Journal of Business and Management Sciences, 2015, Vol. 3, No. 1, 26-31
Available online at http://pubs.sciepub.com/jbms/3/1/4

DOI:10.12691/jbms-3-1-4

A Prediction on Rice Production in India through Multivariate Regression Analysis

Somu Jena*

Department of Business Administration Utkal University, Bhubaneswar India-751004
*Corresponding author: somujena1000@gmail.com

Received March 10, 2015; Revised April 02, 2015; Accepted April 13, 2015

Abstract In this paper an extensive study is carried out to estimate the rice production in India based on current and historical data. The significant factors studied are land used, irrigation and production respectively. To study the strength of interdependence between the factors and estimation of production multivariate correlation analysis and regression analysis have been applied. To predict the future production of rice different time series techniques have been studied for better comparison. To facilitate the entire study and maintain quality, consistency and accuracy of data SPSS- 20.0 version software is used.

Keywords: *rice production, land used, irrigation, multivariate regression analysis, time series analysis*

Cite This Article: Somu Jena, "A Prediction on Rice Production in India through Multivariate Regression Analysis." *Journal of Business and Management Sciences*, vol. 3, no. 1 (2015): 26-31. doi: 10.12691/jbms-3-1-4.

1. Introduction

Rice has more than forty thousand varieties species all across the globe. It is the staple food of the country India. The origin of rice is buried in obscurity and the depths of time. Till today, we do not know when it was first discovered and domesticated and perhaps this is one fact we will never come to know. In the long and turbulent history of the human race, one of the most important developments that led to the development of civilizations was the domestication of rice, for this one single variety of grain has fed and nourished more people over a longer period of time than any other crop.

Rice or *Oryza Sativa* (as botanists prefer to call it) is not a tropical plant but is still associated with a wet, humid climate. It is generally believed that the domestication of rice began somewhere in the Asian arc. From its place of birth, lost forever in the mists of time, the plant and its grain spread all over the world. According to some schools of thought, It is probably a descendent of wild grass that was cultivated in the foothills of the Eastern Himalayas and the upper tracts of the Irrawady and Mekong river basins. Another school of thought believes that the rice plant may have originated in southern India and then spread to the north of the country.

From India, the plant spread to China and then onwards to Korea, the Philippines (about 2000 B.C.), Japan and Indonesia (about 1000 B.C.). The Persians are known to have been importers of this grain. From there its popularity spread to Mesopotamia and Turkestan. It is believed that when Alexander the Great invaded India in 327 B.C., one of the priced possessions he carried back with him was rice.

Arab travelers took it to Egypt, Morocco and Spain and from there it traveled all across Europe. The Portugese and Hollanders took rice to their colonies in West Africa. From Africa it traveled to America through the 'Columbian Exchange' of natural resources - rice being a gift from the Old World to the New. Rice has been cultivated in the United States of America for the last three hundred years.

It was in China that the process of puddling soil and transplanting seedlings was likely refined. With the development of puddling and transplanting, rice became truly domesticated. In China, the history of rice in the river valleys and low-lying areas is older than its history as a dry land crop. In Southeast Asia, by contrast, rice was originally grown under dry land conditions in the uplands, and later it came to occupy the vast river deltas. Migrants from South China or perhaps South East Asia carried the traditions of wetland rice cultivation to the Philippines during the second millennium B.C., and Deutero - Malays may have carried the practice to Indonesia about 1500 B.C. From China or Korea, the crop was introduced to Japan around 100 B.C.

According to some schools of thought, It is probably a descendent of wild grass that was cultivated in the foothills of the Eastern Himalayas and the upper tracts of the Irrawady and Mekong river basins. Another school of thought believes that the rice plant may have originated in southern India and then spread to the north of the country.

The journey of rice around the world has been slow, but once it took root it stayed and became a major agricultural and economic product for the people. In the Indian subcontinent more than a quarter of the cultivated land is given to rice. It is a very essential part of the daily meal in many parts of the country. The rice grain is treated with honour in the subcontinent and in Asia: for here the failure of the rice crop in not only an economic setback but can also create a famine-like situation. Wastage of rice is viewed rather badly in these societies and superstitions about the grain abound.

2. Objective

To predict the future production of rice from the available information and to study the relationship between the various factors affecting the rice production.

3. Data Analysis

From Figure 1, Figure 2 and Figure 3 data related to production, land used and irrigation with respect to the time period that is from the year 1950 to 2012 is shown.

3.1. Time Series Analysis

From this historical data it is easy to predict for future by using least square method of straight line and parabolic method and the interpretations are derived through the comparative study. Table 1 represents the $R^2 = 0.951$ which shows that 95% of the variation in production in India can be explained by the time factors or independent variable. Table 2 is the ANOVA where the significance of F is 0.000. This indicates that the model is statically significant at a confidence level 0f 95%. Table 3 shows the values of the trend line $y = 58.856 + 1.398x$, where y is production and x is the time factor and it shows these two variables are statistically significant for the trend model as P-values are less than 0.05.

Table 4 represents the $R^2 = 0.968$ which shows that 97% of the variation in production of rice can be explained by the time factor or independent variable. Table 5 is the ANOVA where the significance of F is 0.000. This indicates that the Parabolic model is also statistically significant at a confidence level 0f 95%. Table 6 shows the values of the Parabolic trend line $y = 30.039 + 2.417x - 0.073x^2$, where y is tourism and hospitality contribution and x is the time factor and it shows these two variables are statistically significant for the trend model as P-values are less than 0.05. Figure 8 represents the curve fitting of trend line and parabolic curve of Indian Tourism and Hospitality Contribution to GDP CAGR with respect to the time period.

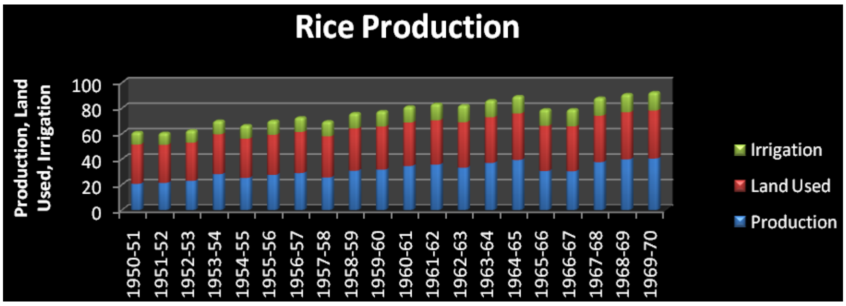

Figure 1. Actual Data of Rice Production, Land Used and Irrigation from the Year 1950-1969

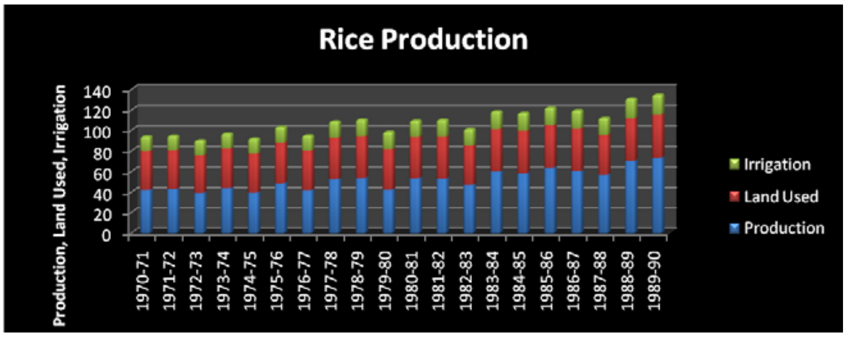

Figure 2. Actual Data of Rice Production, Land Used and Irrigation from the Year 1970-1989

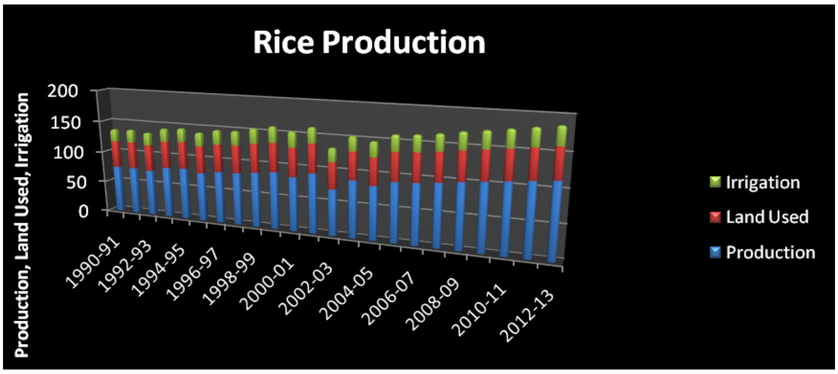

Figure 3. Actual Data of Rice Production, Land Used and Irrigation from the Year 1990-1912

Table 1. Model Summary of Fitting of a Straight Line of Production of Rice with respect to the Year from 1950-2012

R	R Square	Adjusted R Square	Std. Error of the Estimate
.975	.951	.950	5.862

The independent variable is time.

Table 2. ANOVA of Least Square Method of a Straight Line

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	40705.656	1	40705.656	1184.680	.000
Residual	2095.963	61	34.360		
Total	42801.619	62			

The independent variable is time.

Table 3. Coefficients of a Straight Line

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Time	1.398	.041	34.419	34.419	.000
(Constant)	58.856	.739	79.695	79.695	.000

Table 4. Model Summary of Fitting of a Parabolic Curve of Production of Rice with respect to the Year 1950-2012

R	R Square	Adjusted R Square	Std. Error of the Estimate
.984	.968	.967	4.745

The independent variable is time.

Table 5. ANOVA of Least Square Method of a Parabolic Curve

	Sum of Squares	df	Mean Square	F	Sig.
Regression	41450.952	2	20725.476	920.678	.000
Residual	1350.667	60	22.511		
Total	42801.619	62			

The independent variable is time.

Table 6. Coefficients of a Parabolic Curve

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Time	1.398	.033	.975	42.523	.000
time ** 2	.012	.002	.132	5.754	.000
(Constant)	55.009	.897		61.337	.000

Table 7. Model Summary of Fitting of a Straight Line of Land Used with respect to the Year from 1950-2012

R	R Square	Adjusted R Square	Std. Error of the Estimate
.958	.918	.917	1.298

The independent variable is time.

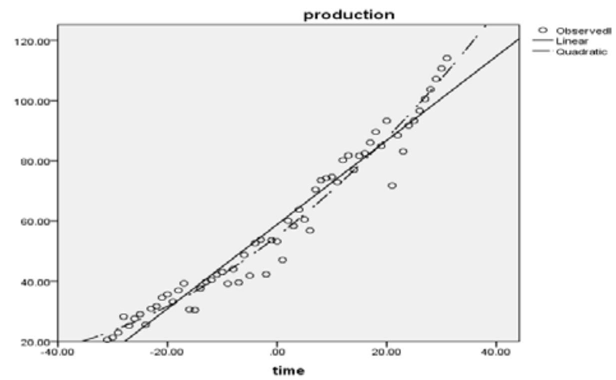

Figure 4. Fitting of Straight Line and Parabolic Curve with Original Data of Production with respect to the Year 1950-2012

Table 8. ANOVA of Least Square Method of a Straight Line

	Sum of Squares	df	Mean Square	F	Sig.
Regression	1153.725	1	1153.725	684.703	.000
Residual	102.785	61	1.685		
Total	1256.510	62			

The independent variable is time.

Table 9. Coefficients of a Straight Line

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Time	.235	.009	.958	26.167	.000
(Constant)	39.220	.164		239.813	.000

Table 10. Model Summary of Fitting of a Parabolic Curve of Land Used with respect to the Year from 1950-2012

R	R Square	Adjusted R Square	Std. Error of the Estimate
.980	.961	.959	.908

The independent variable is time.

Table 11. ANOVA of Least Square Method of a Parabolic Curve

	Sum of Squares	df	Mean Square	F	Sig.
Regression	1207.018	2	603.509	731.632	.000
Residual	49.493	60	.825		
Total	1256.510	62			

Table 12. Coefficients of a Parabolic Curve

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Time	.235	.006	.958	37.399	.000
time ** 2	-.003	.000	-.206	-8.038	.000
(Constant)	40.248	.172		234.444	.000

Table 13. Model Summary of Fitting of a Straight Line of Irrigation with respect to the Year from 1950-2012

R	R Square	Adjusted R Square	Std. Error of the Estimate
.986	.972	.971	.851

The independent variable is time.

Table 14. ANOVA of Least Square Method of a Straight Line

	Sum of Squares	Df	Mean Square	F	Sig.
Regression	1517.965	1	1517.965	2095.276	.000
Residual	44.193	61	.724		
Total	1562.157	62			

The independent variable is time.

Table 15. Coefficients of a Parabolic Curve

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Time	.270	.005	.986	53.308	.000
time ** 2	.001	.000	.088	4.768	.000
(Constant)	17.063	.138		123.514	.000

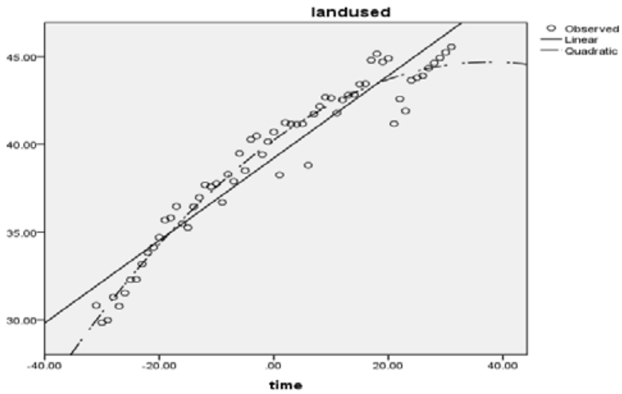

Figure 5. Fitting of Straight Line and Parabolic Curve with Original Data of Land Used with respect to the Year 1950-2012

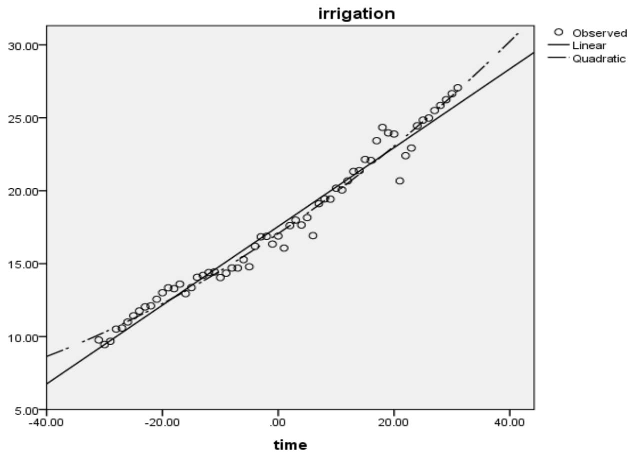

Figure 6. Fitting of Straight Line and Parabolic Curve with Original Data of Irrigation with respect to the Year 1950-2012

3.2. Comparative Study

Fitting of a Straight Line and Parabolic Method of Production, Land Used and Irrigation with respect to the Time Period from 1950-2012 are shown in the Figure 4, Figure 5 and Figure 6 respectively. From these figures it is concluded that for prediction parabolic method of least square is more flexible and appropriate than straight line method of least square.

3.3. Prediction of Rice Production, Land Used and Irrigation for the next 10 Years

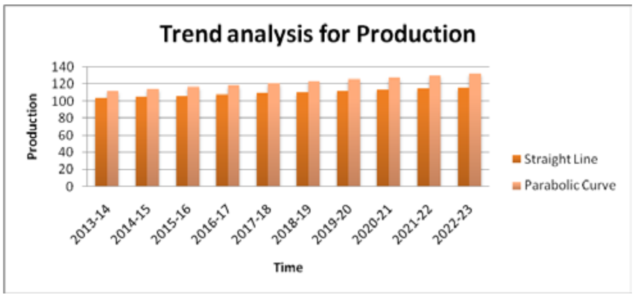

Figure 7. Prediction of Rice Production for the next 10 Years

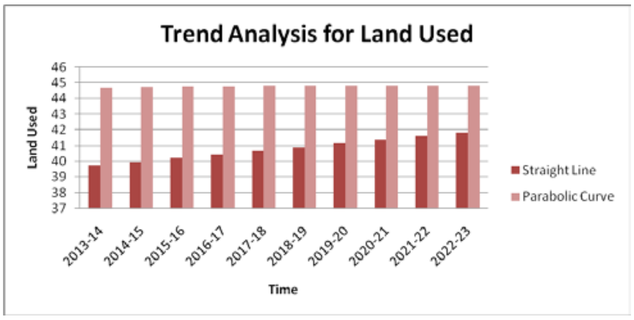

Figure 8. Prediction of Land Used for the next 10 Years

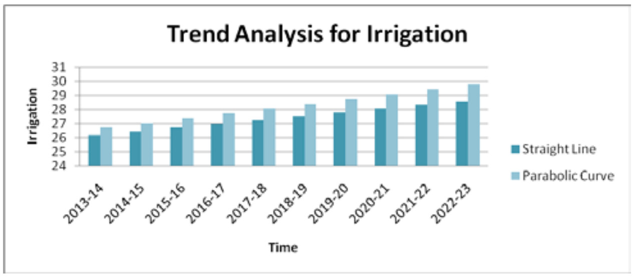

Figure 9. Prediction of Irrigation for the next 10 Years

Figure 7, Figure 8 and Figure 9 show the predicted values of production of rice in India, land used and irrigation in next 10 years respectively by using trend analysis and parabolic curve fitting respectively.

3.4. Multivariate Correlation and Regression Analysis

Hypothesis

***H*₀₁:** Production, land used and irrigation are uncorrelated.

***H*₁₁:** Production, land used and irrigation are correlated.

Level of Significance: α% is 5%.

Table 16. Correlation of Production, Land used and Irrigation

		Production	landused	Irrigation
Production	Pearson Correlation	1	.934**	.990**
	Sig. (2-tailed)		.000	.000
	N	63	63	63
Landused	Pearson Correlation	.934**	1	.950**
	Sig. (2-tailed)	.000		.000
	N	63	63	63
Irrigation	Pearson Correlation	.990**	.950**	1
	Sig. (2-tailed)	.000	.000	
	N	63	63	63

**. Correlation is significant at the 0.01 level (2-tailed).

3.5. Data Interpretation

From Table 16 it is observed that production and land used are positively correlated with high degree, production and irrigation are also positively correlated with high degree and land used and irrigation are positively correlated with high degree. Since all P values are less than 0.05, so it may be concluded that all the variables like production, land used and irrigation are correlated with each other as we may reject the null hypothesis. So it is significant.

Table 17. Model Summary of Regression Analysis

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.991[a]	.981	.981	3.64629	.981	1579.633	2	60	.000	1.493

a. Predictors: (Constant), irrigation, landused
b. Dependent Variable: production

Table 18. ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	42003.891	2	21001.946	1579.633	.000[b]
	Residual	797.727	60	13.295		
	Total	42801.619	62			

a. Dependent Variable: production
b. Predictors: (Constant), irrigation, landused

Table 19. Coefficients of Multivariate Regression Analysis

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-21.937	8.166		-2.686	.009	-38.271	-5.602
	landused	-.421	.329	-.072	-1.278	.206	-1.080	.238
	irrigation	5.543	.295	1.059	18.766	.000	4.952	6.133

Dependent Variable: production

Hypothesis

H_{02}: The model is statistically insignificant
H_{12}:: The model is statistically significant
Level of Significance: α% =5%.

3.6. Interpretation

Table 17 represents the $R^2 = 0.981$ which shows that 98.1% of the variation in production of rice in India for 63 years can be explained by the independent factors like land used and irrigation. Table 18 is the ANOVA where the significance of F is 0.000. This indicates that the model is statically significant at a confidence level of 95%.

Table 19 shows the multivariate regression model $y = -21.937 - 0.421x_1 + 5.543x_2$, where y is the production of rice in india, x_1 is the independent variable like land used and x_2 is other independent variable like irrigation. It shows the one variable like irrigation is statistically significant for the regression model as P-value is less than 0.05 but other variable like land used is statistically insignificant for the regression model as P-value is more than 0.05. The above analysis clearly reveals that the regression model fits well to this concern case. Hence using this model the future production level of rice in India can be easily predicted.

4. Conclusion

Rice is the main grain crop in India. India ranks second in the world in the production in the rice. Rice production is dependent mainly on fertile land, irrigation, suitable climate, use of fertilizers and manures, precipitation, temperature, quality of seeds, technology modern equipments for cultivation and man power. All these factors, if stand to be effective, it can contribute for an increased production of rice. Apart from this emphasis must be given on the proper management of the food grains with safety storage ensured. If the production level of rice increases then it will help in the economy growth of the country and the unutilized resources can be explored. Figure 10 shows the rice growing states of India. India can envisage a better scenario of agrarian economy with an enhanced production of rice. Steps should be taken to engage more man power for rice production.

4.1. Limitations

Due to insufficient data, only two factors of production were taken for studying.

The impact of technological advancement and use of high quality seeds were ignored.

References

[1] Afifi, A., May, S., Practical Multivariate Analysis, Fifth Edition.
[2] All India Rice Exporters Association (AIREA)
[3] Hair, J.F., Ronald L.T., Anderson, R.E., William B., Multivariate Data Analysis.
[4] Levin R.I, Rubin D.S., Statistics for Management, 7th Edition, Pearson, India, 2008.
[5] Nargundkar, R.N., Marketing Reasearch, TMH, Third Edition, 2011.
[6] Pattnaik, M. Inventory Models: A Management Perspective, Lambart Academic Publishing, Germany, 2014.
[7] Richard, A. J., Dean W.W., Applied Multivariate Statistical Analysis.

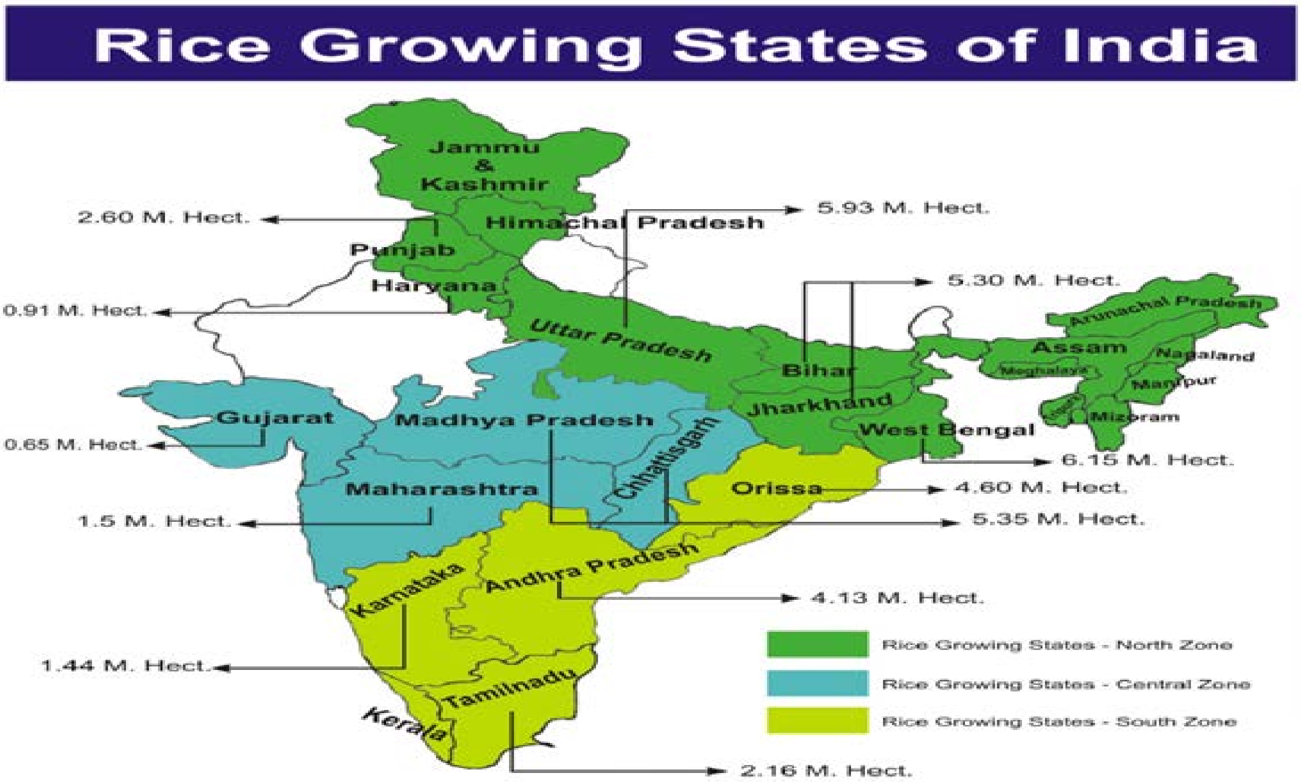

Figure 10. Rice Growing States in India