and Critical values of variables without trend for selected districts	67
Table 3.8	Observed value of the τ_{μ} – statistic with the critical	67

	value	
Table 3.9	Repeated Dickey-Fuller test for order of integration of the variables	68
Table 3.10	OLS estimates of parameters for Linear Trend model	69
Table 3.11	OLS estimates of parameters for Compound Trend model	69
Table 3.12	OLS estimates of parameters for Quadratic Trend model	70
Table 3.13	R^2 , adjusted R^2 , JB and DW statistic by OLS for Linear Trend model	70
Table 3.14	R^2 , adjusted R^2 , JB and DW by OLS for Compound Trend model	71
Table 3.15	R^2 , adjusted R^2 , JB and DW by OLS for Quadratic Trend model	71
Table 3.16	Cochrane and Orcutt estimates by Linear Trend model	72
Table 3.17	Cochrane and Orcutt estimates by Compound Trend model	72
Table 3.18	Cochrane and Orcutt estimates by Quadratic Trend model	73
Table 3.19	R^2 , adjusted R^2 , JB and DW values by Cochrane-Orcutt method for Linear Trend model	73
Table 3. 20	Values of R^2 , adjusted R^2 , JB by Cochrane-Orcutt method for Compound Trend model	74
Table 3. 21	Values of R^2 , adjusted R^2 , JB by Cochrane-Orcutt method for Quadratic Trend model	74
Table 3.22	Method wise forecasted HYV Boro rice production of Rajshahi	75
Table 3.23	Method wise forecasted HYV Boro rice production of Rangpur	76
Table 3.24	Method wise forecasted HYV Boro rice production of	77

	Dinajpur	
Table 4.1	Interpretation of the KMO as characterized	92
Table 4.2	KMO and Bartlett's Test Statistics value for Rajshahi	94
Table 4.3	KMO and Bartlett's Test Statistics value for Rangpur	94
Table 4.4	KMO and Bartlett's Test Statistics value for Dinajpur	94
Table 4.5	Factor wise explained variation for Rajshahi	97
Table 4.6	Item Statements of Factors with Rotated Factor Loadings for Rajshahi	98
Table 4.7	Selected Factors with Rotated Factor Loadings for Rajshahi	99
Table 4.8	Factor wise explained variation for Rangpur	100
Table 4.9	Item Statements of Factors with Rotated Factor Loadings for Rangpur	101
Table 4.10	Selected Factors with Rotated Factor Loadings for Rangpur	102
Table 4.11	Factor wise explained variation for Dinajpur	103
Table 4.12	Item Statements of Factors with Rotated Factor Loadings for Dinajpur	104
Table 4.13	Selected Factors with Rotated Factor Loadings for Dinajpur	105

List of Figures

Figure	Title	Page No.
Fig 1.1	Year wise Population and Rice Production in Bangladesh	9
Fig 1.2	Area under Rice Production in Bangladesh	10
Fig 1.3	Bar diagram of Boro production of different divisions of Bangladesh	12
Fig 3.1	TS Plot of HYV Boro production for the district Rajshahi	63
Fig 3.2	TS Plot of HYV Boro production for the district Rangpur	64
Fig 3.3	TS Plot of HYV Boro production for the district Dinajpur	64
Fig 4.1	Scree Plot for Rajshahi	96
Fig 4.2	Scree Plot for Rangpur	96
Fig 4.3	Scree Plot for Dinajpur	97

Chapter One

Background of the Study

1.1 Introduction

Bangladesh is an over populated country having the density of population approximately 1104 per square kilometer. Agriculture is the most important sector for the economy of Bangladesh. The huge population of the country depends on agriculture sector for their foods and calories. About 84% people of Bangladesh are directly or indirectly engaged in a wide range of agricultural activities (Rahman, 2004). The agriculture sector contributes 19.10% at current prices to GDP and employs 50.28% of the labour force. As a staple food the contribution of rice is remarkable for the country. Rice provides about 75 percent of the calories and 55 percent of the protein in the average daily diet (Bhuiyan *et. al.*, 2002), and ensures political stability for the country. Since the stable supply of rice has great implications for food security, many views food security as synonymous to achieve self-sufficiency in rice production (Hossain *et. al.*, 2005). Food grain occupies the maximum share of agricultural GDP and it employs about 60 per cent of total labour force. (Alam, 2008). Rice is a major source of livelihood in terms of providing food, income and employment in Bangladesh. Rice occupies 10.58

million hectares of land which is about 77 percent of its cultivated area (BBS, 2008).

The population growth rate of Bangladesh is 2 million per year. According to this rate, the total population will become 233.2 million within 2050. In Bangladesh, about 80,000 ha of arable land are going out of production every year. (Mondal, 2010). However, the country faces a tremendous challenge for providing food security to the increasing population. Therefore, it is important to increase rice production in order to meet the growing demand for food emanating from population growth. Although, there have been ups and downs in the domestic production of food grain in Bangladesh, it is important to increase the production by engaging the High Yield Varieties (HYV) of rice to meet the present challenges of food crisis in Bangladesh.

1.1.1 Rice around the world

Rice is the second largest cereal in the world. It is the staple food for more than half of the world's population and is the main food source for seventeen countries in Asia and the Pacific and for eight in Africa (FAO, 2005). At the beginning of the 1990s, world rice annual production in the different countries was around 350 million tons this production is about and by the end of the last century it was 410 million tons. Now this annual production reached about 550 million tons around the world.

Rice is now cultivated in 113 countries in the world. Ninety one percent of the global harvested rice produced in Asia, 3.6% in South America, 3% in Africa,

1.8% in North America and only 1% in Europe. Continental wise production of rice is given in the Table 1.1

Table 1.1 Rice Production by Continent

Asia	574.2 million tons (91% of global rice harvest)
South America	22.6 million tons (3.6% of global rice harvest)
Africa	20 million tons (3% of global rice harvest)
North America	11.1 million tons (1.8% of global rice harvest)
Europe	3.4 million tons (1% of global rice harvest)
In addition, Oceania farms 148,000 tons of rice.	

(source : www.rice-trade.com)

According to USDA, rough rice production (in 000' tons) from 2000 to 2008 of Bangladesh, Asia and the World are given in Table 1.2

Table 1.2 Rough rice production for the period 2000 -2008

Year	World	Bangladesh	Asia
2000	594414	37633	541300
2001	595481	36469	541715
2002	564278	37784	511119
2003	585469	39232	528613
2004	597824	38404	537949
2005	623260	43141	564431
2006	627311	43504	569455
2007	647077	43204	586808
2008	661811	46505	600541

(Source: USDA, PSD June-10, 2009)

According to FAO, top ten world rice producing countries are given in the Table

1.3

Table 1.3 Top ten world rice produce countries 2009-2010

Country	China	India	Indonesia	Bangladesh	Vietnam	Thailand	Myanmar	Philippines	Brazil	Japan
Rice Production (in 000, Mt.)	166417 (32.7%)	132013 (26.0%)	52079 (10.2%)	38060 (7.5%)	34519 (6.8%)	27000 (5.3%)	24640 (4.8%)	14031 (2.8%)	10199 (2.0%)	9740 (1.9%)

(Source: FAO, 2010)

1.2 History of Rice

As a cereal grain, rice is the most important and staple food for a large portion of the world. The scientific name of rice is *Oryza sativa* and it belongs to the grass family. The word Ouliz of the Ningpo dialect of the old Chinese language become Oruz in Arabic and *Oryza* in Greek language which were changed to Ritz and rice. It is originated as dhan from dhanya from Sanskrit.

Oryza sativa is common Asian rice found in most producing countries which originated in the Far East at the foot of the Himalayas. *O. sativa japonica* grew on the Chinese side of the mountains and *O. sativa indica* on the Indian side. Rice has been cultivated in China since ancient times and was introduced to India before the time of the Greeks. Chinese records of rice cultivation go back 4,000 years.

1.3 Cropping pattern of Bangladesh

In Bangladesh, only nine crops- rice (73.94%), wheat (4.45%), jute (3.91%), rape and mustard (3.08%), lentil (1.54%), chickling vetch (1.25%), potato (1.13%),

sugarcane (1.12%), and chilli (1.05%) are grown on 1% or more of the crop acreage 14.61 million hectars and may be considered as major crops.

The minor crops in Bangladesh are gram (0.78%), millets and maize (0.60%), onion (0.58%), black gram (0.51%), sweet potato (0.45%), groundnut (0.40%), green pea (0.36%), sesame (0.33%), linseed (0.30%), garlic (0.20%), pea (0.12%), barley (0.10%), etc. Some crops including vegetables, spices, etc, occupy a very insignificant proportion of the gross cropped area and they altogether account for 1.57 percent. Some other crops of Bangladesh including vegetables, spices, tea, sugarcane, tobacco, betel nut and betel vine etc. occupy a very insignificant proportion of the gross cropped area. Among the minor cereals barley, maize black gram and differ varieties of millets are grown in various locality in Bangladesh.

Rice dominates the cropping pattern in most of the part of Bangladesh and it has been broadly divided into three classes as

- i. Aman
- ii. Boro, and
- iii. Aush

Generally Aman is harvested in December to January, Boro is March to May and Aus is July to August. Respectively and transplanted Aman is the most important and covers about 46.30% of the paddy area, followed by Boro (26.85%), aus (17.59%) and Broadcast Aman (B. Aman) (9.26%). Transplanted Aman (T. Aman) is grown almost everywhere in Bangladesh, while Broadcast Aman is mostly grown in the low-lying areas of the south and northeast. Boro is grown to a

certain extent in every district, especially in the irrigated part, while Aus is a well scattered crop.

Wheat is the second alternative cereal crop in Bangladesh next to rice in Bangladesh. It is grown mainly in the drier parts of the north and is cultivated only as a winter crop.

Jute leads the country's list of export crops but is ranked third in terms of area cultivated. It is confined mainly to the low-lying areas of the Brahmaputra-Jamuna and Padma floodplains.

Mustard is the fourth important crop. It is also grown mainly in the low-lying areas of Brahmaputra-Jamuna and Meghna floodplains.

Masur, Mung, Mashkalai, Chhola Khesari and Motor are the important varieties of pulses produced almost everywhere in Bangladesh. Masur is mainly grown in the Gangetic delta while Khesari is a well-scattered crop in the islands and chars area and the northern part of the country.

Now a day the substitute of rice potato is the most important of the winter vegetables and is broadly grown at Munshiganj, Comilla, Rajshahi, Rangpur, Dinajpur, Bogra, Joypurhat and Nilphamari districts. Sugarcane is the eighth important crop the country. It grows best in Rajshahi, Natore, and Chuadanga districts.

The cropping pattern is dependent on agro-ecological conditions and is specifically influenced by the time of the monsoon rains, the amount and distribution of rainfall, the occurrence of natural calamities such as storms and

cyclones and length of flooding periods. Though there are some local variations for the country as a whole, there are nine basic cropping rotations (Hossain, 1991).

Table 1.4 Cropping pattern of Bangladesh

Cropping Pattern		
Pre-Kharif Season	Kharif Season	Rabi Season
Rice / Jute	Fallow	Rabi crops
Rice / Jute	Rice	Fallow
Rice / Jute	Rice	Rabi crops
Fallow	Rice	Fallow
Rice	Rice	
Fallow	Rabi crops	
Mixed Rice	Fallow	
Rice (B. Aman)	Fallow	
Rice (Boro)	Rice (T. Aman)	
Rice (T.Aman)		
Mixed Aus rice and Kharif Crops.		

Normally, A large number of cropping patterns are follows in Bangladesh depending on the crop production environment which greatly influenced by land type, soil texture flooding regime rainfall and resource base of the farmers. Some dominant cropping patterns under variable crop production environment are as given in following Table 1.5

Table 1.5 Environment and season wise cropping pattern in Bangladesh

	Pre-Kharif Season	Kharif Season	Rabi Season
Rainfed condition	Boro Aus/Jute	Fallow	Wheat/Potato/Pulses/Oilseeds/ Sugercane
Irrigated condition	Fallow T Aus	T Aman Fallow	Wheat/Boro/Wheat/Potato/Tobacco/ Vegetables

1.4 Rice in Bangladesh

Present status of production of rice in Bangladesh and description of Boro rice are discussed below.

1.4.1 Present Status of Rice Production in Bangladesh

Rice production systems make a vital contribution to the reduction of hunger and poverty in Bangladesh. Total rice production in Bangladesh was 10.32 million tons in the year 1975-76 when the country's population was only 79.90 millions and cultivated rice area was 10.32 million ha. However, the country is now producing 27.32 million tons in 10.71 million ha rice area to feed more than 140 million people (BBS and DAE, 2007). Increase in rice production has been possible owing largely to the adoption of modern rice varieties on around 73% of the cultivated rice land which contributes to about 85% of the country's total rice production, modern rice cultivation technology, improvement irrigation facilities and applications of fertilizer and pesticides (BBS, 2006).

T. Aman had a vital role in increasing the total rice production till 1995-96; after 1995-96, Boro rice is playing a significant role in the total rice production. Boro

rice production gradually increased due to increase in cultivation area, high yield varieties, modern technology and improvement in irrigation facilities during winter season. In 2007-2008, Boro rice contributed to more than 58% of the total rice production, whereas, T. Aman 34% and Aus rice only 7%.

The rice cultivation area has mostly remained unchanged over the last few decades but there is a good sign to increase Boro rice area because it makes an important contribution to increase the overall rice production and the main source of livelihood for the farming community in Bangladesh. Rice production in different varieties of rice and population are given in Fig 1.1

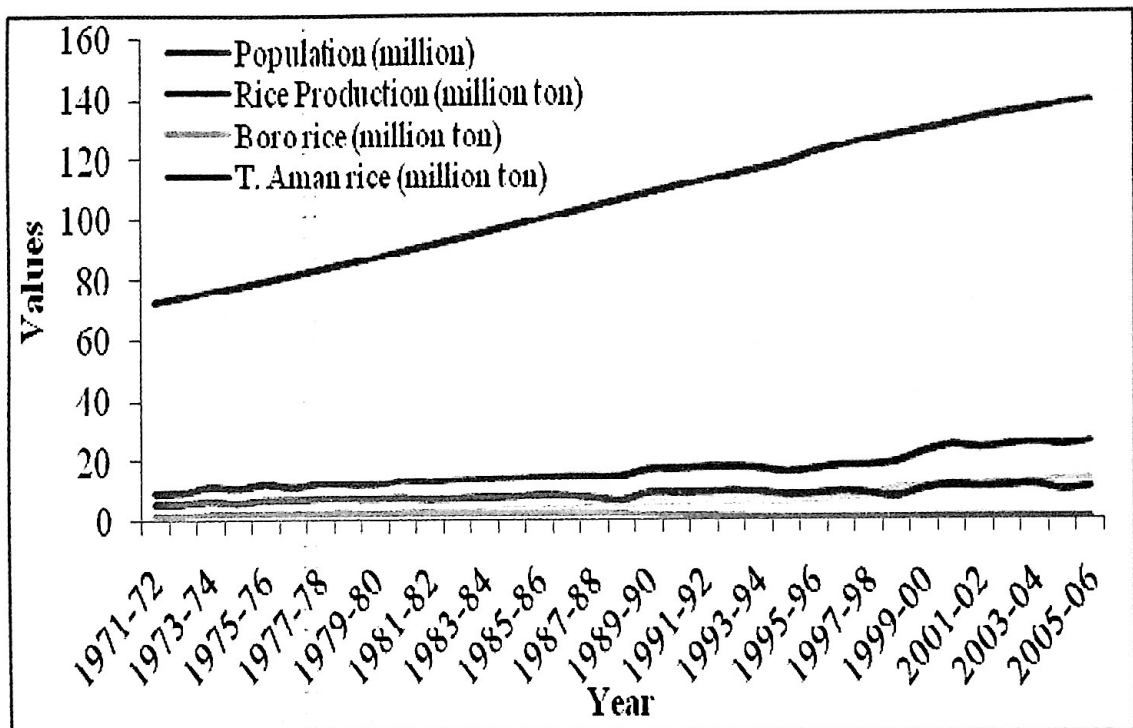


Fig 1.1 Year wise Population and Rice Production in Bangladesh

Year wise production area of different rice varieties are shown in Fig 1.2.

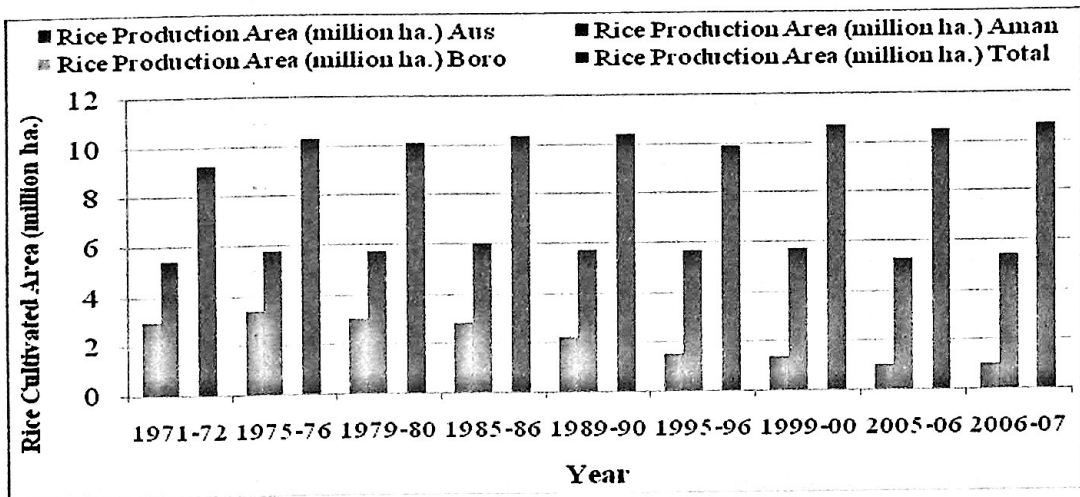


Fig 1.2 Area under Rice Production in Bangladesh

The rice import statistics of Bangladesh from year 1990-91 to 2008-90 are shown in Table 1.6.

Table 1.6 Import of Rice in Bangladesh, 1990-91 to 2008-09

Year	Total Import*
1990-1991	10
1991-1992	39
1992-1993	19
1993-1994	66
1994-1995	814
1995-1996	1144
1996-1997	34
1997-1998	1085
1998-1999	3056
1999-2000	433
2000-2001	562
2001-2002	127
2002-2003	1557
2003-2004	801
2004-2005	1295
2005-2006	53
2006-2007	720
2007-2008	2059
2008-2009	1070

(Source: Bangladesh Food grain Digest different issues * import in '000 metric tons)

1.4.2 Boro Rice in Bangladesh

The Boro rice is commonly known as winter rice which is grown in the winter and pre-monsoon seasons and harvested in April and May. Boro is a Bengali term originated from the Sanskrit word 'BOROB'. Modern rice varieties were introduced for the Boro and Aus seasons in 1967 and for the Amon seasons in 1970 (Hossain *et. al.*, 1994). In Bangladesh Boro rice occupied nearly 35% of the 10.80 million hectares of rice harvested area, and contributed 50% of the 38.7 million tons of rice produced in 2001-2002. The yield in 2001-2002 was 4.9 tons/ha (Singh *et. al.*, 2003b).

The main contributor has been the irrigated Boro rice production, which showed about 6.3 per cent annual growth during the same period. Currently, Boro rice accounts for about 60 per cent of total food grain production. It is notable that Boro production in 2007-08 increased by 18.87 per cent over the previous year and this largely contributed to the rise in total food grain (rice and wheat) production by 6.13 percent during this time. The significant jump in 2008 Boro production was mainly due to favorable weather conditions, timely planting, better care for input application and inter cultural operations by farmers, hard work and vigilance of the Agricultural Extension officials, diesel and electricity supply, and above all extra efforts and constant monitoring by the Ministry of Agricultural (Mondal, 2008).

There has been an overall increase in the production of the crop compared to last year as well as five years average, 1.4 percent and 15.2 percent respectively. Rajshahi division in the North West remains the main basket of Boro rice,

although there has been a slight setback compared to the previous harvest. A multiple bar diagram of Boro production of different divisions are given in Fig 1.3.

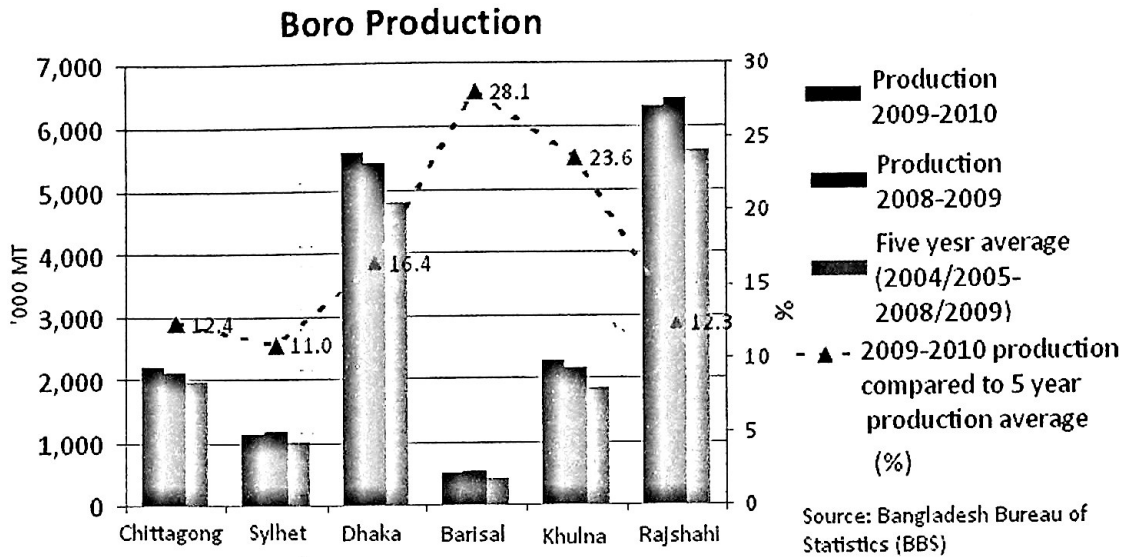


Fig 1.3 Bar diagram of Boro production of different divisions of Bangladesh

Overall increments of Boro production in comparison to the total production of rice are given the following table.

Table 1.7 Contribution of rice to growth in total rice production, 1951–2008

Period	Increase in total rice production (000 t/yr)	Increase in Boro rice production (000 t/yr)	Contribution of Boro to incremental rice production (percent)
1951–70	333	114	34
1971–90	500	312	62
1991–00	867	764	88
2001–08	971	942	97

(Sources: BBS, monthly statistical bulletins and statistical yearbooks, various years.)

1.5 Factors Affecting the HYV Boro Rice

Weather and climate are key determinants of the productivity of crops grown in many region of the world (Challinor *et. al.*, 2003). The production of major crops in Asian countries depends on surrounding climate and weathers. The impacts of climate change on food production are global concerns, and they are very important for Bangladesh.

The production of agriculture crops in Bangladesh like Boro rice are strongly interrelated with the climatic and non climatic factors like as rainfall, temperature, humidity, cloud coverage, sunlight, carbon dioxide emission, evaporation and fertilizer use, human labor, insecticide use, ground water level and seed management system.

1.5. 1 Rainfall

Agricultural production is affected by many uncontrollable climatic factors. Rainfall is one of them. The success of crop establishment and growth depend largely on the availability of adequate rainfall.

The effect of the rainfall is considerably good and beneficial for the crop at the time of germination.

Sastry (1976) emphasized that the success or failure of crops in both rainfed and irrigated agriculture is closely linked to rainfall pattern. Rainfall is one of the main meteorological phenomena affecting the growth and development of the plant and the size and quality of the harvest for a given acreage. Rainfall distribution pattern in relation to physiological phase of the crop variety and soil condition is a prerequisite for successful crop planning. It is usually the dominant climate factor,

not only in partially irrigated environments but also in fully irrigated environments (Bhuiyan and Galang, 1987).

Variability of rainfall distribution is the most important factor limiting yields of rainfed rice. It is about 80% of the growing rice in south and Southeast Asia.

According to Virmani, 1987 rainfall is the most dominant factor for rice culture.

Climate and agriculture are intimately related. Both long-term meteorological factors (climate) and short-term meteorological events (weather) affect crop growth, development, and production.

Deficiency in rainfall creates problems right from sowing to physiological maturity stage of Boro rice are discussed below:

- a) A deficit rainfall during May/early June hampers the summer ploughing which ultimately results in late sowing,
- b) A delay in onset of monsoon results again in delayed sowing.
- c) A break in monsoon immediately after sowing and germination results in decreased plant population.
- d) Any water stress or drought during vegetative and reproductive stages of rice adversely affects the productivity of rice crop.

1.5 .2 Temperature

Temperature is a dominant climatic factor, which affects the growth of rice plant and yield of rice. Growth and maturity of rice plants at different stages are governed by certain critical levels of low and high temperatures, normally below 20°C and above 30°C.

Temperature greatly influences not only the growth duration, but also the growth pattern and the productivity of rice crops. During the growing season, the mean temperature, and the temperature sum, range, distribution pattern, and diurnal changes, or a combination of these may be highly correlated with grain yields (Basak *et al.*, 2009).

Rice plant has nine growth stages with its three distinct growth phases and every stage has an optimum temperature range for its proper development. The critical temperatures for the development of the rice plant at different growth phases like vegetative, reproductive and ripening are shown in Table 1.5.2.1.

Table 1.8 Critical temperature of rice plant at different growth stages

Growth stages	Critical temperature (0C)		
	Low	High	Optimum
Germination	16-19	45	18-40
Seedling emergence	12	35	25-30
Rooting	16	35	25-28
Leaf elongation	7-12	45	31
Tillering	9-16	33	25-31
Initiation of panicle primordia	15	-	-
Panicle differentiation	15-20	30	-
Anthesis	22	35-36	30-33
Ripening	12-18	>30	20-19

(Source: Yoshida, 1978)

These critical temperatures differ according to variety, duration of the critical temperature, diurnal changes and physiological status of the plant (Yoshida, 1981).

Critical low and high temperatures for different growth stages of the rice plant have been identified (Yoshida, 1978). Depending on growth stage injury to the rice plant may occur when the mean daily temperature drops below 20°C (Yoshida, 1981). Low temperature injury to rice plants has been reported by a number of Asian rice growing countries including Bangladesh, China, India, Indonesia, Iran, Japan, Korea, Nepal, Pakistan and Sri Lanka (Yoshida, 1981).

According to Sastri and Minj the impact of increased maximum temperature by 1°C to 2°C during reproductive stage is detrimental as the yields decrease by 3-4 q/ha under 1°C rise of temperature and 5-6 q/ha due to 2°C rise in temperatures during reproductive stage. Such detrimental effect is seen both under rainfed and irrigated conditions. The potential rice yield decrease by about 5 q/ha under both the conditions. Also, if the duration of reproductive stage is decreased by 7 days due to rise in temperatures, the rice yield can drop by 10 q/ha. In the same way if there is forced maturity in the crop due to increase in temperatures during maturity stage, the potential rice yields can go down by 15-18 q/ha.

1.5. 3 Wind Speed

Pollination is central to successful reproduction in some plants. Simply stated, it is the transfer of pollen grains from the stamen of one flower to the stigma of the same or another flower. It is known from the pollen characteristics of rice that any pollen will be viable through weather like Wind Speed and Humidity. The length of time that pollen is viable can be weather dependent. Pollen is typically viable for 10-20 mm, but viability is dependent upon humidity and exposure to UV radiation (Gealy *et. al.*, 2003; Khush, 1993). In order to produce the harvested

crop it is required to successful pollination. Since rice is a self-pollinated or wind-pollinated plant its production may be increased if the ideal condition is present at the time of pollination period.

Only very low wind speeds are required for replenishment of CO₂ supply to the rice plant by turbulence in the crop canopy. Strong wind causes too much fluttering and waving of the crop canopy which

- (i) Interferes with ascent of sap and hence on its mineral nutrition
- (ii) Reduces formation of photosynthesis, and
- (iii) Poorer retention of assimilation product in ears.

Dry winds desiccate ovaries and increase sterility, blow-off the pollen from stigma. Especially those with feathery stigma or with a long gap between the opening and closing of lemma and palea (Saran *et. al.*, 1972).

1.5.4 Relative Humidity

Relative humidity is one of the important factor for the crop cultivation and finally for its production. Relative humidity below 40% inhibit flowering, which is best when relative humidity below is 70- 80% (Angladette, 1966). In the dry weather season growing of rice puddle conditions to ensure requisite relative humidity regime would appear necessary.

According to Huda, *et. al.*, (1975) a change of $\pm 1\%$ of maximum and minimum daily relative humidity from the average during the nursery period slightly affects the yield. The reproductive and ripening phases are susceptible to a change of $\pm 1\%$ maximum and minimum daily relative humidity and also to a change of $\pm 1^\circ\text{C}$ minimum daily temperature from their corresponding averages.

Humidity refers to water vapour content of the atmosphere and the quantity of water held in the atmosphere at any particular time can be expressed as absolute, specific or relative humidity.

$$\text{Absolute Humidity (g/m}^3\text{)} = \frac{\text{Weight of water vapour}}{\text{Volume of air}}$$

$$\text{Specific Humidity} = \frac{\text{Weight of water vapour}}{\text{Weight of air including water vapour}}$$

$$\text{Relative Humidity} = \frac{\text{Water vapour present in the air}}{\text{Water vapour required for saturation}} \times 100$$

1.5. 5 Cloud Coverage

Geographic limits and yields of different crops may be altered by changes in precipitation, temperature, cloud cover and soil moisture as well as increases in CO₂ concentrations. High temperatures and diminished rainfall reduce soil moisture, reducing the water available for irrigation and impairing crop growth in non-irrigated regions. A small increase in cloud cover could, in principle, balance the heating resulting from greenhouse gases. In paddy rice, maximum yields are obtained in the dry season, when cloud cover is less and photosynthetic active radiation is greater than during the wet season. Cloud coverage has an important role in agricultural crop besides the sensitiveness to canopy geometry and moisture contents. Low temperature, rainy and cloudy weather between the boot stages to ripening stage diminish the percentage of matured grain like Boro rice. High level of cloud coverage throughout the year may reduced the production of agricultural crop production

1.5. 6 Carbon Dioxide

Increasing levels of CO₂ in the atmosphere are likely to have a positive impact on crop yields, though there is uncertainty over the extent of this impact. The effects of CO₂ depend on crop type, soil and management factors; they may vary over time and there is interaction with many other environmental factors. Increases in both carbon dioxide levels and temperature will also affect rice production. Higher carbon dioxide levels typically increase biomass production, but not necessarily yield. Higher temperatures can decrease rice yields as they can make rice flowers sterile, meaning no grain is produced. Higher respiration losses linked to higher temperatures also make rice less productive.

The different predictions for elevated temperature, carbon dioxide levels, changes in humidity, and the interactions of these factors make forecasting future rice yields under these conditions challenging. IRRI research indicates that a rise in night time temperature by 1 degree Celsius may reduce rice yields by about 10%.

1.5 .7 Fertilizer Use

Fertilizer is the most important nutrient element in soils and plays the most vital role in crop production in Bangladesh. Fertilizer application mainly depends on the soil types, growing season, irrigation applications and the cultivars used. The demand for fertilizer is also affected by agro climatic conditions. High yielding varieties of Boro rice are highly responsive and need adequate supply of fertilizer to achieve targeted production. Urea (Nitrogen), Triple super phosphate (TSP), Murate of potash (MOP), Gypsum and Dasta (ZnSO₄) are the major fertilizers which are applied in agricultural land in various proportions for Boro rice

production in Bangladesh. Urea is applied in three steps after Boro rice transplanting and the other types of fertilizers are applied during the field preparation. The use of chemical fertilizers in Bangladesh agriculture was started sharply after the year of 1975. Urea application is significantly higher comparing to the other fertilizers application for crop production. Fertilizer application sharply increased with the introduction of high yielding varieties.

Information on nutrient balance for a particular crop production system is very useful in developing nutrient management strategies for sustainable production. The continuous and unbalanced use of the chemical fertilizers under intensive cropping systems has been considered to be the main cause for declining crop yield and environmental degradation. All essential elements must be present in optimum amounts and in forms usable by rice plants. Nitrogen, phosphorus, zinc, potassium and sulfur are nutrients elements most commonly applied by rice farmers. The amount of nutrients for Boro rice production in Bangladesh are given in the Table 1.9

Table 1.9 Recommendation doses of nutrients for Boro rice production

Type of fertilizer	Amount of nutrient (Kg/ha)	Average value of nutrient (Kg/ha)	Average dose of fertilizer (Kg/ha)
Urea (N)	110-130	120	267
TSP (P)	50-55	52.5	109
MOP (K)	79-82	80.5	134
Zinc Sulphate (Zn)	2	2	8
Gypsum (S)	10	10	59

(Source: BRRI, 2010 and Basak calculation)

1.5 .8 Tractor use

To increase production and cropping intensity it is one of the most important issues to make development of agricultural mechanization

Virtually, mechanization in Bangladesh agriculture started in 1960s with the introduction of tractor, power tiller, deep tube well, shallow tube well and low lift pumps on a very limited quantities. After independence in 1971, it gained pace, especially, in irrigation development. Though, tractors, power tillers, and other farm machinery like threshers and weeders increased with time, irrigation equipment increased much faster rates.

The use of modern technology during recent decades resulted in a rapid growth of the farm production. Tractors and farm machinery are important examples of this modern technology (Xinan, *et al.* 2005 and Singh, 2000). The quality of inputs in mechanisation and land and labour productivity may differ considerably (Singh, 1997, 2000; Singh & Chandra, 2002). The mechanisation technologies keep changing with the industrial growth of the country and socio-economic advancement of the farmers.

1.6 Literature review

We could not find any study directly addressing the analysis of trend and impact of some variables on HYV Boro rice production in northern region of Bangladesh. So we were not able to locate any study which has been done jointly to analyze the trend and impact of climatic or non climatic variables on agriculture production. But we have searched out a number of studies which have been conducted separately to detect trend and measure the impact of climatic or non climatic

variables on agriculture production. A number of works have been carried out to study the analysis of trend on various agriculture productions [see (Kumar and Mittal, 2006); (WARDA, 2005); (Wilson *et. al.*, 2004); (Anatoly MANELLYA, 1998); (Gupta *et. al.*, 1999); (Kumar, 1978)] in world wide.

Rajarathinam, *et. al.*, (2010) studied production and productivity trends and growth rates of tobacco (*Nicotiana tabacum*) crop grown in Anand region of Gujarat state, India for the period 1949-50 to 2007-08 based on parametric and nonparametric regression models. In parametric models different linear, non-linear and time-series models were employed and in nonparametric regression optimum bandwidth was computed by cross-validation method.

Omobowale *et. al.*, (2009) used trend analysis to examine the trends of six agricultural staple food crops like maize, cassava, rice, yams, millet, and sorghum of Nigeria and they have found that rice follows a linear trend for their data.

Karmokar and Imon (2008) have used linear, compound and quadratic trend models using all variety of rice production data in Bangladesh. They made a comparison among the selected models and found that the compound trend model outperforms the other two models in all respect and finally they have concluded that Bangladesh is maintaining an encouraging compound growth rate of 1.0227 per year in the production of rice.

Munir *et. al.*, (2008) estimated growth trends in the Production of Mango in Pakistan by Log Linear Model using time series data for the years 1982-2004.

Zare *et. al.*, (2006) examined trends in cultivated area, yield and production of wheat, barley and rice as the main cereal crops of Iran were studied during 1350-

1382 using time series analysis and the situation of future were predicted for 1382-1400 period.

They showed that the last 32 years cereals yield is increased considerably. Mean annual rate of yield increase was much higher for irrigated cereals (62.5kg year⁻¹) and rice (69.7kg year⁻¹) compared to rainfed cereals (12.6kg year⁻¹), this incremental trend will be continued towards year 1400. The increase of cereal yield in spite of constant cultivated area is led to increase in total production and it was predicted that total cereal production of country will be exceed 26 million tons in target year of 1400. However, per capita cereal production which was at its highest in the mid years of the study period, was decreased later and at present is about 310kg which is not much different from the baseline year (1350). Per capita production of cereals will remain unchanged until year 1400 because of high predicted population growth rate.

Nguyen Cong Thanh *et. al.*, (2006) has conducted a study on rice production trend and export in Vietnam last 40 years from 1965 to 2004 has. Using Compound Trend Model they showed that Compound Growth Rates (CGRs) for overall period, of rice area, yield and production in Vietnam were positive and significant growth at the rate of 1.34; 2.60 and 3.97 per cent per annum, respectively. These were all positive and significant growth at 1 per cent level of probability. The instability analysis for overall period showed that overall area, yield, and production were high instability as compared to each sub-period in which most high instability was found in rice production.

Hafner (2003) found a prevalence of linear growth, i.e. constant annual yield increase, on a global scale. Trends of slowing yield growth that are indicated for some countries are caused by economic and biophysical factors. Gupta *et. al.*, (1999) used growth and trend models to analysis the potato production of West Bengal, India. A variety of trend model are now available in the literature for analyzing the production of major crops.

Sahu (2003) studied growth and trend in production of some selected crops like rice, wheat, potato and mustard using linear, quadratic and compound trend model for agricultural production data of West Bengal, India.

Azad *et, al.*, (2007) identified the quantitative and qualitative trends of milk production of Baghabarighat Milk Shed Area (BMSA) of Milk Vita from January, 2001 to December, 2005. They was found a significant trend among the seasons, and milk productions

Bhagyashree and Rajarathinam (2009) reported that nonparametric regression model was found suitable to fit the trends in area under the *bajra* crop in Gujarat state during the period (1949-50 to 2003-04). Bhagyashree (2009) employed nonparametric regression model to study the trends in area, production and productivity of tobacco crop grown in Gujarat state during the period (1949-50 to 2003-04). The results of the present study are in agreement with these researchers.

Finger, (2007) analyzed trends in yield growth and yield variability of barley, maize, oats, rye tritcale and wheat in Switzerland from 1961 to 2006 using linear and quadratic trend models and showed that there have been linear increases in crop yields since the 1960s.

Karim, *et. al.* (2010) studied trend analysis of wheat production of three major districts like Dinajpur, Rajshahi, and Rangpur for the period 1971-72 to 2001-04 using several model selection criteria.

Abedullah and Pandey (2007) estimated a heteroskedastic quadratic production function using panel data gathered from 46 rice farmers in Tarlac, Philippines for the period 1990-1995. Their study shows that aside from farm inputs, rainfall and its interaction with farm inputs (especially fertilizer) also have significant effects on rice production.

Tao, *et. al.* (2008), on the other hand, examined the relationships between rice yields and growing season climate in the main production regions of China over the period 1951-2002. They show that rice yield was significantly influenced by the growing season climate. In particular, growing season temperature, which had a generally significant warming trend, had increased rice yields in northeast China over the period.

Although rice production of Bangladesh is influenced by seasonal characteristics and different variables of climatic and non climatic such as rainfall, temperature, humidity, cloud coverage, sunlight, carbon dioxide emission, evaporation and fertilizer use, human labor, insecticide use, ground water level, seed management, agriculture credit, population growth etc.

Despite technological advances such as improved crop varieties and irrigation systems, weather and climate are important factors, which play a significant role to agricultural productivity. But a reasonable amount of Boro rice production is hampered in every year as a consequence of the above factors interactive over

different scale in different locations. Therefore it is very essential to the researcher to develop dissemination of knowledge and awareness building about climate change and proper utilization of non climatic variables and possibilities adaptation for cultivator.

So we are interested to analyze the study using maximum variables there many techniques of multivariate analysis such as Principal Component Analysis, Factor Analysis, Discriminant Analysis. Among them we have selected factor analysis for our study.

Factor Analysis is done to find out the underlying dimensions or factors that actually contribute to a number of observed attributes and it can be used in any research for data reduction. Factor analysis had successfully adopted and used in various field such scientific fields, behavioral and social sciences fields, marketing and banking sector etc in different countries of the world. Many studies [see Anna *et. al.*, (2005); Babakus, *et. al.*, (1987); Bartholomew *et. al.* (2002); Coughlan, *et. al.*, (2007); Distefano *et. al.*, (2007); Fabrigar *et. al.*, (1999); Floyd & Widaman, (1995); Flora and Curran (2004); Paul and Erik (2005); Michaelides M. (2007); Rigdon and Ferguson (1991); Smith *et. al.*, (2006); Subbarao, *et. al.*, (1995); Widaman, (1993)] have been done to extracted contributory factor using factor analysis. In that view we observe a number studies related to factor analysis in following published by various authors.

Love, *et. al.*, (2006) adopted Factor Analysis in groundwater quality management, it is important to relate the spatial distribution of different chemical parameters to different possible sources, which have different chemical signatures.

Alimba and Akubilo (2002) applied factor analysis to investigate the extent of technological change and the effect on rural farm enterprises in southeastern Nigeria and reported that technological change was neutral to most of the perceived negative consequences associated with change. They also recommend that all the identified agencies, institutional and social problems relating to technological adoption and use in farms in the area must be tackled if the needed transformation of agriculture is to be achieved in the new century.

Tsuyoshi *et. al.*, (2007) used factor analysis to measure the stochastic properties of the daily variations in drug sales in community medicine. They extracted four factors from data of 31 drugs

Bukantis, A. (2002) applied Factor analysis (FA) to identify climate-forming factors and quantitatively evaluate their importance in using average maximum and average minimum of air temperature (T_1 and T_2), atmospheric pressure (P), wind speed at a height of 10 m (v), monthly precipitation (Q), duration of sunshine (S) and partial pressure of water vapour (e) from seven meteorological station data the eastern part of the Baltic Sea region and used.

H. Boyacioglu *et. al.*, (2005) applied the factor analysis technique to the large water quality data sets Buyauk Menderes River Basin in analyzing the surface water contamination and determining the correlations between water quality parameters. Two factors are found out to be significant in representing the water quality parameters recorded in the basin.

1.7 Aim and Objective

The people of Bangladesh are popularly referred to as "Macche-Bhate Bangali" or "fish and rice makes a Bengali." Rice and fish are essential part of the life of Bangladeshi people from time immemorial. Rice farming is the single most important livelihood for a vast majority people of the country. To ensures political stability, achieve development goal and the self-sufficiency in food of Bangladesh it is very important to have an idea about the contribution of rice production especially HYV Boro rice production of Bangladesh from different statistical viewpoints. As the production of HYV Boro rice depends on some important climatic and non climatic factors so it is also very essential to find out the the responsible factors that are very influential for the production of HYV Boro rice in the northern region of Bangladesh. Therefore, HYV Boro rice production trend measure with its future prediction and extraction of influencing factors in northern region of Bangladesh using data from 1971 to 2007 is our objectives for the proposed study.

On the above goals the main aims and objectives of the study are given below:

- To have an idea on the HYV Boro production in the northern area of Bangladesh.
- To judge the pattern/trend of HYV Boro Rice production of study area.
- To forecast of HYV Boro production for the selected region using trend model.
- To construct the contributory factors among a set of climatic and non climatic variables through data reduction technique for HYV Boro Rice production in the northern region of Bangladesh.

Chapter Two

Material and Methods

2.1 Introduction

Definition of Time Series, Time Series Components, Test of Stationary, Trend Analysis, Nonparametric Tests, Data Reduction Technique especially Factor Analysis and some diagnostics tools are discussed in this chapter.

2.2 Definition of Time Series

Some variables related with our everyday life experience changes with the change of time. Arrangement of statistical data according the change of time is known as time series. A time series can be defined as a collection of readings of some phenomenon belonging to different time periods. Data, which are obtained from the observations of a phenomenon over time, are extremely common and named as Time Series (Cryer 1986). A time series is a sequence of observations which are ordered in time (or space). According to Ya-lun Chow "A time-series may be defined as a collection of readings belonging to different time periods, of some economic variable or composite of variables." W.Z. Herisch defined "A time series may be defined as a sequence of values of some variable corresponding to successive points in time."

Change of commodity price, population, yearly agriculture production, monthly sales, weekly share prices, crime figures, yearly export, imports etc. are some examples of time series.

Mathematically, a time series is defined by the relationship

$Y_t = f(t); t = t_1, t_2, \dots, t_n$, where Y_t is the value of phenomenon under consideration at time t .

2.2.1 Time Series Components

An important step in analyzing time series data is to consider the types of data patterns, so that the models most appropriate to those patterns can be utilized. The various factors that affect the values of the phenomenon in time series may be termed as components of time series. The following four types of time series components can be distinguished. These are:

1. **Trend** : In general, natural, human, economic and other processes produce gradual trends. A trend is a long-term component that represents a growth or a decline of a time series over an extended period of time.
2. **Seasonal Variation**: This term of seasonality is used for time series defined at time intervals which are fractions of a year. It is a pattern of change that repeats itself from year to year.
3. **Cyclical Variation** : Changes in time series sometimes show a wavelike fluctuation around a trend, which shows the possible existence of periodicity with longer intervals.

4. Irregular Variation : This is a part of a time series represented by residuals, after the above-mentioned components have been removed.

2.3 Time Series Analysis

Time series analysis consists of identification, isolation, measurement and study of the components of time series data. According to Fan and Yao, (2003) "Time series analysis deals with records that are collected over time." The objective of time series analysis is to discover the magnitude and direction of the components of the series.

There are two main goals of time series analysis:

- (a) Identifying the nature of the phenomenon
- (b) Forecasting.

Both of these goals require that the pattern of observed time series data is identified and more or less formally described. Once the pattern is established, we can interpret and integrate it with other data (i.e., use it in our theory of the investigated phenomenon, e.g., seasonal commodity prices). Regardless of the depth of our understanding and the validity of our interpretation (theory) of the phenomenon, we can extrapolate the identified pattern to predict future events.

2.3.1 Stationary Process

In the time series analysis, a stationary process or strongly stationary process is a stochastic process whose joint probability distribution does not change when shifted in time or space. As a result, parameters such as the mean and variance,

covariance if they exist, also do not change over time or position. That is a stationary process has the property that the mean, variance and autocorrelation structure do not change over time. It is a common assumption that the data are stationary in time series analysis.

According to Challis and Kitney (1991) Stationarity, is defined as a quality of a process in which the statistical parameters (mean and standard deviation) of the process do not change with time.

The most important properties of a stationary process are following:

- (i) The auto-correlation function (ACF) depends on lag alone and does not change with the time at which the function was calculated.
- (ii) A weakly stationary process has a constant mean and ACF (and therefore variance)
- (iii) A truly stationary or strongly stationary process has all higher-order moments constant including the variance and mean.

Stationarity is used as a tool in time series analysis, where the raw data are often transformed to become stationary. When the economic data are often seasonal and/or dependent on a non-stationary, it is essential to make it stationary for the significant decision from the data.

Transformations to Achieve Stationarity: If the time series is not stationary, we can often transform it to stationarity with one of the following method.

- i) We can difference the data. That is, given the series Y_t we create the new series

$\Delta Y_t = Y_t - Y_{t-1}$ and it may be stationary. The differenced data will contain one less point than the original data and usually one difference become sufficient. If we difference the data more than once, the first difference may not be stationary.

ii) If the data contain a trend, we can fit some type of curve to the data and then model the residuals from that fit. Since the purpose of the fit is to simply remove long term trend, a simple fit, such as a straight line, is typically used.

iii) For non-constant variance, taking the logarithm or square root of the series may stabilize the variance. For negative data, you can add a suitable constant to make the entire data positive before applying the transformation. This constant can then be subtracted from the model to obtain predicted (i.e., the fitted) values and forecasts for future points.

The above techniques are intended to generate series with constant location and scale. Although seasonality also violates stationarity, this is usually explicitly incorporated into the time series model.

2.3.2 Test of Stationary

There are two general approaches to testing for stationarity, parametric and nonparametric. From the literature it is seen that parametric approaches are commonly used by researchers working in the time domain, such as economists, who are making certain assumptions about the nature of their data.

Nonparametric approaches are also used by the researchers working in the frequency domain, such as electrical engineers etc. These tests are not based on the knowledge or assumption that the population is normally distributed (Bethea

and Rhinehart 1991). By making no assumptions about the nature of the data, nonparametric tests are more widely applicable than parametric tests which often require normality in the data. While more widely applicable, the trade-off is that nonparametric tests are also less powerful than parametric tests. To arrive at the same statistical conclusion with the same confidence level, nonparametric tests require anywhere from 5% to 35% more data than parametric tests (Bethea and Rhinehart 1991).

2.3.2.1 Autocorrelation Function and Correlogram

An important guide to the persistence in a time series is given by the series of quantities called the sample autocorrelation coefficients, which measure the correlation between observations at different times. The set of autocorrelation coefficients arranged as a function of separation in time is the sample autocorrelation function, or the ACF. Correlogram or Autocorrelation Function (ACF) is a graphic of autocorrelation values from several time intervals in time series data. ACF explains how big the successive data correlation in a time series data. ACF can be used to determine whether a time series data is stationary or not.

One of the simple tests of stationary is based on the so-called autocorrelation function (ACF). The ACF at lag k , denoted by ρ_k and is defined as $\rho_k = \frac{\gamma_k}{\gamma_0}$

where γ_k is the covariance at lag k and γ_0 is the variance of the process. Therefore lies between -1 to +1. Plotting $\hat{\rho}_k$ against k we get the sample autocorrelation function. It is also called Correlogram.

Let us consider be the error from a first order auto-regressive AR(1) and define

$$\varepsilon_t = \rho \varepsilon_{t-1} + v_t$$

Where, $E(v_t) = 0$, $\text{var}(v_t) = \sigma_v^2$ and $\text{cov}(v_t, v_s) = 0$, for $t \neq s$. v_t is a new shock to the level of the variable.

Now, $E(\varepsilon_t) = 0$, $\text{var}(\varepsilon_t) = \sigma_\varepsilon^2 = \frac{\sigma_v^2}{1-\rho^2}$ and $\gamma_k = \text{cov}(\varepsilon_t, \varepsilon_{t-k}) = \sigma_\varepsilon^2 \rho_k$, $k > 0$, k is the lag of order.

$$\text{Hence } \hat{\rho}_k = \text{corr}(\varepsilon_t, \varepsilon_{t-k}) = \frac{\text{cov}(\varepsilon_t, \varepsilon_{t-k})}{\sqrt{\text{var}(\varepsilon_t) \times \text{var}(\varepsilon_{t-k})}} = \frac{\sigma_\varepsilon^2 \rho_k}{\sqrt{\sigma_\varepsilon^2 \times \sigma_\varepsilon^2}}$$

Correlogram or Autocorrelation Function (ACF) is a graphic of autocorrelation values from several time intervals in time series data. ACF explains how big the successive data correlation in a time series data. ACF can be used to determine whether a time series data is stationary or not.

2.3.2.2 Unit Root

A unit root is an attribute of a statistical model of a time series whose autoregressive parameter is one. Let us consider consider the model for the data series $\{y_t\}$ as;

$y_t = y_{t-1} + \varepsilon_t$, where ε is the stochastic error that follows the classical assumptions

- i) ε_t has mean 0
- ii) ε_t has constant variance σ^2 and is non autocorrelated

Here ε_t is some time called a white noise and it is a auto regression of order one that is $AR(1)$. Now if the coefficient of y_{t-1} equal to 1 so the series $\{y_t\}$ has a unit root and this will be treated as a unit root problem on the other hand the series $\{y_t\}$ is nonstationary. Therefore if we find $\rho = 1$ by regressing $\{y_t\}$ on y_{t-1} using the equation $y_t = \rho y_{t-1} + \varepsilon_t$ we may conclude that the stochastic variable y_t has unit root.

2.3.2.2.1 Dickey Fuller and Augmented Dickey Fuller Unit Root Test

A stationary series is generally characterized by a time-invariant mean, variance and covariances. However, most of the macroeconomic time series are found to be nonstationary. The nonstationarity of a time series can be tested by the unit root tests first proposed by Dickey and Fuller (1979) and modified by Said and Dickey (1984). These tests are generally known as Dickey-Fuller (DF) and augmented Dickey-Fuller (ADF) tests. These tests are conducted on the following Dickey-Fuller augmented model

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \alpha_2 t + \sum_{j=1}^p \beta_j \Delta y_{t-j} + \varepsilon_t \quad (2.1)$$

where, y_t is the variable of interest, y_{t-1} is the one-period lag value of y_t , t is the time trend, $\Delta y_t = (y_t - y_{t-1})$, and $\alpha_0, \gamma, \alpha_2, \beta_j$ are parameters, and ε_t is a random disturbance term with mean zero and constant variance.

The fourth term on the right-hand-side of (2.1) is the augmentation term. The Augmented Dickey-Fuller (ADF) test is based on the following hypotheses

$H_0: \gamma = 0$ (nonstationary) against the alternative hypothesis that

$H_1: \gamma < 0$ (stationary) in (2.1)

In the absence of the fourth term on the right-hand-side of the equation (2.1), test of the hypotheses in (2.1) is known as the Dickey-Fuller (DF) test. The third term on the right-hand-side of (2.1) indicates the deterministic trend (if any) in the data set. The intercept term may or may not be present in the model depending upon the data under consideration. When the model contains (i) no intercept and no trend terms, (ii) intercept but no trend term, and (iii) both intercept and trend terms then the DF and the ADF test are leveled τ , τ_μ and τ_τ respectively, for testing the hypotheses in (2.1).

2.4 Trend Analysis

Long term upward or downward tendency of a variable is known as trend. The term "trend analysis" refers to the concept of collecting information and attempting to spot a pattern, or trend, in the information. In some fields of study, the term "trend analysis" has more formally-defined meanings. Although trend analysis is often used to predict future events, it could be used to estimate uncertain events in the past, such as how many ancient kings probably ruled between two dates, based on data such as the average years which other known kings reigned.

In statistics, trend analysis often refers to techniques for extracting an underlying pattern of behaviour in a time series which would otherwise be partly or nearly

completely hidden by noise. A simple description of these techniques is trend estimation, which can be undertaken within a formal regression analysis.

2.4.1 Non parametric Statistical Tests

Besides the graphical presentation a numerous non-parametric tests [Kruskal–Wallis’s test, Wald–Wolfowitz’s test, Whitney–Pettitt’s test (Pettitt, 1979), Mann–Kendall’s test, Spearman’s test] are seen in the literature to detect the trend of a data. Among these statistical tests Mann-Kendall rank statistic and the Spearman rank statistic method are chosen for the present work.

2.4.1.1 The Mann-Kendall Rank Test

The Mann-Kendall test is a non-parametric test for identifying trends in time series data. Kendall & Stuart (1961) has suggested that this is a powerful test when the most likely alternative to randomness is linear or non linear trend. The test compares the relative magnitudes of sample data rather than the data values themselves (Gilbert, 1987).

The Mann-Kendall rank statistic t_m is computed, by first replacing the observations x_i ’s by their ranks k_i ’s such that each term is assigned a number ranging from 1 to n which reflects its magnitude relative to the magnitudes of all other terms. For each element k_i the number N_i is calculated as the number of k_j terms preceding it such that $k_j > k_i$. Then t_m is given by,

$$t_m = \frac{4 \sum_{i=1}^{n-1} N_i}{n(n-1)} \quad (2.2)$$

t_m is distributed very nearly as a normal distribution for large n and can be used as the basis of a significance test,

$$r_m = \pm r_g \sqrt{\frac{4n+10}{9n(n-1)}} \quad (2.3)$$

where r_g is the desired probability point of the normal distribution appropriate to a two-tailed test. If t_m lies inside the range $\pm r_m$ then the time series does not contain a trend.

Mann (1945) originally used this test and kennel (1962) subsequently derived the test statistic distribution. In trend test, the null hypothesis H_0 is that there is no trend of population from which the dataset is drawn; hypothesis H_1 is that there is a trend in the analyzed records.

2.4.1.2 Spearman's Rank Test

Spearman's rank-correlation method is simple and distribution-free, i.e. it does not require the assumption of an underlying statistical distribution. Yet another advantage is its nearly uniform power for linear and non-linear trends. The method is based on the Spearman rank-correlation coefficient, ρ (Dahmen and Hall, 1990; Sneyer, 1990) which is defined as:

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (2.4)$$

where, $d_i = kx_i - ky_i$, when two or more observations, d_i , have the same value, the average rank ky_i is calculated. A test-statistic t_i is used to test the null hypothesis

$H_0 : \rho = 0$ (there is no trend) against the alternative hypothesis $H_1 : \rho \neq 0$ (there is a trend). The test statistic is defined as

$$t_t = \sqrt{\frac{n-2}{1-\rho^2}} \quad (2.5)$$

For large n , the value of r_s can be tested for significance by calculating the quantity t_s given by the equation

$$t_s = r_s \sqrt{\frac{n-2}{1-r_s^2}} \quad (2.6)$$

If the computed t_s lie within the desired confidence limits of the two-tailed test, we can conclude that there is no trend in the data series.

2.4.2 Estimation Methods

Two regression Methods are employed in this study to estimate the parameters of different trend models. The methods are

- (i) Ordinary Least Squares (OLS)
- (ii) Cochrane-Orcutt (CO).

2.4.2.1 Ordinary Least Squares Method

One of the oldest techniques of modern statistics (Harper, 1974), Method of Ordinary Least Square (OLS) was first described by German mathematician Carl Friedrich Gauss around 1794 (Bretsche, 1995). It is more convenient and widely used method among various method of estimation of unknown parameters on regression analysis technique. The principles of Ordinary Least Square are to

estimate the unknown parameters after minimizing the error of the model. This method possesses attractive statistical properties and made it one of the powerful and popular methods of regression analysis.

Let us consider a simple regression model

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i, i = 1, 2, \dots, p.$$

Where, Y_i is the i th dependable variable (response). X_i 's are predictors, β 's unknown parameters and u_i 's are uncorrelated variables errors distributed as $N(0, \sigma^2)$. For different values of i the models are

$$Y_1 = \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + u_1$$

$$Y_2 = \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + u_2$$

.....

$$Y_n = \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + u_n$$

In matrix notation the equations can be written as

$$Y = XB + U$$

where, $X = [X_{ij}]_{n \times k}$, $B = [\beta_j]_{k \times 1}$, $U = [u_i]_{n \times 1}$, $Y_i = [Y_i]_{n \times 1}$ and $E(U) = 0$.

$V(U) = \sigma^2 I$, I is an identity matrix of order n .

In this simplest form of OLS, it is generally assumed that X is a $n \times k$ matrix of full rank, column rank k which implies the $X'X$ is nonsingular. According to the principle of OLS method if $\Psi = U'U = (Y - XB)(Y - XB)'$ is minimizing them we get the solution of B . that is

$$\frac{\partial \Psi}{\partial B} = 0 \text{ satisfy the above condition}$$

$$\Rightarrow 2 (X'X - X'XB) = 0.$$

$$\Rightarrow \hat{B} = (X'X)^{-1} (X'Y)$$

Which is the Least Square Estimate of B.

2.4.2.2 Cochrane-Orcutt Method

Estimation of parameters of a linear regression model is based on the assumption that the errors are normally independently distributed with zero means and constant variance. Sometime data of a regression model involve regressors and response variables that have a natural sequential order over time. Such data are called time series data. The assumption of uncorrected and or independent errors for time series data is often not appropriate. So the errors in time series data exhibit serial correlation (Box and Jenkins, 1976). Serial correlation or autocorrelation refers to the correlation between members of a series of numbers arranged in time. The sources of autocorrelation and consequences of autocorrelation are as follows:

Sources of Autocorrelation

1. Omitted explanatory variables
2. Misspecification of the mathematical form of the model
3. Interpolation in the statistical observations
4. Misspecification of the true random error

Consequences of Autocorrelation

1. the OLS parameter estimates will remain linear and unbiased
2. the estimated variances of the OLS estimators are biased.
3. the variance of the error term may be seriously underestimated (biased)
4. R^2 will be unreliable and will be biased upwards.

The Cochrane and Orcutt method is a technique of estimation of parameters if autocorrelation is appear in the data. The procedure following steps should be followed in such estimation.

1. Run the regression for the original model.
2. Take the estimated residuals from the regression and run the equation $\hat{\varepsilon}_t = v_t + \rho \hat{\varepsilon}_{t-1}$ v_t here is essentially the intercept term.
3. The estimate of ρ obtained from 2 and the generalized differenced form is estimated. From this OLS regression we noted the residuals sum of squares.
4. We take this latter set of residuals and with the aid of the regression model in 2 construct another set of variables and run the generalized model again with these newly constructed variables and note the residual sum of squares.

The above steps will be done repeatedly until the residual sum of squares reduces significantly. This will give the best fit of the model.

2.5 Data Reduction Techniques

Data reduction techniques are applied where the goal is to aggregate or amalgamate the information contained the large data into manageable information nuggets. Data reduction techniques can include simple tabulation, aggregation or more sophisticated techniques like Principal component Analysis and Factor Analysis. As a data reduction technique the factor analysis focuses on the interdependencies between more variables, by reducing the initial data volume to build a smaller set of factors with a minimum loss of information. Therefore the factor analysis could help us for reducing the number of variables and/or for the classification of this variable.

2.5.1 Factor Analysis

Factor analysis is a data reduction technique. It is a statistical approach of finding a way of condensing the information contained in a number of original variables into a smaller set of dimensions (factors) with a minimum loss of information. According to (Hair *et. al.*, 1992) factor analysis is a statistical approach that can be used to analyze interrelationships among a large number of variables and to explain these variables in terms of their common underlying dimensions (factors). Factor analysis extracts maximum common variance from all variables and puts them into a common score. Factor analysis can be considered as an extension of principal component analysis (Johnson and Wichern, 2003). Factor analysis attempts to simplify complex and diverse relationships that exist among a set of observed variables by uncovering common dimensions or factors that link together the same unrelated variables, and consequently provides insight into the underlying structure of the data (Dillon and Goldstein, 1984). Factor analysis extracts principal factors that explain much of the total variation. As the exact external shape of the factor structure is not unique, one factor solution can be transformed into another one or rotated to a terminal solution. This can achieve simpler and more meaningful factor patterns, instead of the highly complex extracted factors that are related to many of the variables rather than to just a few (Kim, 1970; Comrey and Lee, 1992). Factor analysis includes a variety of correlation analyses designed to examine the interrelationships among variables (Carr, 1992; Gorsuch, 1983).

2.5.1.1 Principal Factor Method

Let R be the sample correlation matrix of a data matrix X ($n \times p$) and let us suppose that the factor loadings will be estimated from the correlation matrix R . The principle of principal factor method is that, the factors are to be selected in such a way that they can explain maximum variation in the original data set. Since R is used as input in factor analysis, it is a method of estimation the parameters in case of

$$s = \frac{1}{2}(p - q)^2 - \frac{1}{2}(p + q) > 0$$

Therefore, the method is similar to that of principal component analysis. The difference in both the techniques is that, in principal factor method the diagonal elements in R are replaced by the estimated communalities.

Let the estimated communality of i -th variable be

$$h_i^2 = \frac{T_{ij}T_{ik}}{T_{jk}}$$

Here the variables I_j and I_k are such that they are highly correlated with I_i .

As an estimate of h_i^2 one can use other values also. For example.

(a) h_i^2 is replaced by the square of the multiple correlation coefficient of i -th variable with other variables.

(b) h_i^2 is replaced by the largest correlation coefficient of i -th variable with any other variable .

It has already been mentioned that the amount of variance of f_j in the total variance of X is $V_j = \sum_{i=1}^p X_{ij}^2$ and the total variance of all factors is $V = \sum_{i=1}^p V_j$. Therefore, $(V_j/V) \times 100$ per cent variance of total variance is accounted for by j -th component.

The Total variance of X is p since the variables are standardized. Since we have $V(f_j) = \lambda_j$ hence j -th component explains $(\lambda_j / \sum \lambda_j) \times 100$ per cent variation of data set. The basic characteristic of principal factor method is that, the coefficients $\lambda_{11}, \lambda_{21}, \dots, \lambda_{p1}$ for component f_1 are to be estimated in such a way that in the total variance V that amount of variance of f_1 is maximum. This is done under the condition that total communality will be p and there will be $p(p-1)/2$ correlations.

Harman (1976) has proposed to estimate the values $\lambda_{11}, \lambda_{21}, \dots, \lambda_{p1}$ from reduced correlation matrix $R^* = R - v$. Let I_1 be the largest eigenvalue of R^* and g_1 elements $\lambda_{11}, \lambda_{21}, \dots, \lambda_{p1}$ is $T_1 = \sqrt{I_1}$. Here $V_1 = T_1' T_1$ is the amount of variance of f_1 in the variance of all variables.

In a similar way, the vector of weights of second factor is to be estimated. To estimate this vector R^* is adjusted by

$$R_1^* = R^* - T_1 T_1'$$

Under the conditions that $R_1^* = \sum_{j=2}^q T_j T_j'$ and $V_2 = T_2' T_2$ maximum, the estimate of the vector of factor weights T_2 is to be found out. Let the largest eigenvector of R_1^* be l_2 and its corresponding normalized eigenvector be g_2 so that $T_2 = \sqrt{l_2' g_2}$.

This process of estimation is continued until the factor weights for the factors $f_3, f_4, \dots, f_q$ are estimated.

In this process of estimation of factor weights, it is not pre-decided about the value of q . However, at each step of estimation the reduced correlation matrix is adjusted. Let us consider that m ($\leq q$) factors are selected. Then the adjusted reduced correlation matrix is

$$R_m^* = R^* - \sum_{j=1}^m T_j T_j^1 \quad (2.7)$$

If the values of R_m' are very small, further estimation of factor weights will not be fruitful. The diagonal elements of R' are replaced by estimated communalities, so in most cases these elements become less than 1. Due to the use of estimated communalities, the R' matrix is no longer positive semi-definite and some of its eigenvalues may be negative. The eigenvectors corresponding to negative eigenvalues may be imaginary. In such a situation the negative eigenvalue and its corresponding eigenvector is dropped from the analysis and the number of factors is so decided that the sum of eigenvalues is equal to the communality.

In principal factor method if the communalities are replaced by 1 or values near to 1 and if the factor numbers are equal to the number of variables, then the analytical results will be similar to those obtained by principal component analysis. However, the factors estimated by principal factor method are not scale invariant.

2.5.1.2 Unweighted Least Squares Method

The factor model to be analysed is

$$\Sigma = \Lambda\Lambda' + \psi \quad (2.8)$$

where $\psi = \psi_1 I$. The variances of the unique part of Z_1 's are heterogeneous.

If $\psi = \sigma^2 I$, then the factor model is identifiable and we have the following decomposition

$$(\Sigma - \lambda_1 I)I = [(\Lambda\Lambda' + \sigma^2 I) - \lambda_1 I]T_1 \quad (2.9)$$

where λ_1 and T_1 are the eigenvalue and eigenvector, respectively of Σ . In that case the signvector of $\Lambda\Lambda'$ is $\lambda'_1 = \lambda_1 - \sigma^2$ and the factor loadings are then given by

$$\Lambda = T [\text{diag} (\lambda'_1, \lambda'_2, \dots, \lambda'_q)]^{1/2}$$

where $T = [T_1, T_2, \dots, T_q]$. Thus, we have

$$\Lambda_1 = T_1 \sqrt{\lambda'_1}, i = 1, 2, \dots, q.$$

The sample analog of equation (5.6.1) is

$$(X'X - I, I)gi = (\hat{\Lambda}\hat{\Lambda}' - l'_i I)gi = 0 \quad (2.10)$$

where l_i and g_i are the eigenvalues of $X'X$ and the corresponding eigenvectors, respectively. Also, $l'_i = l_i - s^2, (i = 1, 2, \dots, q)$.

Hence s^2 is estimated by

$$s^2 = \frac{\sum_{i=q+1}^p 4}{p-q}$$

The factor loadings are then estimated by

$$\hat{\Lambda} = qL^{-1/2} \quad (2.11)$$

where $g = [g_1, g_2, \dots, g_q]$, $L' = \text{diag}(l'_1, l'_2, \dots, l'_q)$

This analysis is done assuming a homoscedastic variance matrix for unique part of the variables. This is similar to the principal component analysis. But, in the general factor model we assume a heteroscedastic variance matrix for the unique part of the variables. In ordinary principal component analysis the component loadings are inflated by a factor σ^2 , since the analysis does not adjust the variance matrix ψ .

The decomposition of the model factor for equation (2.8) are shown in the equation (2.9). Thus the minimization of residual sum of squares can be extended to a general factor model, where the residuals to be minimized are given by according to Joreskog (1977), the total derivative of U is

$$\begin{aligned} dU &= d[\text{Tr}(S-\Sigma)^2] \\ &= \text{Tr}[d(S-\Sigma)^2] \\ &= -2\text{Tr}[(S-\Sigma)d\Sigma] \\ &= -2\text{Tr}[(S-\Sigma)(\Lambda d\Lambda')] \end{aligned} \quad (2.12)$$

so that $\frac{\partial U}{\partial \Lambda} = -4 (\Sigma - S)\Lambda$

Now, $\frac{\partial U}{\partial \Lambda} = 0$ provides

$$\begin{aligned} (\hat{\Sigma} - S)\hat{\Lambda} &= [(\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}) - S]\hat{\Lambda} \\ &= \hat{\Lambda}(\hat{\Lambda}'\hat{\Lambda}) - (S\hat{\Lambda} - \hat{\Psi}\hat{\Lambda}) \\ &= 0 \end{aligned} \quad (2.13)$$

where $\hat{\Lambda}'\hat{\Lambda} = L$ is diagonal. From (5.6.4), we can write

$$(S - \hat{\Psi})\hat{\Lambda} = \hat{\Lambda}L \quad (2.14)$$

Here L and Λ are the matrices with eigenvalues of $(S - \psi)$ in the diagonal position and with the eigenvectors, respectively. Therefore, equation (2.14) represents a principal component analysis with covariance matrix $(S - \psi)$. This matrix is a corrected covariance matrix with corrected variances in the diagonal places. The non-zero eigenvalues of this corrected covariance matrix are q in number.

In the equation (2.14) ψ and Λ are two unknown quantities and as ψ is not known beforehand, ψ and Λ are estimated simultaneously. As a method of estimation principal component technique can be used. Let $\psi_{(1)}$ be the estimate of ψ using first q components. Then, we have

$$\hat{\psi}_{(1)} = \text{diag}(S - \hat{\Lambda}\hat{\Lambda}') \quad (2.15)$$

since ψ_i are uncorrelated, only the diagonal elements of S are replaced by estimated values. Once the initial estimate ψ_i is available, it is used in equation (2.14) and a second round estimate $\psi_{(2)}$ is to be obtained. The process of estimation is continued until the sequence of ψ and Λ converges to stable values. Note that the factor loadings estimated thus are not scale invariant.

2.5.1.3 Generalized Least Squares Method

The least squares method discussed in the previous section can be generalized to minimize the weighted least squares criterion, where the quantity to be minimize is

$$G = \text{Tr}[S^{-1}(S - \Sigma)S^{-1}(S - \Sigma)] \quad (2.16)$$

Here S^{-1} is the weight matrix [Joreskog (1977)]. The total derivative of G is

$$\begin{aligned} dG &= d[\text{Tr}(S^{-1}\Sigma - I)^2] \\ &= \text{Tr}[d(S^{-1}\Sigma - I)^2] \\ &= 2 \text{Tr}[(S^{-1}\Sigma - I)d(S^{-1}\Sigma - I)] \\ &= 2 \text{Tr}[(S^{-1}\Sigma - I)S^{-1} d\Sigma] \end{aligned} \quad (2.17)$$

If ψ in the model is assumed fixed,

$$\begin{aligned} d\Sigma &= d(\Lambda\Lambda' + \psi) \\ &= \Lambda(d\Lambda') + (d\Lambda)\Lambda' \\ &= 2\Lambda(d\Lambda') \end{aligned} \quad (2.18)$$

Thus, we, have

$$dg = 4 \text{ Tr } [(S^{-1}\Sigma^{-1}) S^{-1}\Lambda(d\Lambda')]$$

$$\frac{dG}{d\Lambda} = 4S^{-1}(\Sigma^{-1}S) S^{-1}\Lambda$$

Putting $\frac{dG}{d\Lambda} = 0$, we get

$$S^{-1}(\Sigma^{-1}S)S^{-1}\Lambda = 0$$

$$\Sigma S^{-1}\Lambda = \Lambda \quad (2.19)$$

This equation is equivalent to that obtained in (2.19). This equation is also equivalent to

$$(\psi^{1/2}S^{-1}\psi^{1/2})\psi^{-1/2}\Lambda = \psi^{-1/2}\Lambda(I + \Lambda'\psi^{-1}\Lambda)^{-1}$$

Since $\Sigma^{-1} = \psi^{-1} - \psi^{-1}\Lambda(I + \Lambda'\psi^{-1}\Lambda)^{-1}\Lambda'\psi^{-1}$. Therefore, Λ is obtained by interactive process. Here $\Lambda\psi^{-1}\Lambda$ can be taken as diagonal. The columns of $\psi^{-1/2}\Lambda$ are thus latent vectors of $\psi^{1/2}S^{-1}\psi^{1/2}$, and non-zero diagonals of $(I + \Lambda'\psi^{-1}\Lambda)^{-1}$ are the eigenvalues.

2.6 Akaike's Information Criterion

Akaike's information criterion, developed by Hirotugu Akaike under the name of "an information criterion" (AIC) in 1971 and proposed in (Akaike 1974), is a measure of the goodness of fit of an estimated statistical model. It is grounded in the concept of entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality and can be said to describe the

tradeoff between bias and variance in model construction, or loosely speaking that of accuracy and complexity of the model.

The AIC is not a test of the model in the sense of hypothesis testing; rather it is a test between models - a tool for model selection. Given a data set, several competing models may be ranked according to their AIC, with the one having the lowest AIC being the best. From the AIC value one may infer that e.g. the top three models are in a tie and the rest are far worse, but it would be arbitrary to assign a value above which a given model is 'rejected' (Burnham, Anderson, 1998)

In the general case, the AIC is

$$AIC = 2k - 2\ln(L) \quad (2.20)$$

where k is the number of parameters in the statistical model, and L is the maximized value of the likelihood function for the estimated model.

2.7 Normality Checking

A standard assumption is that the random shocks are normally distributed. This permits us to perform approximate t tests on coefficient significance at the estimation stage. One way to check the normality is to examine the histogram of the residuals. Another is to plot the residuals in a normal probability plot [see (Weisberg, 1980); (Liu and Hudak, 1986)]. Both of the procedure provides a helpful graphical presentation of the data; they do not, however, provide any formal test for normality. There various formal test for normality. The most recent and efficient test for normality of the fitted residuals is described as follows.

2.7.1 Jarque-Bera Normality Test

In statistical analysis, the test for normality is an important problem. Many statistical tests have been proposed by many authors to find out whether a sample is drawn from a normal distribution or not. Here we discuss the Jarque-Bera (JB) test which is based on the classical measures of skewness and kurtosis, a test of normality proposed by Carlos Jarque and Anil K. Bera in 1980 and complete paper of the JB Test was published in 1987 which is useful tool in Econometrics Jarque (1987). In this section we discuss JB normality tests for the following null and alternative hypothesis:

H_0 : The data are sampled from a normal distribution

H_1 : The data are not sampled from a normal distribution

Let us consider the usual regression model

$$Y = X\beta + \varepsilon \quad (2.21)$$

Where $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2 I$. In addition, ε normally distributed the third and fourth moments for in ε elements are as

$$\mu_3 = E(\varepsilon_i^3) = 0$$

$$\text{And } \mu_4 = E(\varepsilon_i^4) = 3\sigma^4$$

A large number of tests for normality are based on how far estimates of the third and fourth order moments, $\tilde{\mu}_3$ and $\tilde{\mu}_4$ deviate from 0 and $3\sigma^4$ respectively,

where $\tilde{\sigma}^2$ is an estimate of $E(\varepsilon_i^2) = \sigma^2$. In this regards measures of skewness and kurtosis are

$$\sqrt{b_1} = \frac{\mu_3}{\sigma^3}$$

$$b_2 = \frac{\mu_4}{\sigma^4}$$

Relative to the normal distribution, a measure of excess kurtosis is given by

$$(b_2 - 3) = \frac{\mu_4 - 3\sigma^4}{\sigma^4}$$

The JB test is a joint test of whether or not estimate of $\sqrt{b_1}$ and/or $(b_2 - 3)$ are significantly differ from 0. Since the error vector ε is unobservable, estimates of

$\sqrt{b_1}$ and $(b_2 - 3)$ are biased on least square residuals. Specially, if $\hat{\varepsilon}_i = (y_i - x_i')$

where $b = (x'x)^{-1} x'y$ and x' is the t-th row of the $X_{(T \times K)}$, then

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_{i=1}^T \hat{\varepsilon}_i^2, \quad \tilde{\mu}_3 = \frac{1}{T} \sum_{i=1}^T \hat{\varepsilon}_i^3, \quad \tilde{\mu}_4 = \frac{1}{T} \sum_{i=1}^T \hat{\varepsilon}_i^4$$

$$\sqrt{\tilde{b}_1} = \frac{\tilde{\mu}_3}{\tilde{\sigma}^3}, \quad (\tilde{b}_2 - 3) = \frac{\tilde{\mu}_4 - 3\tilde{\sigma}^4}{\tilde{\sigma}^4}$$

Under H_0 errors are normally distributed, JB statistic is as

$$JB = T \left[\frac{\tilde{\mu}_3^2}{6\tilde{\sigma}^6} + \frac{\tilde{\mu}_4 - 3\tilde{\sigma}^4}{\tilde{\sigma}^4} \right] \quad (2.22)$$

which asymptotically follows the chi-squared distribution with two degrees of freedom.

2.8 Sampling Design

The sampling design is a fundamental part of data collection for statistically based decision making. A well-developed sampling design plays an important role in ensuring that data are sufficient to draw the conclusions needed. The goal of sampling is that the selected data should be representative of the population and it will also be unbiased so that the data have the integrity. There are number of issues that researcher should be aware of when selecting the samples. This includes determining:

- a) The appropriate type and source of data which permit investigator to adequately answer the stated research questions,
- b) Suitable procedures in order to obtain a representative sample.

To draw the sample for the present study we have used the Simple random sampling.

2.8.1 Simple Random Sampling

Simple random sampling (SRS) is one of the most popular types of random or probability sampling. In this technique, each member of the population has an equal chance of being selected as subject. The entire process of sampling is done in a single step with each subject selected independently of the other members of the population. Conceptually, Simple random sampling is the simplest of the probability sampling techniques used in case of homogeneous sampling units. It requires a complete sampling frame, which may not be available or feasible to construct for large populations.

Even if there may have some differences among the greater districts all of them are homogeneous in per hector production. So these 5 greater districts have been consecutively numbered from 1 to 5 and using random number from random number table we have selected three districts. The selected districts are Rajshahi, Rangpur and Dinajpur.

2.9 Study Area

The districts Rajshahi, Rangpur and Dinajpur marked by red circle in the following map are the area selected for the study.



2.10 Sources of Data and Construction of Variables

For the proposed study the production data of HYV Boro rice have been used from three northern districts, namely Rajshahi (Yhraj), Rangpur (Yhran), and Dinajpur (Yhdin). The HYV Boro rice has three different stages in its production period like vegetative stage, growing stage and reproductive stage. The climatic variables rainfall (r), temperature (t), wind speed (w), humidity (h) and cloud coverage (cc) for these three stages explained as 1 for vegetative stage, 2 for growing /milk stage and 3 for reproductive stage in each of the variables. Another climatic variable Co2 emission (co) has used same for the three stages. The two non climatic variables have been used on the fertilizer use (fu) and tractor use (tu) from three different districts, namely Rajshahi (fur & tur), Rangpur (furan & turan) and Dinajpur (fudin & tudin) respectively. The data have been used for all the variables from the year 1971 to 2007.

The production data from the project 'Early Warning and Food Information System (EWFIS)' conducted by the Food Planning Monitoring Unit (FPMU) of the Ministry of Food, government of Bangladesh and Bangladesh Bureau of Statistics.

The climatic data (Monthly Minimum Temperature, Monthly Rainfall, Monthly Humidity, Monthly Cloud Coverage, Monthly Wind Speed) and a non climatic data Fertilizer Use are jointly collected from Bangladesh Metrological Department (BMD), The Year Book of Agriculture Statistics of Bangladesh

published by Bangladesh Bureau of Statistics (BBS), Planning Division, Ministry of Planning, People's Republic of Bangladesh

The variables quantity of production of HYV Boro rice by the three different districts is labeled in italic given in parenthesis next to the district name. The variables are measured in metric tons per hectare (mt/hect.), temperature in Celsius ($^{\circ}\text{C}$) rainfall in millimeter (mm) wind speed in meter per second (m/sec), cloud coverage in octas humidity in percentage (%) fertilizer use in Kilogram per hectare (Kg/hect.) and CO_2 emission in metric tons of carbon per capita tractor use in number per hectare.

2.11 Statistical Software

SHAZAM (version 10), a popular and comprehensive software package for econometric and statistical computing was created in 1977 by Ken White have used for data analysis. Writing and composing of the whole thesis have been done by MS Word 2003 editor.

Chapter Three

Trend Analysis

3.1 Introduction

Different trend models and their estimation of parameters, the empirical results obtained from trend analysis and different forecasting errors are presented in this chapter.

3.2 Trend Models

A variety of trend models are available in the literature for analyzing the trend of agricultural production data. Among them linear trend model, compound trend model and quadratic trend model are commonly used in the literature (Karmokar and Imon 2008).

Let y_{ij} ($i = \text{yhraj, yhran, yhdin}; j = 1, 2, \dots, n$) denote the HYV Boro production of rice, and t be the year of production then the linear trend model becomes

$$y_{ij} = a + bt + \epsilon_{ij} \quad (3.1)$$

The quadratic trend model can be defined as

$$y_{ij} = a + bt + ct^2 + \epsilon_{ij} \quad (3.2)$$

The compound trend model takes the form

$$y_{ij} = ab^t \epsilon_{ij} \quad (3.3)$$

Taking logarithmic on both sides of equation (3.3) we get,

$$y_{ij}^* = a^* + b^* t + \epsilon_{ij}^* \quad (3.4)$$

where $y_{ij}^* = \ln y_{ij}$, $a^* = \ln a$, $b^* = \ln b$, $\epsilon_{ij}^* = \ln \epsilon_{ij}$. For all of the above models we have included the error term ϵ_{ij} and the basic assumption about the error term ϵ_{ij} is that it follows a Normal distribution with zero mean and constant variance.

3.3 Estimation of Parameters

To estimate the parameter of model (3.1), we have to minimize

$$\varphi = \sum (y_{ij} - \hat{y}_i)^2 = \sum (y_{ij} - \hat{a} - \hat{b}t)^2 \text{ according to the principle of OLS. The estimates}$$

of a and b are given in below respectively

$$\hat{a} = \bar{y}_i - \hat{b}\bar{t} \quad \text{and}$$

$$\hat{b} = \frac{\sum y_i t - \sum y_i \sum t}{\sum t^2 - \frac{(\sum t)^2}{n}}$$

Then the estimated liner trend model becomes $\hat{y}_i = \hat{a} + \hat{b}t$.

Similarly applying OLS, we get the estimates of parameters of equation (3.2) and the estimates are

$$\hat{c} = \frac{\sum t^2 y_i - (\sum t^2)(\sum y_i)}{\sum t^4 - \frac{(\sum t^2)^2}{n}}, \quad \hat{b} = \frac{\sum y_i t}{\sum t^2} \quad \text{and} \quad \hat{a} = \bar{y}_i - \frac{\hat{c} \sum t^2}{n}$$

Then the estimated quadratic trend model becomes $\hat{y}_i = \hat{a} + \hat{b}t + \hat{c}t^2$.

Similarly, from equation (3.3) we get the estimated parameters are as

$$b = \text{Anti log} \left(\frac{\sum t \log y_i}{\sum t^2} \right) \text{ and } a = \text{Anti log} \left(\frac{\sum \log y_i}{n} \right)$$

Hence the estimated quadratic trend model becomes $\hat{y}_i = \hat{a}\hat{b}'$.

3.4 Results of Trend Models

Table 3.1 reports the summary statistics of variables for the production of high yielding variety of Boro rice, yhraj, yhran and yhdin respectively for the districts Rajshahi, Rangpur and Dinajpur respectively.

Table 3.1 Descriptive Statistics for HYV Boro production

Variable	Min	Max	Mean	SE	CV (%)	Skew	Kur
yhraj	2.05	3.81	2.848	0.466	16.34	0.342	2.327
yhran	1.68	3.91	2.675	0.514	19.20	0.285	3.256
yhdin	2.03	3.81	2.715	0.420	15.45	0.917	3.859

From the above the table we see that the maximum production of HYV Boro is 3.81 metric tone per hector and the minimum HYV Boro production is 2.05 metric tone per hector for Rajshahi. The maximum production of HYV Boro is 3.91 metric tones per hector and the minimum HYV Boro production is 1.68 metric tones per hector for Rangpur and the maximum production of HYV Boro is 3.81 metric tones per hector and the minimum HYV Boro production is 2.03 metric tones per hector for Dinajpur for the period 1971 to 2007. The average production of HYV Boro rice for Rajshahi, Rangpur and Dinajpur are 2.8483 m.ton/hector, 2.6746 m.ton/hector and 2.7150 m.ton/hector respectively thought the study period. The standard deviation for Rajshahi is 0.46551 m.ton, for

Rangpur is 0.51369 m.ton and for Dinajpur is 0.41970 m.ton. The coefficient of variation for Rajshahi, Rangpur and Dinajpur are obtained as 16.34%, 19.20% and 15.45% respectively. According to the coefficient of variation, the production of HYV Boro rice for Rangpur shows the maximum variation and Dinajpur shows the minimum variation. The skewness of the production of HYV Boro rice for Rajshahi, Rangpur and Dinajpur are 0.342, 0.285 and 0.917 respectively. This coefficients are positive indicate that the data set are positively skewed. The kurtosis values for Rajshahi, Rangpur and Dinajpur are obtained as 2.327, 3.256 and 3.859 respectively.

The time series (TS) plot of the production of HYV Boro rice of the selected districts for the year from 1971 to 2007 are shown in the following figures Fig. 3.1-3.3.

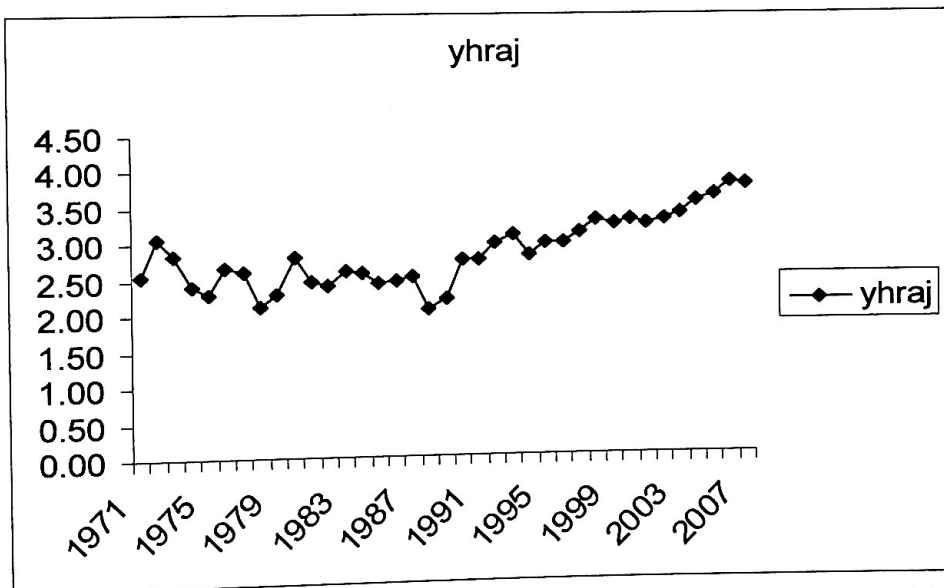


Fig 3.1. TS Plot of HYV Boro production for the district Rajshahi

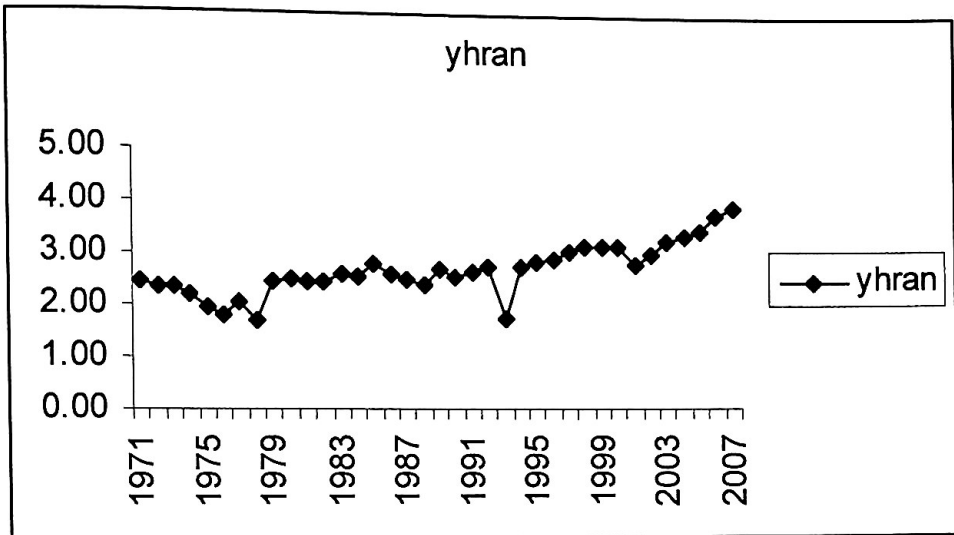


Fig 3.2. TS Plot of HYV Boro production for the district Rangpur

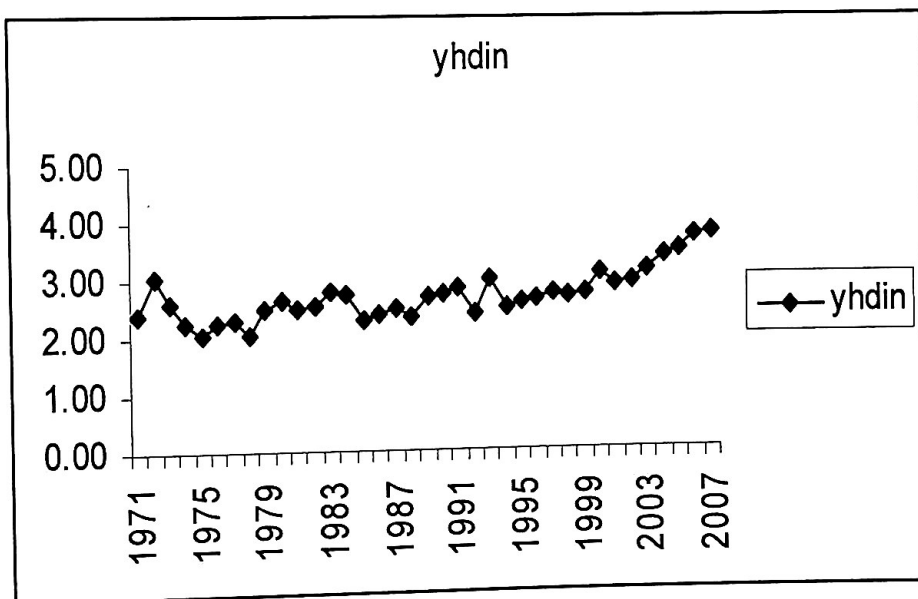


Fig 3.3. TS Plot of HYV Boro production for the district Dinajpur

The Figures (Fig. 3.1 - 3.3) shows that the production of HYV Boro rice for the selected districts have upward trend. The non parametric statistical test like,

Mann-Kendall test and Spearman test have been done to check the validity of the trend of the study variables and the results are show in the following tables.

Table 3.2 Mann-Kendal test statistic values for the period 1971 to 2007

Variable	Range (Years)	Mann-Kendal statistic t_m	Value at 95% level $\pm r_m$	Trend Exist
yhraj	37	-0.8431	0.338	Yes
yhran	37	-0.9477	0.338	Yes
yhdin	37	-0.9346	0.338	Yes

Table 3.3 Spearman test statistic values for the period 1971 to 2007

Variable	Range (Years)	Spearman statistic t_s	Value at 95% level $\pm r_s$	Trend Exist
yhraj	37	-2.1130	0.834	Yes
yhran	37	-2.7722	0.834	Yes
yhdin	37	-2.1463	0.834	Yes

From Table 3.1 we see that all the values of t_m statistics fall outside of the 95% confidence interval. Also from Table 3.2, all the value of t_s statistics fall outsides of 95% confidence interval. So we may reject the null hypothesis (H_0 : There is no trend in the data set) for both the tests. Therefore, the trend exists for all the variables used in the data set for the districts Rajshahi, Rangpur and Dinajpur.

Since we are using the time series data it is necessary to check the time series properties (discussed in Chapter 2, Section 2.3) before applying the trend models. The lag choice is a matter of fact for different test in time series analysis. Akaike's

information criterion (AIC) is one of the simple ways to selection the lag length. Because of the small sample size, we are strict to the seventh lag for the AIC test. AIC values with trend and without trend are given in Table 3.4 and Table 3.5. The test values (τ_μ) for with trend and without trend are reported in the Table 3.6 and Table 3.7.

Table 3.4 AIC values of variables with trend for selected districts

Variable	Lag-1	Lag-2	Lag-3	Lag-4	Lag-5	Lag-6	Lag-7	Min(AIC) (Lag Choice)
yhraj	-2.540	-3.312	-3.226	-3.157	-3.162	-2.973	-3.167	-3.312 (Lag 2)
yhran	-2.496	-2.434	-2.391	-2.299	-2.204	-2.103	-2.072	-2.496 (Lag 1)
yhdin	-2.912	-2.837	-2.766	-2.831	-2.724	-3.135	-3.029	-3.135 (Lag 6)

Table 3.5 AIC values of variables without trend for selected districts

Variable	Lag-1	Lag-2	Lag-3	Lag-4	Lag-5	Lag-6	Lag-7	Min(AIC) (Lag Choice)
yhraj	-2.757	-3.070	-3.033	-3.070	-2.981	-2.973	-3.064	-3.07 (Lag2)
yhran	-2.310	-2.237	-2.161	-2.132	-2.133	-2.020	-2.076	-2.31 (Lag1)
yhdin	-2.637	-2.649	-2.706	-2.810	-2.736	-3.179	-3.079	-3.179 (Lag 6)

Table 3.6 Test and Critical values of variables with trend for selected districts

Variable	Critical value	τ_{μ} – Test values						
		Lag-1	Lag-2	Lag-3	Lag-4	Lag-5	Lag-6	Lag-7
yhraj	4.03	5.2220	4.3400	3.3163	2.5355	1.9392	4.3040	2.6646
yhran	4.03	3.3979	3.4770	3.7390	3.1628	2.6369	2.4217	2.6048
yhdin	4.03	4.3226	3.2915	2.2062	1.8471	1.6726	2.7757	1.5916

Table 3.7 Test and Critical values of variables without trend for selected districts

Variable	Critical value	τ_{μ} – Test values						
		Lag-1	Lag-2	Lag-3	Lag-4	Lag-5	Lag-6	Lag-7
yhraj	3.78	0.40235	1.0808	0.86397	1.5968	1.0212	1.6806	1.9717
yhran	3.78	0.67031	0.79442	0.82391	1.2580	2.0805	1.7521	3.2203
yhdin	3.78	0.28270	0.70185	1.4432	1.6059	1.8642	4.0126	2.2749

The unit root test is applied for the series of data for Rajshahi, Rangpur and Dinajpur in levels to test the stationarity. The Dickey-Fuller test values and the associated 10% critical values are used reported in Table 3.8.

Table 3.8 Observed value of the τ_{μ} – statistic with the critical value

Variable	Critical value of test	Test Value (τ_{μ})
yhraj	4.03	4.340
yhran	4.03	3.398
yhdin	4.03	2.776

From the above table, the test statistic value for variable yhraj falls in the rejection region but the test statistic value for yhdin and yhran fall in the acceptance region hence we may accept our null hypothesis for yhdin and yhran.

The ADF test (discussed in Section 2.3.2.2.1) test statistics values are done for all the data set of selected districts and that are given in the Table 3.9

Table 3.9 Repeated Dickey-Fuller test for order of integration of the variables

Variable	Δy	$\Delta^2 y$	Critical value	Order of Integration
yhraj	6.2301	8.5855	4.03	$I(0)$
yhran	4.6617	5.9613	4.03	$I(0)$
yhdin	12.820	26.223	4.03	$I(0)$

The ADF statistics indicates that the null hypothesis (H_0 : nonstationary of data) is rejected for all the variables yhraj, yhran and yhdin. Therefore, the data of each of the series are stationary. Since all the values of 1st difference and 2nd difference in Table 3.9 greater than critical values, it indicates that all the variables are stationary in level.

District-wise OLS estimate of parameters, standard error of estimates, T-Ratio and corresponding p-value of the estimates for linear, compound and quadratic trend models are shown in Table 3.10 to Table 3.12.

Table 3.10 OLS estimates of parameters for Linear Trend model

Model	Variable	Co-efficient	Standard Error	T Ratio	P-Value
yhraj	t	0.028	0.005	5.495	0.000
	Constant	2.285	0.101	22.600	0.000
yhran	t	0.032	0.005	6.763	0.000
	Constant	2.015	0.963	20.930	0.000
yhdin	t	0.0193	0.004	4.458	0.000
	Constant	2.292	0.867	26.440	0.000

Table 3.11 OLS estimates of parameters for Compound Trend model

Model	Variable	Co-efficient	Standard Error	T Ratio	P-Value
yhraj	t	0.010	0.002	5.120	0.000
	Constant	0.839	0.038	22.040	0.000
yhran	t	0.013	0.002	5.938	0.000
	Constant	0.715	0.042	16.880	0.000
yhdin	t	0.007	0.002	4.457	0.000
	Constant	0.831	0.033	25.140	0.000

Table 3.12 OLS estimates of parameters for Quadratic Trend model

Model	Variable	Co-efficient	Standard Error	T Ratio	P-Value
yhraj	t	-0.050	0.015	-3.294	0.002
	t ²	0.002	0.000	5.258	0.000
	Constant	2.754	0.116	23.670	0.000
yhnan	t	0.004	0.019	0.187	0.853
	t ²	0.001	0.001	1.537	0.135
	Constant	2.188	0.147	14.900	0.000
yhdin	t	-0.020	0.017	-1.221	0.231
	t ²	0.001	0.000	2.460	0.020
	Constant	2.528	0.126	20.150	0.000

The values of coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2), Jarque-Bera statistic (JB) and Durbin-Watson statistic (DW) by OLS method for Linear Trend model, Compound Trend model and Quadratic Trend model for the selected districts are reported in the Table 3.13-Table 3.15.

Table 3.13 R^2 , adjusted R^2 , JB and DW statistic by OLS for Linear Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.4855	0.469	1.400	0.496	0.799
yhran	0.5884	0.576	39.706	0.000	1.390
yhdin	0.3831	0.364	3.882	0.144	1.373

Table 3.14 R^2 , adjusted R^2 , JB and DW by OLS for Compound Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.4503	0.433	3.249	0.197	0.869
yhran	0.5242	0.509	48.407	0.000	1.465
yhdin	0.3831	0.364	2.074	0.355	1.403

Table 3.15 R^2 , adjusted R^2 , JB and DW by OLS for Quadratic Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.7281	0.711	0.867	0.648	1.494
yhran	0.6175	0.593	38.452	0.000	1.484
yhdin	0.4839	0.451	1.049	0.592	1.624

From Table 3.13 to 3.15 it is seen that the DW values of all methods and for all districts are less than 2 indicate that the data are affected by autocorrelation problem. To avoid this problem Cochrane-Orcutt estimation (discussed in chapter two, section 2.4.2.2) technique have been applied. The Cochrane and Orcutt estimates of parameters for three methods and for the selected districts are computed and reported in the Table 3.16-Table 3.18. Method-wise R^2 , adjusted R^2 , JB statistic and DW statistic are given in Table 3.19-Table 3.21.

Table 3.16 Cochran and Orcutt estimates by Linear Trend model

Model	Variable	Co-efficient	Standard Error	T Ratio	P-Value
yhraj	t	0.028	0.009	3.171	0.003
	Constant	2.297	0.183	12.580	0.000
yhran	t	0.032	0.006	5.239	0.000
	Constant	2.027	0.124	16.390	0.000
yhjin	t	0.020	0.006	3.526	0.001
	Constant	2.286	0.114	20.000	0.000

Table 3.17 Cochran and Orcutt estimates by Compound Trend model

Model	Variable	Co-efficient	Standard Error	T Ratio	P-Value
yhraj	t	0.010	0.003	3.041	0.005
	Constant	0.844	0.066	12.860	0.000
yhran	t	0.013	0.003	4.724	0.000
	Constant	0.720	0.053	13.630	0.000
yhjin	t	0.008	0.021	3.531	0.001
	Constant	0.830	0.043	19.270	0.000

Table 3.18 Cochran and Orcutt estimates by Quadratic Trend model

Mode l	Variable	Co- efficient	Standard Error	T Ratio	P-Value
yhraj	t	-0.048	0.019	-2.568	0.015
	t^2	0.002	0.001	4.178	0.000
	Constant	2.737	0.143	19.170	0.000
yhran	t	0.004	0.024	0.026	0.984
	t^2	0.001	0.001	1.375	0.179
	Constant	2.188	0.147	14.900	0.000
yhdin	t	-0.021	0.019	-1.080	0.289
	t^2	0.001	0.001	2.171	0.038
	Constant	2.528	0.145	17.390	0.000

Table 3.19 R^2 , adjusted R^2 , JB and DW values by Cochran-Orcutt method for
Linear Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.664	0.6535	2.561	0.277	1.7207
yhran	0.6198	0.6079	1.987	0.370	1.9577
yhdin	0.4308	0.413	2.073	0.355	1.8666

Table 3. 20 Values of R^2 , adjusted R^2 , JB by Cochrane-Orcutt method for Compound Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.6202	0.6083	1.252	0.535	1.7200
yhran	0.5526	0.5386	0.197	0.906	1.9780
yhdin	0.4290	0.4112	1.401	0.496	1.8774

Table 3. 21 Values of R^2 , adjusted R^2 , JB by Cochrane-Orcutt method for Quadratic Trend model

Component	R^2	Adjusted R^2	JB		DW
			Statistic	P value	
yhraj	0.7445	0.7281	0.963	0.618	1.8152
yhran	0.6411	0.6180	0.578	0.749	2.0104
yhdin	0.4991	0.4668	1.776	0.411	1.9352

It is clear from Table 3.19–Table 3.21 that the DW statistic values for all the models is near to 2 indicate that the data now free from autocorrelation problem and the respective R^2 and adjusted R^2 values improves than OLS method. The JB test shows the normality of the data set.

Both deterministic and stochastic measures can be used to evaluate performance of a forecasting technique. In this study, the MSE and AME are used as a deterministic measure for the prediction performance of the models. Table 3.1.32 to 3.1.34 show the method wise forecasted HYV Boro rice production of the districts Rajshahi, Rangpur and Dinajpur. Method-wise Mean Square Error (MSE)

and Absolute Mean Error (AME) have estimated to check the forecasting performance of the selected trend methods. The MSE and AME results for different trend models are also displayed in the Tables 3.19 –3.21.

Table 3.22 Method wise forecasted HYV Boro rice production of Rajshahi

Year of Production	Actual Production (y _{hraj})	Forecasted by Linear Trend Model	Forecasted by Compound Trend Model	Forecasted by Quadratic Trend Model
2004	3.550	3.260	3.242	3.575
2005	3.640	3.290	3.274	3.675
2006	3.810	3.310	3.306	3.779
2007	3.780	3.340	3.339	3.887
2008	-	3.370	3.372	4.000
2009	-	3.400	3.405	4.117
2010	-	3.430	3.438	4.238
2011	-	3.460	3.472	4.363
2012	-	3.480	3.506	4.493
2013	-	3.510	3.540	4.627
2014	-	3.540	3.575	4.765
2015	-	3.570	3.610	4.908
2016	-	3.600	3.646	5.055
2017	-	3.630	3.681	5.206
2018	-	3.650	3.718	5.362
2019	-	3.680	3.754	5.522
2020	-	3.710	3.791	5.686
2021	-	3.740	3.828	5.854
MSE		0.054	0.042	0.043
AME		0.210	0.141	0.164

Table 3.23 Method wise forecasted HYV Boro rice production of Rangpur

Year of Production	Actual Production (yhran)	Forecasted by Linear Trend Model	Forecasted by Compound Trend Model	Forecasted by Quadratic Trend Model
2004	3.340	3.119	3.131	3.275
2005	3.460	3.152	3.170	3.338
2006	3.770	3.184	3.210	3.403
2007	3.910	3.216	3.250	3.470
2008	-	3.248	3.290	3.538
2009	-	3.280	3.331	3.608
2010	-	3.312	3.373	3.680
2011	-	3.344	3.415	3.754
2012	-	3.376	3.457	3.830
2013	-	3.409	3.500	3.907
2014	-	3.441	3.544	3.986
2015	-	3.473	3.588	4.067
2016	-	3.505	3.633	4.150
2017	-	3.537	3.678	4.235
2018	-	3.569	3.724	4.321
2019	-	3.601	3.770	4.410
2020	-	3.633	3.817	4.500
2021	-	3.666	3.865	4.592
MSE		0.071	0.070	0.095
AME		0.189	0.189	0.271

Table 3.24 Method wise forecasted HYV Boro rice production of Dinajpur

Year of Production	Actual Production (yhdin)	Forecasted by Linear Trend Model	Forecasted by Compound Trend Model	Forecasted by Quadratic Trend Model
2004	3.39	2.965	2.963	3.156
2005	3.52	2.985	2.985	3.215
2006	3.75	3.005	3.008	3.276
2007	3.81	3.025	3.031	3.339
2008	-	3.044	3.054	3.405
2009	-	3.064	3.077	3.473
2010	-	3.084	3.100	3.543
2011	-	3.104	3.123	3.616
2012	-	3.124	3.147	3.691
2013	-	3.144	3.171	3.768
2014	-	3.164	3.195	3.848
2015	-	3.184	3.219	3.930
2016	-	3.204	3.243	4.014
2017	-	3.224	3.268	4.100
2018	-	3.244	3.293	4.189
2019	-	3.264	3.318	4.280
2020	-	3.284	3.343	4.373
2021	-	3.304	3.368	4.469
MSE		0.056	0.053	0.051
AME		0.184	0.183	0.184

We see from Table 3.22 that the MSE's are 0.054, 0.042 and 0.043 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.210, 0.141 and 0.164. Hence the Compound Trend model possesses the lowest MSE

and AME values for HYV Boro rice production of Rajshahi district. Quadratic Trend model possesses next lowest MSE and AME values while Linear Trend model possesses the highest MSE and AME values among the three methods for the data of Rajshahi district.

In Table 3.23, the MSE's are 0.071, 0.070 and 0.095 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.189, 0.189 and 0.271. Here also the Compound Trend model possesses the lowest MSE and AME values for HYV Boro rice production of Rajshahi district but the Linear Trend model possesses the highest MSE and AME values among the three methods for the data of Rajshahi district.

For the HYV Boro production of Dinajpur, the MSE's are 0.056, 0.053 and 0.051 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.184, 0.183 and 0.184. Therefore Quadratic Trend model has minimum MSE value than Linear and Compound Trend model but considering the AME's the Compound Trend have the lowest forecasting error for Dinajpur district. The other two methods show the same AME values for this district. The growth rate for compound trend model is 1.009 for Rajshahi, is 1.0124 for Rangpur and is 1.0075 for Dinapur.

Chapter Four

Factor Analysis

4.1 Introduction

The Model for Factor Analysis, Factor Loading Estimation, Eigen values, Scree Plot, Kaiser-Meyer-Olkin Statistics and Bartlett's Test of Sphericity and corresponding results with discussion are discussed in this chapter.

4.2 The Model for Factor Analysis

The common factor analytic model assumes that a variable consists of common and unique parts. Let us suppose the factor model as,

$$X = \Lambda F + U \quad (4.1)$$

$$\text{where } X = [X_1, X_2, \dots, X_p]', F = [f_1, f_2, \dots, f_q]', U = [e_1, e_2, \dots, e_p]'$$

X is the data vector, F is the vector of common factor and U is the unique factor

$$\text{and } \Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1q} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2q} \\ \dots & \dots & \dots & \dots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pq} \end{bmatrix} \text{ is a } (p \times q) \text{ matrix of factor loadings, where}$$

λ_{ij} 's ($i=1,2,\dots,p; j=1,2,\dots,q$) are unknown.

It is assumed that $e_1, e_2, \dots, e_p$ are mutually independent and they are also independent of the elements in F . That is $\text{cov}(U, F') = 0$. Also, it is assumed that

$$E[UU'] = \begin{bmatrix} \psi_1 & 0 & 0 & \dots & 0 \\ 0 & \psi_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \psi_p \end{bmatrix}, \quad E[F] = 0, \quad V(F) = I$$

On the basis of the assumptions of the model (4.1), we have

$$\Sigma = \Lambda\Lambda' + \Psi \quad (4.2)$$

However, if $\text{cov}[F] = \begin{bmatrix} 1 & & & & \\ \phi_{21} & 1 & & & \\ \phi_{31} & \phi_{32} & 1 & & \\ \dots & \dots & \dots & \dots & \\ \phi_{q1} & \phi_{q2} & \dots & \dots & 1 \end{bmatrix} = \phi$, we have

$$\Sigma = \Lambda\phi\Lambda' + \Psi \quad (4.3)$$

Here, Σ is the variance-covariance matrix of X . From the factor model (4.1), we can write

$$X_i = \sum_{j=1}^q \lambda_{ij} f_j + e_i, \quad i=1, 2, \dots, p \quad (4.4)$$

The equation (4.1) represents the factor-pattern. Thus the factors can be estimated from the observable random variables. The estimators of factors are the linear combinations of observable random variables, where the general form of the j -th factor is

$$f_j = \sum_{i=1}^p W_{ij} X_i = W_{1j} X_1 + W_{2j} X_2 + \dots + W_{pj} X_p \quad (4.5)$$

4.3 Estimation of Factor Loadings

We can write from model (4.5),

$$V(X_i) = V\left[\sum_{j=1}^q \lambda_{ij} f_j\right] + V(e_i)$$

$$\sigma_i^2 = \sum_{j=1}^q \lambda_{ij}^2 + \psi_i, \text{ since } V(f_j) = 1 \text{ and } \text{cov}(f_i, f_j) = 0$$

$$\text{cov}(X_i, X_j) = \text{cov}\left[\sum_{j=1}^q \lambda_{ij} f_j + e_i, \sum_{j=1}^q \lambda_{j'j} f_j + e_{j'}\right]$$

$$\sigma_{ij} = \sum_{j=1}^q \lambda_{ij} \lambda_{j'j}$$

$$\therefore V(X) = \Sigma$$

$$= \begin{bmatrix} \sum \lambda_{1j}^2 & \sum \lambda_{1j} \lambda_{2j} & \sum \lambda_{1j} \lambda_{3j} & \dots & \sum \lambda_{1j} \lambda_{pj} \\ \sum \lambda_{1j} \lambda_{2j} & \sum \lambda_{2j}^2 & \sum \lambda_{2j} \lambda_{3j} & \dots & \sum \lambda_{2j} \lambda_{pj} \\ \dots & \dots & \dots & \dots & \dots \\ \sum \lambda_{pj} \lambda_{1j} & \sum \lambda_{pj} \lambda_{2j} & \dots & \dots & \sum \lambda_{pj}^2 \end{bmatrix} + \begin{bmatrix} \psi_1 & 0 & 0 & 0 & 0 \\ 0 & \psi_2 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \psi_p \end{bmatrix}$$

$$\Rightarrow \Sigma \Lambda \Lambda' + \Psi \quad (4.6)$$

Here the diagonal elements of $\Lambda \Lambda'$ are $\sigma_i^2 - \psi_i = \sum_{j=1}^q \lambda_{ij}^2 = h_i^2, i = 1, 2, \dots, p$.

Here $h_i^2 = \sum_{j=1}^q \lambda_{ij}^2$ is called communality. From model (4.5) it is clear that the variable X_i has two uncorrelated parts, one part is a function of common factors, i.e., $h_i^2 = \sum_{j=1}^q \lambda_{ij} f_j$, that is part of X_i which is common to other $(p-1)$ variables.

Another part of X_i is e_i and this part of X_i is unique. Also, we have seen $V(X_i) = V(\sum \lambda_{ij} f_j) + V(e_i)$

Here $V(\sum \lambda_{ij} f_j)$ is the common variance and $V(e_i)$ is the unique variance of X_i . Here the common variance of X_i is called the communality of the variable. Therefore,

the communality of X_i is the portion of total variance of X_i which is accounted for by the common factors. The unique variance of X_i is ψ_i and it is called the uniqueness of the variable. This ψ_i reflects that portion of variance of X_i which is not accounted for by the common factors.

The total variance of X is $Tr \Sigma$. From model (4.6), we have

$$Tr \Sigma = \sum_i^p \sum_j^q \lambda_{ij}^2 + \sum_{j=1}^p \psi_i \quad (4.7)$$

Therefore, the total communality is $V = \sum_{i=1}^p h_i^2$

The contribution of factor f_j to the total variance of X is, say V_j , where

$$V_j = \sum_{i=1}^p \lambda_{ij}^2 \Rightarrow V = \sum_{j=1}^q V_j \quad (4.8)$$

Therefore, the portion of variance of all variables which is accounted for by f_j is V_j/V .

Again, we have

$$\text{cov} [X_i, f_j] = E \left[\sum_{j=1}^q \lambda_{ij} f_j, f_j \right], \because E(X_i) = 0 \text{ and } E(f_j) = 0$$

$$= \lambda_{ij}$$

$$\Rightarrow \text{cov} (X, F') = \Lambda \quad (4.9)$$

Thus the Λ matrix is the matrix of covariances of i -th variables and j -th factor. If X_i 's are standardized, Σ matrix is transform to correlation matrix. In that case A_{ij} is the correlation coefficient of i -th variables with j -th factor. Therefore, the

estimation of factor loadings is nothing but the estimation of correlation coefficient of i -th variables with j -th factor.

The elements in Λ matrix are to be estimated under the condition that

$$\sigma_i^2 = \sum_{j=1}^q \lambda_{ij}^2 + \psi_i \quad (4.10)$$

The normal equation to estimate Λ is

$$\Sigma = \Lambda \Lambda' + \Psi$$

In the parameter matrix Λ there are pq unknown elements. In the left-hand side of the equation, there are $\frac{1}{2}p(p+1)$ different variances and covariances and hence there are $\frac{1}{2}p(p+1)$ equations in (4.6). But to estimate the parameters the equations must be identifiable. The equations are identifiable if the number of parameters is less than the number of equations. Therefore, $pq+p$ should be less than $\frac{1}{2}p(p+1)$, or $q < \frac{1}{2}p(p+1)$. This means that q should be less than p . However, the smaller value of q does not imply that the solution of the equations will be available.

Lawley (1940, 1942, 1943) proposed to estimate λ_{ij} by ML method. In that case, we have assume that there is a data vector of $N = n + 1$ sample observations which are drawn from $N_p(\mu, \Sigma)$, where $\text{rank}(\Sigma) = p$. The sample varinace-covariance matrix of data vector is S , where S has Wishart distribution. The density function of S is

$$F(S) = K |S|_2^{1(n-p-1)} |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} n \text{Tr} \Sigma^{-1} S\right) \quad (4.11)$$

Here K is that part of $F(S)$ which is free of S . Now, the log of likelihood function on sample variance-covariances is

$$L(\Lambda, \psi) = \ln K + \frac{1}{2}(n-p-1)\ln|S| - \frac{n}{2}\ln|\psi + \Lambda\Lambda'| - \frac{n}{2}\text{Tr}[\psi + \Lambda\Lambda']^{-1}S$$

To estimate the $p(q+1)$ unknown parameters in Λ and ψ the normal equations are

$$\frac{\partial L(\Lambda(\psi))}{\partial \psi_i} = 0 \text{ and } \frac{\partial L(\Lambda(\psi))}{\partial \psi_{ij}} = 0$$

$$\begin{aligned} \text{We have, } \frac{\partial L(\Lambda(\psi))}{\partial \psi_i} &= \frac{n}{2} \frac{1}{|\psi + \Lambda\Lambda'|} \frac{\partial |\psi + \Lambda\Lambda'|}{\partial \lambda_i} + \frac{n}{2} \text{Tr}[\psi + \Lambda\Lambda']^{-1} \frac{\partial \psi}{\partial \psi_i} [\psi + \Lambda\Lambda']^{-s} \\ &= -\frac{n}{2} \frac{1}{|\psi + \Lambda\Lambda'|} |\psi + \Lambda\Lambda'|_{ii} + \frac{n}{2} \text{Tr}[\psi + \Lambda\Lambda']^{-1} S |\psi + \Lambda\Lambda'|^{-1} \frac{\partial \psi}{\partial \psi_i} \end{aligned} \quad (4.12)$$

Here $\partial \psi / \partial \psi_i$ is a diagonal matrix of order $(p \times p)$ with 1 in the diagonal position and zero in the off-diagonal position. Again, $|\psi + \Lambda\Lambda'|_{ii}$ is the cofactor of the i -th diagonal element of $[\psi + \Lambda\Lambda']$.

Now, putting $\frac{\partial L(\Lambda(\psi))}{\partial \psi_i} = 0$, we get

$$\begin{aligned} \text{diag} [\{\psi + \Lambda\Lambda'\}^{-1} (I - S[\psi + \Lambda\Lambda']^{-1})] &= 0 \\ \text{or, } \text{diag} (\hat{\Sigma}^{-1}) &= \text{diag} (\hat{\Sigma}^{-1} S \hat{\Sigma}^{-1}) \end{aligned} \quad (4.13)$$

Here $\hat{\Sigma} = \hat{\psi} + \hat{\Lambda}\hat{\Lambda}'$ is the sample variance-covariance matrix. But $\hat{\Sigma}$ is not the ML estimator of Σ . Again,

$$\frac{\partial L(\Lambda(\psi))}{\partial \lambda_{ij}} = -\frac{n}{2} \frac{1}{|\psi + \Lambda\Lambda'|} \sum_{g=1}^p \sum_{h=1}^p |\psi + \Lambda\Lambda'|_{gh} \frac{\partial \sigma_{gh}}{\partial \lambda_{ij}} + \frac{n}{2} \text{Tr}[\psi + \Lambda\Lambda']^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}} [\psi + \Lambda\Lambda']^{-1} \quad (4.14)$$

The first quantity of the right hand of (4.13) is $-\frac{n}{2} \text{Tr} \Sigma^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}}$, where,

$$\frac{\partial \Sigma}{\partial \lambda_{ij}} = \begin{bmatrix} 0 & \dots & 0 & \lambda_{ij} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_{i-1,j} & 0 & \dots & 0 \\ \lambda_{1j} & \dots & \lambda_{i-1,j} & 2\lambda_{ij} & \lambda_{i+1,j} & \dots & \lambda_{pj} \\ 0 & \dots & 0 & \lambda_{i+1,j} & 0 & \dots & 0 \\ 0 & \dots & 0 & \lambda_{pj} & 0 & \dots & 0 \end{bmatrix}$$

The above matrix is a symmetric matrix with zeros in places except i -th row and j -th column. Let the (i,j) -th element of Σ^{-1} be σ^{ij} . Then

$$\begin{aligned} \text{Tr} \Sigma^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}} &= \sigma^{1i} \lambda_{1j} + \sigma^{2i} \lambda_{2j} + \dots + 2\sigma^{ii} \lambda_{ij} + \dots + \sigma^{ip} \lambda_{pj} + \dots + \sigma^{ip} \lambda_{pj} \\ &= 2(\sigma^{1i} \lambda_{1j} + \dots + \sigma^{ip} \lambda_{pj}) \\ &= 2\delta^i T_j \end{aligned}$$

where δ^i is the i -th row of Σ^{-1} and T_j is the j -th column of A . Therefore, the first term of the derivative shown in (4.14) can be written as $-n \Sigma^{-1} A$

The second term can be written as

$$\frac{n}{2} \text{Tr} \Sigma^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}} \Sigma^{-1} S = \frac{n}{2} \text{Tr} \Sigma^{-1} \frac{\partial \Sigma}{\partial \lambda_{ij}}$$

For simplicity, let us write $\Sigma^{-1} S \Sigma^{-1} = Z = (Z_{gh})$.

$$\text{Then } \text{Tr} Z \frac{\partial \Sigma}{\partial \lambda_{ij}} = 2Z'_{Tj}$$

where Z'_i is the i -th row of Z . Therefore, the second term of the derivative shown in (4.14) can be written as ZA . The equations obtained from derivative (4.14) are to be satisfied by the estimates obtained from.

$$(\hat{\Sigma}^{-1} - \hat{\Sigma}^{-1} S \hat{\Sigma}^{-1}) \hat{\Lambda} = 0$$

$$\text{or, } S \hat{\Sigma}^{-1} \hat{\Lambda} = \hat{\Lambda} \quad (4.15)$$

The estimates of factor loadings will be available by solving the set of equations given in (4.13) and in (4.15). Pre- and post-multiplying both sides of (4.15) by the diagonal matrix $\hat{\Psi} = \hat{\Sigma} + \hat{\Lambda} \hat{\Lambda}'$, we get

$$\text{diag}(\hat{\Sigma} - 2\hat{\Lambda} \hat{\Lambda}' + \hat{\Lambda} \hat{\Lambda}' \hat{\Sigma}^{-1} \hat{\Lambda} \hat{\Lambda}') = \text{diag}(S - S \hat{\Sigma}^{-1} \hat{\Lambda} \hat{\Lambda}' - \hat{\Lambda} \hat{\Lambda}' \hat{\Sigma}^{-1} S + \hat{\Lambda} \hat{\Lambda}' \hat{\Sigma}^{-1} \hat{\Lambda} \hat{\Lambda}')$$

Putting the value of (4.15), we get

$$\text{diag}(\hat{\Sigma}) = \text{diag}(S) \quad (4.16)$$

The equation (4.15) can be written as

$$S \hat{\Psi}^{-1} \hat{\Lambda} (I + \hat{\Lambda}' \hat{\Psi}^{-1} \hat{\Lambda}) = \hat{\Lambda} \quad (4.17)$$

$$\text{or } S \hat{\Psi}^{-1} \hat{\Lambda} = \hat{\Lambda} (I + \hat{\Lambda}' \hat{\Psi}^{-1} \hat{\Lambda}) \quad (4.18)$$

This is true, since

$$(\hat{\Psi} + \hat{\Lambda} \hat{\Lambda}')^{-1} \hat{\Lambda} = \hat{\Psi}^{-1} \hat{\Lambda} (I + \hat{\Lambda}' \hat{\Psi}^{-1} \hat{\Lambda})^{-1} \quad (4.19)$$

From equation (4.19) we have $\hat{\Lambda}$. Let the value of $\hat{\Lambda}$ be $\hat{\Lambda}_0$. Now, assuming $\hat{\Lambda}_0$ known, we have from $\partial L(A, \psi) / \partial \psi_i$

$$\text{diag}(\hat{\Lambda}_0 \hat{\Lambda}_0' + \hat{\Psi} - S) = 0 \text{ or, } \hat{\Psi} = \text{diag}(S - \hat{\Lambda}_0 \hat{\Lambda}_0') \quad (4.20)$$

However, the solution of equation (4.20) is not easy and it is available by iterative process [Lawley and Maxwell (1971), Joreskog and Lawley (1968)]. Rao (1955) and Maxwell (1964) have discussed a numerical solution to obtain ψ .

To obtain the numerical solution, pre-multiplying both sides of equation (5.3.10) by ψ^{-1} and rearranging, we get

$$[\hat{\psi}^{-1/2}(S - \hat{\psi})\hat{\psi}^{-1/2}\hat{\Lambda}J] \quad (4.21)$$

where $J = \hat{\Lambda}'\hat{\psi}^{-1}\hat{\Lambda}$ and for unique solution of $\hat{\Lambda}$, the J matrix must be diagonal. Under this condition, the first q characteristics of $(S - \hat{\psi})\hat{\psi}^{-1}$ and $\hat{\psi}^{-1/2}(S - \hat{\psi})\hat{\psi}^{-1/2}$ are equal to the successive elements of J, and the i-th column of $\hat{\psi}^{-1/2}\hat{\Lambda}$ is the characteristic vector corresponding to the i-th largest root of $\hat{\psi}^{-1/2}(S - \hat{\psi})\hat{\psi}^{-1/2}$. As the values of $\hat{\psi}$ are unknown, the values of this vector are not easily available. The values of $\hat{\psi}$ are obtained from $\hat{\psi} = \text{diag}(A = \hat{\Lambda}\hat{\Lambda})$. At this stage, it needs to assume a value of $\hat{\psi}$ and on the basis of this assumed value the value of eigenvector and $\hat{\Lambda}$ are to be obtained.

At the initial step of iterative process, let us consider a one-factor model and for this one-factor model let us take the eigenvector as the initial vector, where the eigenvector corresponds to the largest eigenvalue of S of one-factor model. The values of this eigenvector are so scaled that their sum of squares equals to eigenvalue. The iterative process will follows:

1. Find the eigenvalues of S. Let t_{10} be the greatest eigenvalue and its corresponding eigenvector is a_{10} . The eigenvector is scaled so that $a'_{10}a_{10} = 1_{10}$.
2. Putting the value of a_{10} , $\psi'_{10} = (S - a_{10}a'_{10})$ is obtained. Now, using the value of ψ'_{10} , a matrix $\psi'^{-1/2}_{10}(S - \psi'_{10})\psi'^{-1/2}_{10}$ is formed. Consider that the largest eigenvalue of this newly formed matrix is l_{11} .
3. The eigenvector to l_{11} is so scaled that the scale vector provides $a'_{11}a_{11} = 1_{10}$.

4. The vector a_{11} is pre-multiplied by $\hat{\Psi}_{10}^{1/2}$. This will give approximately the values of single column matrix $\hat{\Lambda}_1$. Let this column be $\hat{\Lambda}_{11}$.

5. Using $\hat{\Lambda}_{11}$ we can get $\hat{\Psi}_{11} = \text{diag}(S - \hat{\Lambda}_{11}\hat{\Lambda}'_{11})$

Let the eigenvector corresponding to the largest eigenvalue of $\hat{\Psi}_{11}$ is a_{12} . The values of a_{12} are scaled so that $a'_{12}a_{12} = l_{12}$, where l_{12} is the largest eigenvalue of $\hat{\Psi}_{11}$. The eigenvector a_{12} is pre-multiplied by $\hat{\Psi}_{11}^{1/2}$, The product will give the second approximate value of $\hat{\Lambda}_{11}$. The process is continued until the successive iterates $\hat{\Lambda}_{11}$ and $\hat{\Lambda}_{11+1}$ do not differ by more than some pre-determined amount. The column vector so obtained is the ML estimates of loadings of one-factor model.

The factor loading thus obtained will not be sufficiently close to those obtained by principal component analysis, especially if it used for q factors model. For factors model the residual matrix is to be found out after estimating $\hat{\lambda}_i$ - at this state, let the eigenvector corresponding to largest eigenvalue of S_1 be a_{20} , where a_{20} is such that $a'_{20}a_{20} = l_{20}$. Here l_{20} is the largest eigenvalue of S_1 . Let us form a $(p \times 2)$ matrix with $\hat{\Lambda}_1$ and a_{20} such that

$$\hat{\Lambda}_{20} [\hat{\Lambda}_1, a_{20}]$$

From $\hat{\Lambda}_{20}$ let us find

$$\hat{\Psi}_{20} = \text{diag}(S - \hat{\Lambda}_{20}\hat{\Lambda}'_{20}) \text{ and } \hat{\Psi}_{20}^{-1/2}(S - \hat{\Psi}_{20})\hat{\Psi}_{20}^{-1/2}$$

This latter matrix is symmetric. Let us find the largest two eigenvalues and their corresponding eigenvectors of this latter matrix and scale those two eigenvectors. Arrange these two scaled vectors into a $(p \times 2)$ matrix are not same. To obtain the factor loading $\hat{\Lambda}_{21}$ for two-factor model pre-multiply $[a_{11}, a_{21}]$ by $\hat{\psi}_{20}^{-1/2}$. This process of iteration is continued until all elements of iterates $\hat{\Lambda}_{21}$ have converged with specified accuracy to the two-factor solution $\hat{\Lambda}_2$.

The process is applied in solving the q -factor likelihood equations. The q -th trial vector is found from the eigenvector of the largest eigenvalue of

$$S_{q-1} = S - \hat{\Lambda}_{q-1} \hat{\Lambda}_{q-1}'$$

It has already been mentioned that, the value of Λ can be estimated from $\Sigma = \Lambda \Lambda' + \psi$. In this equation there are $pq + p$ parameters in Λ and ψ . The parameters in Σ are $\frac{1}{2}p(p+1)$. The number of equations are not equal to the number of parameters. Therefore, to solve the equation, we need some restrictions. Let the restriction be $\Lambda' \psi^{-1} \Lambda$ or $\Lambda' D^{-1} \Lambda$ diagonal, where $D = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$.

Under this restriction $\frac{1}{2}q(q-1)$ parameters in the factor model become zero. Now, let us find out the difference between the parameters in Σ and in the factor model.

This difference is

$$\begin{aligned} s &= \frac{1}{2}p(p+1) - \left\{pq + p - \frac{1}{2}q(q-1)\right\} \\ &= \frac{1}{2}(p-q)^2 - \frac{1}{2}(p+q) \end{aligned} \quad (4.22)$$

If $s < 0$, the unique solution of the equations is not available. This is arisen if the factor model is not well defined. The value $s = 0$, implies that the exact solutions of Λ and ψ are available. The approximate solutions are available if $s > 0$.

One of the important properties of ML estimators of the parameters in a factor model is that, the estimator of variance of i - t variable is

$$\delta_1^3 = \sum_{j=1}^q \lambda_{ij}^2 + \hat{\psi}_1^2$$

Eigenvalues

The eigenvalue for a given factor reflects the variance in all the variables, which is accounted for by that factor. A factor's eigenvalue may be computed as the sum of its squared factor loadings for all the variables. The ratio of eigenvalues is the ratio of explanatory importance of the factors with respect to the variables. If a factor has a low eigenvalue, then it is contributing little to the explanation of variances in the variables and may be ignored. Eigenvalues is also called characteristic roots. The eigenvalue for a given factor reflects the variance in all the variables, which shows variance explained by that particular factor out of the total variance. A factor's eigenvalue may be computed as the sum of its squared factor loadings for all the variables. From the commonality column, we can know how much variance is explained by the first factor out of the total variance. For example, if our first factor explains 68% variance out of the total, this means that 32% variance will be explained by the other factor.

4.4 Statistical Test used in Factor analysis

Factor analysis involves generating one or more unobserved independent variables that correlate with the observed measures. Commonly used in survey research and other applications, factor analysis can be considered a data reduction technique because it reduces a large number of variables that often overlap to a smaller number of factors. As a first step of factor analysis it is need to measure the sampling adequacy by Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity

Kaiser-Meyer-Olkin (KMO) Test

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy is a statistic which indicates the proportion of variance in the variables which is common variance, i.e. which might be caused by underlying factors. It is defined as:

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}, \quad (i \neq j = 1, 2, 3, \dots, p) \quad (4.23)$$

where r_{ij} is the correlation coefficient of i -th and j -th variable and a_{ij} is the partial correlation coefficient of those variables. Generally it used to test the null hypothesis, H_0 : there is no significant contribution of the study variables to the factors.

High values generally (close to 1.0) indicate that a factor analysis may be useful for the data. The Interpretation of the KMO as characterized by Kaiser, Meyer, and Olkin is given below in Table 4.1.

Table 4.1 Interpretation of the KMO as characterized

KMO Value	Degree of Common Variance
0.90 to 1.00	Marvelous
0.80 to 0.89	Meritorious
0.70 to 0.79	Middling
0.60 to 0.69	Mediocre
0.50 to 0.59	Miserable
0.00 to 0.49	Don't Factor

Bartlett's Test of Sphericity

Bartlett's test of sphericity indicates whether the correlation matrix is an identity matrix, which would indicate that all the variables are uncorrelated; each variable correlates perfectly with itself ($r = 1$) but has no correlation with the other variables ($r = 0$). The p value of the Bartlett's test of sphericity indicates relationship among the study variables. Bartlett's test is used to test the null hypothesis that the variables are uncorrelated in the population (Hair *et al*, 1998). i.e. H_0 : there is no significant relationship among the study variables to the factors.

The Bartlett's test statistic is defined as:

$$\chi^2 = - \left[(n-1) - \frac{1}{6} \left(2p+1 + \frac{2}{p} \right) \right] \left[\ln|S| + p \ln \left(\frac{1}{p} \right) \sum l_j \right] \quad (4.24)$$

where, p = number of variables,

K = number of components,

l_j = jth eigenvalue of S

$$df = \frac{(p-1)(p-2)}{2}$$

Scree Plot

The Scree Plot is a technique of graphical representation of eigen values against all the factors. In factor analysis is often applied for the purposes of data reduction or structure detection or simplification. It is most commonly used method for determining how many factors to retain. The number of factors is taken as the factor number that appears just before the “elbow” in the plot. So it is a way to determine the appropriate number of factors to extract in the final solution.

4.5 Results of Factor Analysis

Three northern districts like Rajshahi, Rangpur and Dinajpur have been selected of the present study. The production of HYV Boro rice of these districts may depend on some variables climatic and non climatic variables. The climatic variables rainfall (r), temperature (t), wind speed (w), humidity (h) and cloud coverage (cc). Although these variables have been consider for three stages like vegetative, growing /milk and reproductive (1 for vegetative stage, 2 for growing /milk stage and 3 for reproductive stage) the climatic variable Co2 emission (co) and other two variables, fertilizer use (fu) and tractor use (tu) have been once for the whole production time of the HYV Boro rice. The results are presented in this section.

The KMO, a measure of sampling adequacy, is an index for comparing the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. A large KMO measure indicates that the factor analysis of the variables is good and meaningful, since correlations between pairs of variables can be explained by the other variables (See Norussis, 1985). The

values of KMO and Bartlett's Test of Sphericity for three selected districts are presented in Table 4.2-Table 4.4.

Table 4.2 KMO and Bartlett's Test Statistics value for Rajshahi

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.615
Bartlett's Test of Sphericity	Approx. Chi-Square	653.069
	df	153
	Sig.	0.000

Table 4.3 KMO and Bartlett's Test Statistics value for Rangpur

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.612
Bartlett's Test of Sphericity	Approx. Chi-Square	592.105
	df	153
	Sig.	.000

Table 4.4 KMO and Bartlett's Test Statistics value for Dinajpur

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.597
Bartlett's Test of Sphericity	510.987	574.987
	153	153
	.000	.000

Table 4.2 shows that the KMO statistic for sampling adequacy measure for Rajshahi is 0.615. This value indicates that the data used in this test is suitable for factor analysis. From Table 4.3 and Table 4.4, the KMO statistic for sampling adequacy measure of Rangpur and Dinajpur are 0.612 and 0.597 respectively. A value of 0.60 or above from the KMO measure of sampling adequacy indicates that the data are adequate for the purposes of exploratory factor analysis

(Tabachnick and Fidell 2001). Therefore the KMO measures of our selected districts are in acceptable level for factor analysis.

The Bartlett's test of sphericity is used to test the hypothesis that the correlation matrix is an identity matrix. That is, all diagonal terms are unity and all off diagonal terms are zero. The Bartlett's test of sphericity for Rajsahi obtained this value as 653.069. This value is significant at 1% level of significance. The Bartlett's test of Sphericity values for Rangpur and Dinajpur are 592.105 and 574.987 respectively. These values are also significant at 1% level of significance. Therefore, it is unlikely that the population correlation matrix is an identity. If the hypothesis that the population correlation matrix is an identity cannot be rejected because the observed significance level is large, so according to Norussis 1985 we may use of the factor model. The value of KMO obtained by this test characterized by Norussis as "mediocre" and the highly significant level of the test of sphericity, are both very comfortable indications that the given set of data of the three selected districts are adequate for factor analysis.

The Scree test is use to determine the number of common factors necessary to represent the data. The Scree test advocated by Cattell (in Kim, 1978, p. 44), directs one to examine the graph of eigenvalues, and stop factoring at a point where the eigenvalues begin to level off forming a straight line with an almost horizontal slope. According to Kim, 1978, this method is often superior to other tests where the interest is in locating only the major common factors.

The Scree plot for determining the factors influencing the HYV Boro production of the selected districts are presented in Fig 4.1- Fig 4.3. The Scree plot justifies

the five factors as a influential to the HYV Boro production for each districts separately.

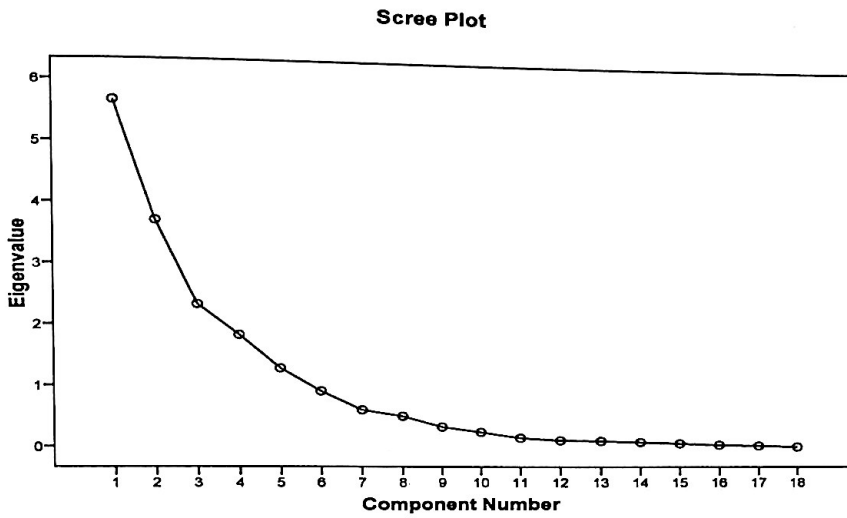


Fig 4.1 Scree Plot for Rajshahi

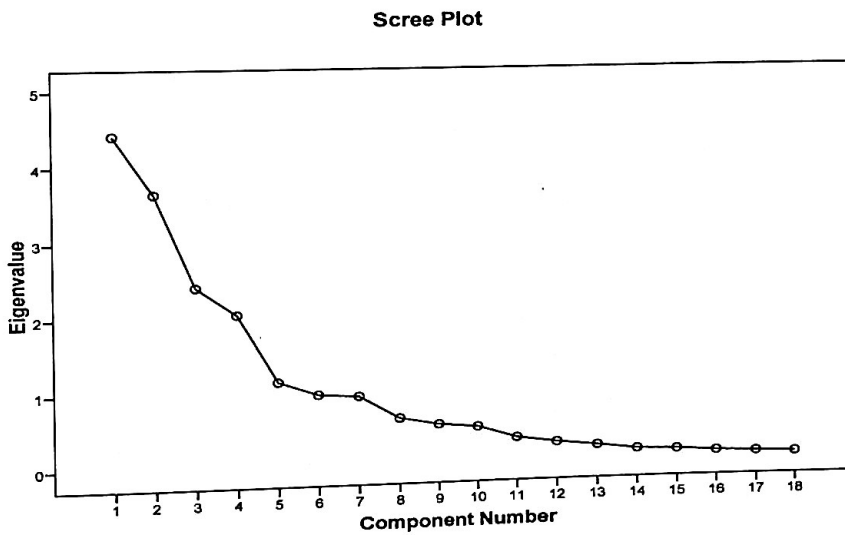


Fig 4.2 Scree Plot for Rangpur

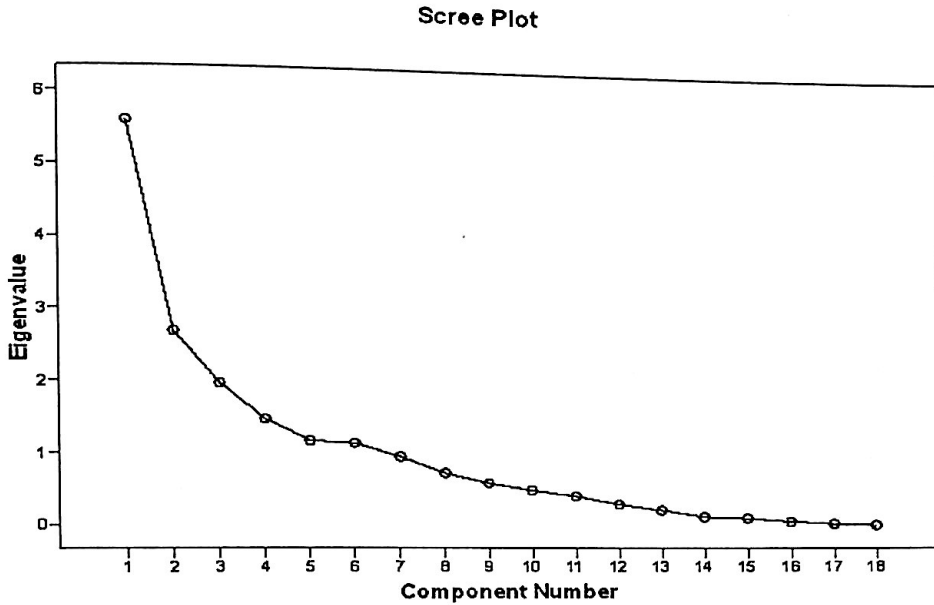


Fig 4.3 Scree Plot for Dinajpur

Factor wise eigenvalues, percentages of variance accounted for the factors and cumulative percentages of variance explained by the factors selected for the district Rajshahi are reported in the Table 4.5.

Table 4.5 Factor wise explained variation for Rajshahi

HYV Boro production influencing Factors	Eigenvalue	Percent of Variance Accounted for	Cumulative Percent of Variance
1	5.659	31.440	31.440
2	3.707	20.593	52.033
3	2.331	12.952	64.986
4	1.833	10.185	75.171
5	1.288	7.156	82.327

Table 4.5 shows that the five factors account for 82.33% of the total variance, with Factor I accounting for 31.44%, Factor II, 20.59%, Factor III, 12.95%, Factor IV, 10.18% and Factor V 7.16%. These five factors are more than adequate to represent the data considering the convention that at least 75% (two third percents) of the total variance is needed to capture information for the production of HYV Boro rice of Rajshahi district.

Table 4.6 Item Statements of Factors with Rotated Factor Loadings for Rajshahi

Number	Variable Name	Factor Loading
1	Temperature in vegetative stage	0.240
2	Temperature in growing stage	0.249
3	Tractor use	0.649
4	Cloud Coverage in growing Stage	0.649
5	Temperature in Reproductive Stage	0.649
6	Cloud Coverage vegetative stage	0.745
7	Humidity in vegetative Stage	0.767
8	Humidity in Reproductive Stage	0.770
9	Humidity in growing Stage	0.785
10	Cloud Coverage in Reproductive Stage	0.807
11	Rainfall in vegetative Stage	0.822
12	Wind Speed in Reproductive Stage	0.886
13	Rainfall in growing Stage	0.896
14	Wind Speed in growing stage	0.902
15	Rainfall in Reproductive Stage	0.921
16	Fertilizer use	0.935
17	Carbon dioxide Emission	0.935
18	Wind Speed in vegetative stage	0.944

Table 4.7 Selected Factors with Rotated Factor Loadings for Rajshahi

Factors	Variable Name	Factor Loading
1	Fertilizer use	0.935
	Carbon dioxide Emission	0.935
2	Rainfall in Reproductive Stage	0.921
	Cloud Coverage in Reproductive Stage	0.807
	Humidity in Reproductive Stage	0.770
3	Wind Speed in growing stage	0.902
	Wind Speed in vegetative stage	0.944
	Wind Speed in Reproductive Stage	0.886
4	Rainfall in growing Stage	0.896
	Humidity in growing Stage	0.785
5	Rainfall in vegetative Stage	0.822
	Humidity in vegetative Stage	0.767

Five influential factors for the HYV Boro rice production of Rajshahi district are selected and reported in Table 4.5. Item statements of factors with rotated factor loadings for Rajshahi are presented in Table 4.6. Selected factors with rotated factor loadings are given in Table 4.7. Therefore the first factor accounts for the largest proportion, 31.44% of the total variance. This factor includes two variables like Fertilizer Use and Carbon Dioxide Emission with same factor loading of 0.935. Factor 2 accounts for 20.59%, the second largest proportion of the total variance consisting of the three variables like, Rainfall in Reproductive Stage, Cloud Coverage in Reproductive Stage and Humidity in Reproductive Stage with factor loadings 0.902, 0.807 and 0.770 respectively. The third factor accounts for

12.95% of the total variance consisting of three variables of Wind Speed at three stages of the HYV Boro production. The factor loadings of Wind Speeds are 0.944 for Vegetative Stage, 0.902 for Growing Stage and 0.886 for Reproductive Stage. Factor 4 accounts for the next 10.19% variation of the total variance, consisting of two variables, Rainfall in Growing Stage and Humidity in Growing Stage with factor loadings 0.896 and 0.785 respectively. Factor 5 is the last factor accounts for 7.16% of the total variation consisting of two variables, Rainfall in Vegetative Stage and Humidity in Vegetative Stage of factor loadings 0.822 and 0.767 respectively.

Factor wise eigenvalues, percentages of variance accounted for the factors and cumulative percentages of variance explained by the factors selected for the district Rangpur are reported in the Table 4.8.

Table 4.8 Factor wise explained variation for Rangpur

HYV Boro production influencing Factors	Eigenvalue	Percent of Variance Accounted for	Cumulative Percent of Variance
1	4.419261	24.55145	24.55145
2	3.644948	20.24971	44.80116
3	2.404627	13.35904	58.1602
4	2.032834	11.29352	69.45372
5	1.126867	6.26037	75.71409

Table 4.8 shows that the five factors account for 75.71% of the total variance, with Factor I accounting for 24.80%, Factor II, 20.25%, Factor III, 13.36%, Factor IV, 11.29% and Factor V 6.26%. These five factors are more than adequate to

represent the data considering the convention that at least 75% (two third percents) of the total variance is needed to capture information for the production of HYV Boro rice of Rangpur district.

Table 4.9 Item Statements of Factors with Rotated Factor Loadings for Rangpur

Number	Variable Name	Factor Loading
1	Temperature in growing stage	0.249
2	Tractor use	0.649
3	Cloud Coverage in growing Stage	0.649
4	Temperature in Reproductive Stage	0.649
5	Cloud Coverage vegetative stage	0.745
6	Humidity in vegetative Stage	0.767
7	Humidity in growing Stage	0.785
8	Cloud Coverage in Reproductive Stage	0.807
9	Humidity in Reproductive Stage	0.81
10	Rainfall in vegetative Stage	0.822
11	Wind Speed in Reproductive Stage	0.886
12	Rainfall in growing Stage	0.896
13	Wind Speed in growing stage	0.902
14	Temperature in vegetative stage	0.912
15	Rainfall in Reproductive Stage	0.921
16	Fertilizer use	0.935
17	Carbon dioxide Emission	0.935
18	Wind Speed in vegetative stage	0.944

Table 4.10 Selected Factors with Rotated Factor Loading for Rangpur

Factors	Variable Name	Factor Loading
1	Humidity in growing Stage	0.785
	Humidity in Reproductive Stage	0.810
	Rainfall in growing Stage	0.896
2	Temperature in Reproductive Stage	0.649
	Temperature in vegetative stage	0.912
3	Wind Speed in Reproductive Stage	0.886
	Wind Speed in growing stage	0.902
4	Fertilizer use	0.935
	Carbon dioxide Emission	0.935
5	Cloud Coverage vegetative stage	0.745

Five influential factors for the HYV Boro rice production of Rangpur district are selected and reported in Table 4.8. Item statements of factors with rotated factor loadings for Rangpur are presented in Table 4.9. Selected factors with their rotated factor loadings are given in Table 4.10. It shows that Factor 1 accounts for the largest proportion, 24.55% of the total variance. This factor includes three variables like Humidity in growing Stage, Humidity in Reproductive Stage and Rainfall in growing Stage with factor loadings of 0.785, 0.810 and 0.896 respectively. Factor 2 accounts for 20.25%, the second largest proportion of the total variance consisting of the two variables like, Temperature in Reproductive

Stage and Temperature in Vegetative Stage with factor loadings 0.649 and 0.912 respectively. The third factor accounts for 13.36% of the total variance consisting of two variables of Wind Speed in Reproductive Stage and Wind Speed in growing stage of the HYV Boro production. The factor loadings of Wind Speeds are 0.886 and 0.902. Factor 4 accounts for the next 11.29% variation of the total variance, consisting of two variables, Fertilizer use and Carbon dioxide Emission with same factor loadings 0.935. The last factor (Factor 5) accounts for 6.26% of the total variation consists of only one variable, Cloud Coverage Vegetative Stage with loading 0.745.

Factor wise eigenvalues, percentages of variance accounted for the factors and cumulative percentages of variance explained by the factors selected for the district Dinajpur are reported in the Table 4.11.

Table 4.11 Factor wise explained variation for Dinajpur

HYV Boro production influencing Factors	Eigenvalue	Percent of Variance Accounted for	Cumulative Percent of Variance
1	5.588	31.044	31.044
2	2.686	14.922	45.966
3	1.965	10.916	56.881
4	1.470	8.167	65.048
5	1.173	6.519	71.567

Table 4.11 shows that the five factors account for 71.57% of the total variance, with Factor I accounting for 31.04%, Factor II, 14.92%, Factor III, 10.92%, Factor IV, 8.17% and Factor V 6.52%. These five factors are more than adequate to

represent the data considering the convention that more than 70% of the total variance is needed to capture information for the production of HYV Boro rice of Rangpur district.

Table 4.12 Item Statements of Factors with Rotated Factor Loadings for Dinajpur

Number	Variable Name	Factor Loading
1	Temperature in growing stage	0.691
2	Tractor use	0.433
3	Cloud Coverage in growing Stage	0.703
4	Temperature in Reproductive Stage	0.457
5	Cloud Coverage vegetative stage	0.595
6	Humidity in vegetative Stage	0.911
7	Humidity in growing Stage	0.674
8	Cloud Coverage in Reproductive Stage	0.459
9	Humidity in Reproductive Stage	0.662
10	Rainfall in vegetative Stage	0.721
11	Wind Speed in Reproductive Stage	0.657
12	Rainfall in growing Stage	0.637
13	Wind Speed in growing stage	0.538
14	Temperature in vegetative stage	0.456
15	Rainfall in Reproductive Stage	0.844
16	Fertilizer use	0.805
17	Carbon dioxide Emission	0.807
18	Wind Speed in vegetative stage	0.592

Table 4.13 Selected Factors with Rotated Factor Loading for Dinajpur

Factors	Variable Name	Factor Loading
1	Humidity in Vegetative Stage	0.911
	Carbon dioxide Emission	0.807
	Fertilizer use	0.805
2	Cloud Coverage in Growing Stage	0.703
3	Rainfall in Vegetative Stage	0.721
	Rainfall in Growing Stage	0.637
4	Wind Speed in Reproductive Stage	0.657
5	Temperature in Growing Stage	0.691

Five influential factors for the HYV Boro rice production of Dinajpur district are selected and reported in Table 4.11. Item statements of factors with rotated factor loadings for Dinajpur are presented in Table 4.12. Selected factors with rotated factor loadings are given in Table 4.13. It shows that Factor 1 accounts for the largest proportion, 31.04% of the total variance. This factor includes three variables like Humidity in Vegetative Stage, Fertilizer Use and Carbon Dioxide Emission with factor loadings 0.911, 0.805 and 0.807 respectively. Factor 2 accounts for 14.92%, the second largest proportion of the total variance consisting of only one variable Cloud Coverage in Growing Stage with factor loadings 0.702. The third factor accounts for 10.92% of the total variance consisting of two variables of Rainfall in Vegetative Stage and Rainfall in Growing Stage. The factor loadings are 0.721 and 0.637 respectively. Factor 4 accounts for the next

8.17% variation of the total variance, consisting only one variable, Wind Speed in Reproductive Stage with factor loading 0.657. Factor 5 is the last factor accounts for 6.52% of the total variation consisting of the variable Temperature in Growing Stage with its loading 0.691.

Chapter Five

Result Discussion and Conclusion

5.1 Introduction

The findings of the present study and concluding remarks are upheld in this chapter.

5.2 Discussion of Results

The production of HYV Boro rice may differ due to impact of some climatic and non climatic factors throughout the production period. The climatic factors may be rainfall (r), temperature (t), wind speed (w), humidity (h), cloud coverage (cc) and Co2 emission (co). The non climatic factors may be fertilizer use (fu), technological changes etc. In the present study we are interested to find future production through appropriate trend model for three selected districts like Rajshahi, Rangpur and Dinajpur. Also we are interested to extract the factors that are influential to the HYV Boro rice production of these districts. The detailed results of this study are discussed in this section.

5.2.1 Results of Trend Analysis

The descriptive statistics of the production of HYV Boro rice for the period 1971 to 2007 of the selected districts, Rajshahi, Rangpur and Dinajpur are reported in Table 3.1. The result from this table shows that the maximum production of HYV

Boro is 3.81 metric tone per hector and the minimum production is 2.05 metric tone per hector for Rajshahi district. The maximum production for Rangpur and Dinajpur districts are 3.91 and 3.81 metric tones per hector respectively. The minimum productions of these districts are 1.68 and 2.03 metric tones per hector respectively. The average productions are 2.8483 m.ton/hector, 2.6746 m.ton/hector and 2.7150 m.ton/hector respectively for Rajshahi, Rangpur and Dinajpur. The standard deviations are 0.46551, 0.51369 and 0.41970 and the coefficients of variations are 16.34%, 19.20% and 15.45% respectively for Rajshahi, Rangpur and Dinajpur. Hence the production of HYV Boro rice for Rangpur shows the maximum variation and Dinajpur shows the minimum variation. The skewness of the production of HYV Boro rice for Rajshahi, Rangpur and Dinajpur are 0.342, 0.285 and 0.917 and the kurtosis values for are 2.327, 3.256 and 3.859 respectively. Therefore, the data are positively skewed. The distributions of HYV Boro production for Rajshahi and Rangpur are platykurtic but the distribution of Dinajpur leptokurtic.

After examining the TS plot (Fig. 3.1 - 3.3) Mann-Kendall Test and Spearman Tests have been done to check the validity of the trend of the study variables. Table 3.1 shows that all the values of t_m statistics fall outside of the 95% confidence interval and all the value of t_s statistics of Table 3.2 fall outside of 95% confidence interval. Hence we may reject the null hypothesis (H_0 : There is no trend in the data set) for both Mann-Kendall Test and Spearman Tests. Therefore, there exists trend among the data sets for the selected districts of this study.

We have used the Akaike's information criterion to choosing the lag length of the variables are given in Table 3.4 to Table 3.5. The test values (τ_μ) for with trend and without trend are reported in the Table 3.6 and Table 3.7.

The DF Unit Root statistics with associated 10% critical values are reported in Table 3.8. The test statistic value for variable yhraj falls in the rejection region but the test statistic value for yhdin and yhran fall in the acceptance region hence we may accept our null hypothesis for yhdin and yhran. The ADF test rejects the null hypothesis for all the variables yhraj, yhran and yhdin. Therefore, the data of each of the series are stationary. Table 3.9 indicates that all the variables are stationary in level.

District-wise OLS estimate of parameters, standard error of estimates, T-Ratio and corresponding p-value of the estimates for linear, compound and quadratic trend models are shown in Table 3.10 to Table 3.12. Examining Table 3.10 we can say that the estimates of parameter for Linear trend models for Rajshahi, Rangpur and Dinajpur are positive and significant at 1% level of significance. So the productions of these districts are influenced by linear time. Table 3.11 shows that the estimates of parameter for Compound trend models for Rajshahi, Rangpur and Dinajpur are positive and significant at 1% level of significance. So the compound time trend is present for the productions of these districts. Table 3.12 shows that the estimates of parameter for Compound trend models for Rajshahi, Rangpur and Dinajpur are positive and significant at 1% level of significance.

Various diagnostic tools like, coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2), Jarque-Bera statistic (JB) and Durbin-Watson statistic (DW) have been used in OLS method for Linear Trend model, Compound Trend model and Quadratic Trend model for the selected districts are reported in the Table 3.13-Table 3.15.

In Table 3.13, the R^2 values are 0.486, 0.588, 0.383 and adjusted R^2 are 0.469, 0.576, 0.364 respectively for yhraj, yhran and yhdin by OLS for Linear Trend model. Therefore about 46.9% of total variation for yhraj, about 57.6% of total variation for yhran and about 36.4% of total variation for yhdin are explained by Linear Trend model. The DW values of all models for yhraj are less than 2 indicate that the data are affected by autocorrelation.

In Table 3.14, the R^2 values are 0.450, 0.524, 0.383 and adjusted R^2 are 0.433, 0.509, 0.364 respectively for yhraj, yhran and yhdin by OLS for Compound Trend model. Therefore about 43.3% of total variation for yhraj, about 50.9% of total variation for yhran and about 36.4% of total variation for yhdin are explained by Compound Trend model. The DW values of all models for yhraj are less than 2 indicate that the data are affected by autocorrelation.

In Table 3.15, the R^2 values are 0.728, 0.618, 0.484 and adjusted R^2 are 0.711, 0.593, 0.451 respectively for yhraj, yhran and yhdin by OLS for Quadratic Trend model. Therefore about 71.1% of total variation for yhraj, about 59.3% of total variation for yhran and about 45.1% of total variation for yhdin are explained by Quadratic Trend model. The DW values of all models for yhraj are less than 2 indicate that the data are affected by autocorrelation.

The Cochrane and Orcutt estimates of parameters (Section 2.4.2.2) for all trend models and for the selected districts are computed and reported in the Table 3.16-Table 3.18. It is seen from Table 3.16 that the Cochrane and Orcutt estimates of the parameters of Linear Trend model shows that the coefficient of t and constant used in the model for y_{harj} , y_{hran} and y_{hdin} are significant at 1% level. From Table 3.17 we see that the Cochrane and Orcutt estimates of the parameters of Compound Trend model shows that the coefficient of t and constant used in the model for y_{harj} , y_{hran} and y_{hdin} are significant at 1% level. Also from Table 3.18, the Cochrane and Orcutt estimates of the parameters of Quadratic Trend model shows that the coefficient of t and constant used in the model for y_{harj} , y_{hran} and y_{hdin} are significant at 1% level.

Based on the results provided in Table 3.19-Table 3.21 the R^2 values are 0.664, 0.6198, 0.4308 and adjusted R^2 are 0.6535, 0.6079, 0.413 respectively for y_{harj} , y_{hran} and y_{hdin} by Cochrane and Orcutt method for Linear Trend model. Therefore about 65.4% of total variation for y_{harj} , about 60.8% of total variation for y_{hran} and about 41.3% of total variation for y_{hdin} are explained by Linear Trend model. DW statistics for this trend model are 1.72, 1.96, 1.87 respectively for y_{harj} , y_{hran} and y_{hdin} respectively. This implies that the data sets are not affected by autocorrelation problem. For Compound Trend model the R^2 values are 0.6202, 0.5526, 0.4290 and adjusted R^2 are 0.6083, 0.5386, 0.4112 respectively for y_{harj} , y_{hran} and y_{hdin} by Cochrane and Orcutt method. Therefore about 60.83% of total variation for y_{harj} , about 53.86% of total variation for y_{hran} and about 41.12% of total variation for y_{hdin} are explained by Compound Trend

model. DW statistics for compound trend model are 1.72, 1.98, 1.88 respectively for yhrāj, yhran and yhdin respectively. This implies that the data sets are not affected by autocorrelation problem. The R^2 values are 0.7445, 0.6411, 0.4991 and adjusted R^2 are 0.7281, 0.6180, 0.4668 respectively for yhrāj, yhran and yhdin by Cochran and Orcutt method for Quadratic Trend model. About 72.813% of total variation for yhrāj, about 61.80% of total variation for yhran and about 46.68% of total variation for yhdin are explained by Quadratic Trend model. It is noted from Compound Trend Model that the DW statistics of the data sets yhrāj, yhran and yhdin are 1.82, 2.01 and 1.94. Since this statistics are near to 2 this indicates that these data sets are not affected by autocorrelation problem.

The forecasting performance of the trend models for HYV Boro production of Rajshahi have been done and the results are demonstrated in Table 3.22. The MSE's are 0.054, 0.042 and 0.043 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.210, 0.141 and 0.164. Hence the Compound Trend model possesses the lowest MSE and AME values for HYV Boro rice production of Rajshahi district. Quadratic Trend model possesses next lowest MSE and AME values while Linear Trend model possesses the highest MSE and AME values among the three methods for the data of Rajshahi district.

Table 3.23 provides the forecasted production of HYV Boro rice with corresponding MSE and AME values for Rangpur district. The MSE's are 0.071, 0.070 and 0.095 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.189, 0.189 and 0.271. Therefore, the Compound Trend model possesses the lowest MSE and AME values for HYV Boro rice production

of Rajshahi district but the Linear Trend model possesses the highest MSE and AME values among the three methods for the data of Rajshahi district.

The forecasted production of HYV Boro rice and method-wise MSE and AME values for Dinajpur district are given in Table 3.24. The MSE's are 0.056, 0.053 and 0.051 respectively for Linear, Compound and Quadratic Trend models and the AME's are 0.184, 0.183 and 0.184 for Dinajpur district. It is observed that the Quadratic Trend model has minimum MSE value than other two trend models but Compound Trend has the lowest AME value. The Linear and Compound methods have same AME values.

Finally the growth rates of these districts are computed for Compound Trend Model. The rates are 1.009 1.0124 and 1.0075 respectively for Rajshahi, Rangpur and Dinapur.

5.2.2 Results of Factor Analysis

The KMO statistics for sampling adequacy measure of Rajshahi, Rangpur and dinajpur are summarized in Table 4.2-Table 4.4. It is seen from these Tables that the KMO statistic for Rajshahi is 0.615, the KMO statistic for Rangpur is 0.612 and the KMO statistic for Dinajpur is 0.597. According to Tabachnick and Fidell (2001), a value of 0.60 or above from the KMO measure of sampling adequacy indicates that the data of our selected districts are adequate for the exploratory factor analysis.

To test the all diagonal terms are unity and all off diagonal terms are zero we uses the Bartlett's test of sphericity for Rajsahi The results are reported in Table 4.2. The chi square value for Rajshahi is 653.069 idicate that the test is significant at 1% level of significance. Therefore, according to Norussis 1985 we may use of the

factor model. From the Table 4.3.-Table 4.4, the Bartlett's test of Sphericity values for Rangpur and Dinajpur are 592.105 and 574.987 respectively. These values are also significant at 1% level of significance. Hence the factor analysis is may be used for this districts.

The Scree plot is a very useful graphical presentation is use to determine the number of common factors necessary for any data set. For the present study the Scree plot for determining the factors influencing the HYV Boro production of the selected districts are presented in Fig 4.1- Fig 4.3. The Scree plot justifies the five factors as an influential to the HYV Boro production for each district separately.

The number of HYV Boro production influencing factors to extract to the study with eigenvalues, percentages of variation account by the factors and their cumulative percentages of explained variances for Rajshahi are reported in Table 4.5. The result shows that there are five factors account for 82.33% of the total variance, with Factor I accounting for 31.44%, Factor II, 20.59%, Factor III, 12.95%, Factor IV, 10.18% and Factor V 7.16%. These five factors are more than adequate to represent the data considering the convention that at least 75% (two third percents) of the total variance is needed to capture information for the production of HYV Boro rice of Rajshahi district. The variables included in these factors are Fertilizer use, Carbon dioxide Emission, Rainfall in Reproductive Stage, Cloud Coverage in Reproductive Stage, Humidity in Reproductive Stage, Wind Speed in growing stage, Wind Speed in vegetative stage, Wind Speed in Reproductive Stage, Rainfall in growing Stage, Humidity in growing Stage, Rainfall in vegetative Stage and Humidity in vegetative Stage. Item Statements of

Factors with Rotated Factor Loadings for Rajshahi are presented in Table 4.6. Factors with Rotated Factor Loadings are given in Table 4.7. Therefore the first factor accounts for the largest proportion, 31.44% of the total variance. This factor includes two variables like Fertilizer Use and Carbon Dioxide Emission with same factor loading of 0.935. Factor 2 accounts for 20.59%, the second largest proportion of the total variance consisting of the three variables like, Rainfall in Reproductive Stage, Cloud Coverage in Reproductive Stage and Humidity in Reproductive Stage with factor loadings 0.902, 0.807 and 0.770 respectively. The third factor accounts for 12.95% of the total variance consisting of three variables of Wind Speed at three stages of the HYV Boro production. The factor loadings of Wind Speeds are 0.944 for Vegetative Stage, 0.902 for Growing Stage and 0.886 for Reproductive Stage. Factor 4 accounts for the next 10.19% variation of the total variance, consisting of two variables, Rainfall in Growing Stage and Humidity in Growing Stage with factor loadings 0.896 and 0.785 respectively. Factor 5 is the last factor accounts for 7.16% of the total variation consisting of two variables, Rainfall in Vegetative Stage and Humidity in Vegetative Stage of factor loadings 0.822 and 0.767 respectively.

The number of HYV Boro production influencing factors to extract to the study with eigenvalues, percentages of variation account by the factors and their cumulative percentages of explained variances for Rangpur are reported in Table 4.8. The result shows that that the five factors account for 75.71% of the total variance, with Factor I accounting for 24.80%, Factor II, 20.25%, Factor III, 13.36%, Factor IV, 11.29% and Factor V 6.26%. These five factors are more than

adequate to represent the data considering the convention that at least 75% (two third percents) of the total variance is needed to capture information for the production of HYV Boro rice of Rangpur district.

The variables included in these factors are Humidity in growing Stage, Humidity in Reproductive Stage, Rainfall in growing Stage, Temperature in Reproductive Stage, Temperature in vegetative stage, Wind Speed in Reproductive Stage, Wind Speed in growing stage, Fertilizer use, Carbon dioxide Emission and Cloud Coverage in Vegetative Stage. Factors with Rotated Factor Loadings for Rangpur are presented in Table 4.10. From this table it shows that Factor 1 accounts for the largest proportion, 24.55% of the total variance. This factor includes three variables like Humidity in growing Stage, Humidity in Reproductive Stage and Rainfall in growing Stage with factor loadings of 0.785, 0.810 and 0.896 respectively. Factor 2 accounts for 20.25%, the second largest proportion of the total variance consisting of the two variables like, Temperature in Reproductive Stage and Temperature in Vegetative Stage with factor loadings 0.649 and 0.912 respectively. The third factor accounts for 13.36% of the total variance consisting of two variables of Wind Speed in Reproductive Stage and Wind Speed in growing stage of the HYV Boro production. The factor loadings of Wind Speeds are 0.886 and 0.902. Factor 4 accounts for the next 11.29% variation of the total variance, consisting of two variables, Fertilizer use and Carbon dioxide Emission with same factor loadings 0.935. The last factor (Factor 5) accounts for 6.26% of the total variation consists of only one variable, Cloud Coverage Vegetative Stage with loading 0.745.

The number of HYV Boro production influencing factors with their eigenvalues, percentages of variation and their cumulative percentages of explained variances for Dinajpur are exhibited in Table 4.11. From this table it is seen that five influential factors for the HYV Boro rice production of Dinajpur district explained 71.57% of total variation. Rotated component matrix and selected factor wise factor loadings of Dinajpur are given in Table 4.12 and Table 4.13 respectively. It is noted that the first factor accounts for the largest proportion, 31.04% of the total variance. This factor includes three variables like Humidity in Vegetative Stage, Fertilizer Use and Carbon Dioxide Emission with factor loadings 0.911, 0.805 and 0.807 respectively. The second factor accounts for 14.92%, the second largest proportion of the total variance consisting of only one variable Cloud Coverage in Growing Stage with factor loadings 0.702. The third factor accounts for 10.92% of the total variance consisting of two variables of Rainfall in Vegetative Stage and Rainfall in Growing Stage. The factor loadings are 0.721 and 0.637 respectively. Factor 4 accounts for the next 8.17% variation of the total variance, consisting only one variable, Wind Speed in Reproductive Stage with factor loading 0.657. The last factor accounts for 6.52% of the total variation consisting of the variable Temperature in Growing Stage with its loading 0.691.

5.3 Conclusion

High yielding Boro rice produces enormously in the Northern region of Bangladesh. On the basis of per hector production three HYV Boro rice producing districts, Rajshahi, Rangpur and Dinajpur have been selected by Simple Random Sampling as a study area for the present study. Productivity of HYV Boro rice

may depend on several climatic and non climatic variables like, rainfall, temperature, wind speed, humidity, cloud coverage, Co2 emission, fertilizer use and tractor. Rice productions in the recent years are hampering and the production variability is occurring due to changing climatic conditions and some other causes. So it is our concern to know the trend of HYV Boro rice production of the selected districts. Therefore to fulfill the objectives of the present study Trend Analysis and Factor Analysis methodologies have been anticipated.

Linear, Compound and Quadratic trend models have used to forecast the production up to the next 2021 years on basis of data the data set from 1971 to 2003. The non parametric statistical tests Mann-Kendall test and Spearman are used to the validity of trends in the variables used in the present study and the results are show in the chapter two. The result shows that both the test statistics Mann-Kendal and Spearman fall outside of the 95% confidence interval of the respective test statistics indicate the rejection of null hypothesis. Therefore, we may say that the trend exists for all the variables used in the data set of the districts Rajshahi, Rangpur and Dinajpur. The HYV rice productions for different districts are time series data so this study therefore emphasises on tougher justification of time series properties. A good understanding of the time-series properties of agricultural productions are prerequisite to useful analyses of risk management and forecasting problems, and one question is, are these data stationary? So we have used DF and ADF tests of stationarity of this study.

The DF and ADF statistics reject the null hypothesis indicate that no unit root exists in the data sets that is the data of each of the series are stationary. Therefore among

a variety of trend models available in the literature the commonly used linear trend model, compound trend model and quadratic trend models have been used for the preset study. The OLS estimates of parameters of all the models with standard error and p-values are reported in detailed in the chapter two. The DW statistics shows that all methods for all districts are affected by autocorrelation problem.

To overcome from this situation Cochrane-Orcutt estimation technique have been applied and the results indicate that the DW statistics more or less equal to 2. Also the JB test shows Normality of the data sets of the three districts. The R^2 values and adjusted R^2 values improve for the Cochrane-Orcutt estimation techniques. Large increments 17%-18% are notable of Rajshahi districts for both the Linear and Compound models. For rest of the models of other districts and all the districts for quadratic trend model it increases within 2%-4% of the total variations.

The HYV Boro productions of the selected districts for three trend models by OLS techniques and Cochrane-Orcutt estimation techniques have been forecasted using the data from 1971 to 2003 and the results also given in detailed in chapter 2. The forecasting performances have been checked by MSE AME. Therefore it is seen from the result that the Compound Trend model performs better maintaining the growth rate 1.009 for Rajshahi, 1.0124 for Rangpur and 1.0075 for Dinapur.

Factor analysis has been used to judge of influences of study variables on HYV Boro production of the selected districts. Factor analysis is a statistical technique that permits the reduction of a large number of correlated variables to a smaller number of latent dimensions to achieve parsimony by using the smallest number

of explanatory concepts to explain the maximum amount of common variance in a correlation matrix.

Bartlett proposed a significance test for a covariance or correlation matrix and a significance test to be applied to the residual matrix after a given number of principal components had been extracted is called Bartlett's test of Sphericity. This test is applicable to determining whether or not any factors or components should be extracted from the correlation matrix. We have used this test for the selected districts. The Bartlett's test values for Rajshahi, Rangpur and are 653.069, 592.105 and 574.987 respectively and each of statistics have the same p-value, 0.000. Therefore, all the tests are significant at 1% level of significance. The value of KMO obtained by this test characterized by Norussis as "mediocre" and the highly significant level of the test of sphericity, are both very comfortable indications that the given set of data of the three selected districts are justified for factor analysis.

Determining the number of factors to use has been an issue since the beginning of factor analysis. There is no perfect way to determine how many factors to be used for factor analysis. The Scree test is use to determine the number of common factors necessary to represent the data. The Scree plot for determining the factors influencing the HYV Boro production of the selected districts are presented in Fig 4.1- Fig 4.3 (in chapter 4). The Scree plot justifies the five factors as an influential to the HYV Boro production for each district separately.

Five influential factors for the HYV Boro rice production of Rajshahi district account for 82.33% of the total variance. Among them the first factor accounts for

the largest proportion, 31.44% of the total variance. This factor includes two variables like Fertilizer Use and Carbon Dioxide Emission with same factor loading of 0.935. Factor 2 accounts for 20.59%, the second largest proportion of the total variance consisting of the three variables like, Rainfall in Reproductive Stage, Cloud Coverage in Reproductive Stage and Humidity in Reproductive Stage with factor loadings 0.902, 0.807 and 0.770 respectively. The third factor accounts for 12.95% of the total variance consisting of three variables of Wind Speed at three stages of the HYV Boro production. The factor loadings of Wind Speeds are 0.944 for Vegetative Stage, 0.902 for Growing Stage and 0.886 for Reproductive Stage. Factor 4 accounts for the next 10.19% variation of the total variance, consisting of two variables, Rainfall in Growing Stage and Humidity in Growing Stage with factor loadings 0.896 and 0.785 respectively. Factor 5 is the last factor accounts for 7.16% of the total variation consisting of two variables, Rainfall in Vegetative Stage and Humidity in Vegetative Stage of factor loadings 0.822 and 0.767 respectively.

Five influential factors for the HYV Boro rice production of Rangpur district that account for 75.71% of the total variance are selected. It shows that Factor 1 accounts for the largest proportion, 24.55% of the total variance. This factor includes three variables like Humidity in growing Stage, Humidity in Reproductive Stage and Rainfall in growing Stage with factor loadings of 0.785, 0.810 and 0.896 respectively. Factor 2 accounts for 20.25%, the second largest proportion of the total variance consisting of the two variables like, Temperature in Reproductive Stage and Temperature in Vegetative Stage with factor loadings

0.649 and 0.912 respectively. The third factor accounts for 13.36% of the total variance consisting of two variables of Wind Speed in Reproductive Stage and Wind Speed in growing stage of the HYV Boro production. The factor loadings of Wind Speeds are 0.886 and 0.902. Factor 4 accounts for the next 11.29% variation of the total variance, consisting of two variables, Fertilizer use and Carbon dioxide Emission with same factor loadings 0.935. The last factor (Factor 5) accounts for 6.26% of the total variation consists of only one variable, Cloud Coverage Vegetative Stage with loading 0.745.

Similarly another five influential factors to the HYV Boro rice production of Dinajpur district that account for 71.57% of the total variance are selected.

Among them the first factor accounts for the largest proportion, 31.04% of the total variance. This factor includes three variables like Humidity in Vegetative Stage, Fertilizer Use and Carbon Dioxide Emission with factor loadings 0.911, 0.805 and 0.807 respectively. Factor 2 accounts for 14.92%, the second largest proportion of the total variance consisting of only one variable Cloud Coverage in Growing Stage with factor loadings 0.702. The third factor accounts for 10.92% of the total variance consisting of two variables of Rainfall in Vegetative Stage and Rainfall in Growing Stage. The factor loadings are 0.721 and 0.637 respectively. Factor 4 accounts for the next 8.17% variation of the total variance, consisting only one variable, Wind Speed in Reproductive Stage with factor loading 0.657. Factor 5 is the last factor accounts for 6.52% of the total variation consisting of the variable Temperature in Growing Stage with its loading 0.691.

5.4 Policy Implication

The findings should be useful for providing a better understanding of the situations of the HYV Boro rice in the Northern area of Bangladesh. The results obtained from this study may be helpful for the planners and policy makers to allocate and achieve production targets and in long term planning at national level so that appropriate action can be taken as and when necessary.

References

- Abedullah and S. Pandey (2007). The Value of Rainfall Forecasts in the Rainfed Rice Areas of the Philippines. *The Lahore Journal of Economics* 12(2), pp. 69-81.
- Akaike, H. (1970). Statistical predictor identification. *Ann. Statis. Math.* 22, pp. 203-217.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19 (6): pp.716–723.
- Alam, M. S. Islam, M. R. Jabber, M. A. Islam M. S. and M. A Salam (2008). Institutional backup towards food security in Bangladesh, In *Proceedings BKAS 13th National Conference and Seminar on Climate Changes: Food Security in Bangladesh*, Vol. 13, Dhaka, Bangladesh.
- Alimba, J. O. and C. J. C. Akubuilu (2002). Assessing the consequences of technological change on farm enterprises in southeastern Nigeria by use of factor analysis. *Agricultural Systems*. 74, pp. 257-270.
- Anatoly, M. (1998). Analysis of Agriculture Production with Statistical Method. *Russian Federation, St. Petersburg*, 29 June - 3 July 1998.
- Angladette, A. (1966). P. Le Riz Maisonneuve and Larose, Paris, 930
- Anna, B. Costello and J.W. Osborne (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment Research & Evaluation*, Vol 10(7).
- Azad, M.A. Hasanuzzaman, M. Azizunnesal, Shil G.C. and M.A. Barik (2007). Milk Production Trend, Milk Quality and Seasonal Effect on it at Baghabarighat Milk Shed Area, Bangladesh. *Pakistan Journal of Nutrition* 6 (4): pp. 363-365, 2007, ISSN 1680-5194.

- Babakus, E., Carl E. J. Ferguson, and K.G. Joreskog (1987). The sensitivity of Confirmatory Maximum Likelihood Factor Analysis to violations of measurement scale and distributional assumptions, *Journal of Marketing Research*, 24 (May), pp.222-28.
- Bartholomew, David J., Fiona S., Irini M., and J. I. Galbraith (2002). *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Boca Raton, Florida: Chapman & Hall.
- Basak, J.K., Ali, M.A., Islam, M.N. and M.J.B. Alam (2009). Assessment of the effect of climate change on boro rice production in Bangladesh using CERES-Rice model, *Proceedings of the International conference on Climate Change Impacts and Adaptation Strategies for Bangladesh*, 18-20 February 2009, pp. 103-113.
- Cure, J.D. (1985). Direct Effects of Increasing Carbon Dioxide on Vegetation. (eds.). U.S. Dept. of Energy, Washington, D.C., Publication DOE/ER-0238, 1985, pp. 155-170.
- BBS (2006). Bangladesh Bureau of Statistics, Planning Division, Ministry of Planning, Government of the People's Republic of Bangladesh, Dhaka. *Statistical Year Book of Bangladesh*.
- BBS (2007). Bangladesh Bureau of Statistics, Planning Division, Ministry of Planning, Government of the People's Republic of Bangladesh, Dhaka. *Statistical Year Book of Bangladesh*.
- BBS (2008). Bangladesh Bureau of Statistics, Planning Division, Ministry of Planning, Government of the People's Republic of Bangladesh, Dhaka. *Statistical Year Book of Bangladesh*.
- Bethea, R. M., and R. R. Rhinehart (1991). *Applied Engineering Statistics*, Marcel Dekker, Inc., New York, NY.

- Bhagyashree, S.D. and A. Rajarathinam (2009). Statistical modeling on area, production and productivity trends of bajra crop grown in Gujarat state. *IJTA*, 27, pp. 291-296.
- Bhagyashree, S.D. (2009). Application of parametric and nonparametric regression models for area, production and productivity trends of major crops of Gujarat. M.Sc. Thesis, Submitted to *Anand Agricultural University*, Anand, Gujarat.
- Bhuiyan, N. I., D. N. R. Paul and M. A. Jabbar. (2002). Feeding the extra millions by 2025 challenges for rice research and extension in Bangladesh. A keynote paper presented on *national workshop on rice research and extension* held on 29-31 January 2002 at BRRI, Gazipur. p9.
- Bhuiyan, S. I. and A. L. Galang (1987). Impact of climatic factors on irrigated rice culture. Proc. Internet. Workshop on "The Impact of Weather parameters on Growth and Yield Rice". *IRRI, Los Banos, Laguna, Puilippines*. PP. 221-234.
- Bukantis, A. (2002). Application of factor analysis for quantification of climate-forming processes in the eastern part of the Baltic Sea region. *Clim Res*. Vol. 20: pp.135-140,
- Burnham, A. (1998). *Model Selection and Inference - A practical information-theoretic approach*. ISBN 0-387-98504-2.
- Carr, S.C. (1992). A primer on the use of Q sort technique factor analysis. *Measurement and Evaluation in Counseling and Development*, 25, 133-138.
- Challinor, A. J., Slingo, T. R. Wheeler, P. Q. Craufurd and D. I. F. Grimes (2003). Toward a combined seasonal weather and crop production forecast system: Determination of the working special case. *American Meteorological Society*. PP. 175-191.

- Comrey, A. L. and, Lee, H. B. (1992). *A First Course in Factor Analysis*. Lawrence Erlbaum associates Publishers, Hillsdale, NJ, USA.
- Coronel, V.P., Dai, Q.J. and B.S. Vergara (1990). Preliminary study on response of rice seedlings to enhanced UV-B radiation . *Intl. Rice Res. News*, 15:37.
- Cryer, J. D. (1986). *Time Series Analyses*, Duxbury Press, Boston, pp-1-5 Dahmen and Hall, 1990.
- DAE (2007). Directorate of Agricultural Extension, Government of the People's Republic of Bangladesh. Ministry of Agricultural, Dhaka.
- Dahmen, E. R. and M. J Hall (1990). Screening of Hydrological Data: Test for Stationarity and Relative Consistency, *Publication* 49, ILRI, Wasington, The Netherlands, pp.28.
- Dickey, D. A. and W. A. Fuller (1979). Distribution of the Estimates for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*. 74, Pp-427-431.
- Dillon, W.R. and M. Goldstein (1984). *Multivariate Analysis: Methods and Application*. John Willey & Sons, New York.
- DiStefano, C., Diane M. M., John M., Patricia M., and T. Dickenson (2007). Using School Climate Surveys to Categorize Schools and Examine Relationships with School Achievement, Paper presented at *the annual conference of the American Educational Research Association*, April, 2007, Chicago, Illinois.
- Evenson, R.E., C. Pray and M.W. Rosegrant, (1999). *Agricultural Research and Productivity Growth in India*. Research Report No. 109. Washington, D.C: *International Food Policy Research Institute*.
- FAO (2005): FAOSTAT database (available at www.fao.org).

- Finger, R. (2007). Evidence of slowing yield growth – the example of Swiss cereal yields. MPRA Paper No. 9475, posted 07. July 2008 / 17:02. Online at http://mpra.ub.uni-muenchen.de/9475/mpra.ub.uni-muenchen.de/9475/1/MPRA_paper_9475.pdf
- Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Gupta, D. S., Sahu, P.K., Sanyal, M. K. and S.R. Pal (1999). Growth and trend analysis of potato production in *West Bengal*, *J. Interacad*, 3 (3& 4), pp.345-350.
- Hafner, S. (2003). Trend in maize, rice and wheat yield 188 nations over the past 40years : a prevalence of linear growth. *Agriculture, Ecosystem and Environment* 97 pp. 275-283.
- Hossain, M. (1984). Increasing Food Availability in Bangladesh: Constraint and Possibilities. *A Report Prepared for the Ministry of Agriculture, Government of the People's Republic of Bangladesh, and ILO, Geneva.*
- Hossain, M., F. Naher and Q. Shahabuddin (2005). Food security and nutrition in Bangladesh: Progress and determinants. *Electronic journal of Agricultural and Development Economics* 2(2): pp.103-132.
- Huda, A.K. Samsul, Ghildyal, B.P., Tomar, V.S. and R.C. Jain (1975). Contribution of climatic variables in predicting rice yield, *Science Direct, Agricultural Meteorology* Volume 15, Issue 1, June 1975, pp. 71-86.
- Huda, M. Z. (2001). Regional development of irrigation technologies and its impact on food grain production in Bangladesh, Unpublished MS Thesis, Department of Agricultural Economics, BAU Mymensingh, Bangladesh.
- Islam, M.N., Ahmed, G.J.U, Panaullah, G.M. and M.A.R Gomasta. (2001). Crop-soil water management technology for rice production. In the *Proceedings*

of the workshop on modern rice cultivation in Bangladesh, 14-16 February 1999.

- Jabber, M. A. and R. P. Jones. (1997). The Growth of MV Rice Production and Adoption in Bangladesh. *The Bangladesh Journal of Agricultural Economics* 20(2), pp: 1-19.
- Blignaut, J., Liza Ueckermannb and J. Aronsonc (2009). Research Articles. Agriculture production's sensitivity to changes in climate in South Africa. *South African Journal of Science* 105, pp. 61-72
- Jarque, C.M., A.K. Bera (1987). A test for normality of observations and regression residuals. *International Statistical Review* 55 (2): pp.163–172. JSTOR 1403192.
- Johnson, R. A. and D. W. Wichern (2003). Applied Multivariate Statistical Analysis, Prentice-Hall of India.
- Karim, M.R., Awal, M.A. and M. Akter (2010). Forecasting of Wheat production in Bangladesh. *Bangladesh J. Agril. Res.* 35(1):pp. 17-28. ISSN 0258-7122.
- Karmokar, P. K. and A.H.M.R. Imon (2008). Treed Analysis of the production of rice in Bangladesh. *International Journal of Statistical Sciences*. Vol. 8, 2008, pp. 103-110. ISSN 1686-5603.
- Kendall, M. G and A. Stuart (1961). *The Advance Theory of Statistics*. Vol. 2, Hafner Publishing Company, New York. USA.
- Kim, J. (1970). *Factor analysis*. In: Nie, Hull, C. H., Jenkins, J. G., Steinberger, K., Bent, D.
- Kumar, P. and S. Mittal (2006). Agricultural Productivity Trends in India: Sustainability Issues. *Agricultural Economics Research Review*, Vol. 19. pp. 71-88.

- Kumar, B. L. (1978). Declining Trend in Production of Pulses and Factors Affecting It. JSTOR: Economic and Political Weekly. Vol. 13, No. 27 (Jul. 8, 1978), pp. 1112-1114 Published by: *Economic and Political Weekly* Stable URL: <http://www.jstor.org/stable/4366793>.
- Michaelides, M. (2007). Early Detection of Ovarian Cancer Using Fluorescence Spectroscopy and Parallel Factor Analysis. A thesis of the master of science. The University of Arizona.
- Mondal, M. H. (2010). Crop Agriculture of Bangladesh: Challenges and Opportunities. *Bangladesh Journal of Agricultural Research*, Vol 35, No 2. pp235-245.
- Mondal, M. A. S. (2008). World Food Security: The Challenges of Climate Change and Bioenergy Bangladesh Perspectives. Key note paper presented at the world food day 2008 seminar on "World Food Security: The Challenges of Climate Change and Bioenergy", organized by the Ministry of Agriculture, Govt. of the People's Republic of Bangladesh. BARC, Farmgate, Dhaka, Bangladesh.
- Munir, K, Mustafa K., Shah M., Naushad K. and Z. K. Jehan (2008). Forecasting Mango Production In Pakistan An Econometric Model Approach. *Sarhad J. Agric.* Vol.24, No.2, 2008. pp. 363-369.
- Nguyen, C. T. and B. Singh (2006). Trend In Rice Production and Export In Vietnam. *Omonrice* 14 pp. 111-123.
- Omobowale, O., Ephraim N., John P., Dayo, P. and E. Kato (2009). Trends and Drivers of Agricultural Productivity in Nigeria. Nigeria Strategy Support Program (NSSP). Report 001. December 2009. *International Food Policy Research Institute*. 2033 K Street NW. Washington, DC 20006-1002 USA. www.ifpri.org.

- Paul, D.G. and E. Meijer (2005). Time Series Factor Analysis with an Application to Measuring Money. University of Groningen, Research School SOM. Research Report 05F10
- Rajarathinam, A., R.S. P. and P.R. Vaishnav (2010). Estimating models for area, production and productivity trends of tobacco (*Nicotiana tabacum*) crop for Anand Region of Gujarat State, India. *J. Applied Sci.*, 10: 2419-2425.
- Sahu, P.K (2003). Production model: In search of alternative food. Presented in Professor K.M. Hossain *International seminar-cum-workshop, Department of Statistics, University of Rajshahi, Bangladesh*, October 2003.
- Sarma, A. A. L. N., Lakshmi, T. V. Kumar and K. Koteswararao (2008). Development of an Agroclimatic Model for the Estimation of Rice Yield. *Journal of the Indian Geophysical Union* 12(2): pp. 89-96.
- Sarma, M. (2005). Study on agricultural growth performance of Assam. *Economic. Affairs*, 50: pp.38-41.
- Sastri, A.S.R.A.S. and M. Serjius (2009). Regional climate changes and their impact on agriculture-A case study for Chhattisgarh State. www.cse.org.in/climate_research/pdf/
- Sastry, P.S.N.(1976). Climate and crop planning with particular reference to rainfall. In: *Climate and Rice, Proceedings of the Symposium, International Rice Research Institute*, 1976, pp. 5163.
- Singh, G. (1997). Data book on mechanization and agro processing in India after independence. Technical Bulletin CIAE/97/71, *Central Institute of Agricultural Engineering, Nabi Bagh, Bhopal, India*.
- Singh, G. (2000). Modernization of Agriculture in India. *Farm Mechanization. Agricultural Situation in India*, Ministry of Agriculture, New Delhi.

- Singh, G. (2006). Estimation of a mechanization index and its impact on production and economic factors a case study in India. *Biosystems Engineering*, 93: 99–106.
- Singh, G. and H. Chandra (2002). Production and economic factors growth in Indian agriculture. *Technical Bulletin No. CIAE/ 2002/91, Central Institute of Agricultural Engineering, Nabi Bagh, Bhopal, India.*
- Singh, G., and D. De (1999). Quantification of a mechanization indicator for Indian agriculture. *Applied Engineering in Agriculture*, 15: pp. 197–204.
- Singh, R.B. (2000). Environmental consequences of agricultural development: A case study from the Green Revolution state of Haryana, India. *Agriculture, Ecosystems and Environment*, 82: pp. 97–103.
- Singh, P.K., Hossain, M. and R. Thakur (2003a). In: Boro rice (RK Singh, M Hossain and R Thakur, eds). *IRRI-India Office.*
- Singh, R.K., Thakur, R. and S.D. Chatarjee. (2003b). Harnessing boro rice potential for increasing rice production in deepwater areas of Eastern India, An overview. In: Boro rice (RK Singh, M Hossain and R Thakur, eds). *IRRI-India Office.* Pp. 51-72.
- Singh, U.P. (2002). Boro Rice in Eastern India. Rice-Wheat Consortium Regional Technical Coordination Committee seminar. 10-14 February 2002. *Rice-Wheat Consortium for the Indo-Gangetic Plains*, New Delhi, India.
- Smith, C., Wendy G., Hammond J. and L. Brooks (2006). The Application of An Exploratory Factor Analysis To Investigate The Inter-Relationships Amongst Joint Movement During Performance of A Football Skill. *The 8th Australasian Conference on Mathematics and Computers in Sport*, 3-5 July 2006, Queensland, Australia. Published (online): 15 December 2006
- Sneyer, R. (1990). On the Statistical Analysis of Series of Observations. *WMO No. 415, Technical Note No. 143*, Geneva.

- Tao, F., M. Yokozawa, J. Liu, and Z. Zhang (2008). Climate-Crop Yield Relationships at Provincial Scales in China and the Impacts of Recent Climate Trends. *Climate Research* 38: 83-94.
- Tsuyoshi, et. al., (2007). Factor Analysis of Drug Supply Time series at Pharmacies. *Journal of Health Science*, 54(1), pp- 107-111.
- Virmani, S. M. (1987). Climatic analyses and cropping systems in the semiarid tropics. Weather and Rice: Proceeding of the International Workshop on *the Impact of Weather Parameters on Growth and Yield of Rice*; Publ. of IRRI, Philippines, 1987, pp. 215-220.
- Virmani, S.S., Viraktamath, B.C., Casal, C.L., Toledo, R.S., Lopez, M.T., and J.O. Manalo (1997). *Hybrid rice breeding manual*. IRRI Philippines: 7-8
- WARDA (2005). West Africa Rice Statistics Data Bank. *Rice Trends in Sub-Saharan Africa* - The Africa Rice Center
- Wilson, C.E. J. and J.W. Branson (2004). Trends in Arkansas Rice Production. *AAES Research Series* 529. B.R. Wells Rice Research Studies.
- Xinan, D., Yuzhou L., Suocheng D. and Y. Xiusheng (2005). Impact of resources and technology on farm production in northwestern China. *Agricultural Systems*, 84: pp. 155-169.
- Zare, F. A. A., Kouchaki A. R. and M. M. Nasiri (2006). Trend analysis of yield, production and cultivated area of cereal in Iran during the last 50 years and prediction of future situation. 13. *Iranian Journal of Field Crops Research* 2006; 4(1): pp. 49-69.