

# Statistical Method to predict Agricultural CPI

Yash Kumar Arora, Santosh Kumar, Umesh Kumar Tiwari, Aditi Goswami and Jasmeet Kalra

**Abstract:** *The growth of any country depends on its economy and economic growth is nothing but an increase in the inflation i.e. adjusted market value of the goods and services produced by an economy over time. Statisticians conventionally measure such inflation using the price indices. They are mainly WPI (Wholesale Price Index and CPI (Consumer Price Index). WPI is now known to be an older method of computation because the main focus has to be on consumer prices. CPI is a measure of consumer prices over a certain period. Changes in the CPI are used to assess price changes associated with the cost of living. It can be calculated for rural, urban areas as well as for both. In CPI rural, the workers and labourers are benefitted as their daily wages can be predicted by this approach. The CPI by state data represents the inflation of each of the states giving a concise view of the country. The data is collected and analysed using a mathematical approach called linear regression in future prediction for rural labours based on previous data.*

**Index Terms:** CPI, Inflation, Linear Regression, RMS

## I. INTRODUCTION

A versatile country of different states and union territories depends upon its economy and one's economy depends upon the inflation rate which is in turn the measure of Consumer Price Index (CPI). Price stability is one of the important goals of any policy formation as it helps in promoting growth.

Inflation can be defined as a percentage of change in prices of goods consumed over a certain period. The previously used index to determine inflation was the Weighted Price Index (WPI) but the problem with WPI was it included the prices in which the manufacturer provided the goods to wholesalers whereas the main idea should be related to customer prices only. Therefore, in the year 2012 consumer price index was chosen as a standard method for calculating inflation. CPI is a measure which calculates the percentage change in the general prices of a basket of goods and services consumed by the household.

CPI can be categorized or separated in CPI-Rural, CPI-Agricultural, CPI-Combined. CPI has six components

- Housing -10.07%
- Food and Beverages- 45.86%
- Clothing and footwear - 6.53%
- Tobacco etc- 2.38%
- Fuel, light etc -6.84%

**Revised Manuscript Received on October 20, 2019.**

**Yash Kumar Arora**, Computer Science and Engineering, Graphic Era deemed to be University, Dehradun, India.

**Santosh Kumar**, Computer Science and Engineering, Graphic Era deemed to be University, Dehradun, India.

**Umesh Kumar Tiwari**, Computer Science and Engineering, Graphic Era deemed to be University, Dehradun, India.

**Aditi Goswami**, Computer Science and Engineering, Graphic Era deemed to be University, Dehradun, India.

**Jasmeet Kalra**, Mechanical Engineering, Graphic Era Hill University, Dehradun, India.

- Miscellaneous-28.32%

The calculation of CPI is very rigorous as it includes many categories and sub-categories so after calculating sub-indices and indices a fully overall index is determined by national agencies. CPI thus used by the government and other agencies as a promoter of economic stability and growth.

CPI for agriculture workers and rural labourers were framed taking as 1986-87 as the base year. CPI is based upon its index value i.e. base index value which is 100 that means if the reading is 175 it indicates a rise in the level of Inflation as 75%. The inflation rate for an ideal economic growth has to be 2-4%. Inflation reduces the real income of those whose incomes are fixed and increases the real before-tax income of those whose income are variable. Computing the inflation over allocations of goods becomes an easy task by using a statistical model that's why the linear regression model has been used [1]. The data used for predicting the future CPI should be trustworthy as with this available past data future data is being calculated.

## II. RELATED WORK

The economy for a country depends on the inflation rate of each state individually.

In research by Dennis R. Starleaf [2], he stated that inflation is the increase in the prices of goods in general. As inflation increases over zero percent, we expect to see a rapid rise in the price of a commodity. It affects the real income of those whose incomes are fixed and tends to increase real before-tax income of those whose income is variable. In a journal by Deepika Goel [3], it was shown how RBI focused on WPI for measuring headline inflation but it had some limitation and therefore a new CPI was formed for giving a combined stats of Indian rural and urban households CPI-IW was used and the study shows the complete divergence between the two indices. An article in 2018 showed that inflation turned to be lower than expected in 2018, primarily due to the moderation of food products, while inflation in other products has been higher.

In an Indian Journal of Economics and Development [4], it was tried to find out the relationship between inflation and agriculture production and vice versa. To examine this cointegration technique is used which is used to testing co-relation between non-stationary time series variable.

Stabilizing food inflation has emerged a major challenge for handling economies over the past decade. An IIBM Management review [5] showed to understand the shocks of mark-up due to seeking activities of an agent at wholesaler and retailer level which can lower wholesaler and retail food inflation in the country.



# Statistical Method to predict Agricultural CPI

JoydebSasmal [6] states that the increase in per capita income has significantly increased the demand for food but agricultural production has failed to keep pace with rising demand. So, a study of CPI for agriculture becomes vital and it co-relates to rural laborers as the income or their wages also play a handy role in the country economy because both are getting down and are failing to meet demands.

Robert A. Pollak[7] tells CPI is often used as a cost-of-living index and it was recommended to take the cost-of-living index as an objective to calculate the consumer price index. The research shows many different indices for constructing political problems of CPI and industrial taxes. Hence, CPI was weighted as a long-run task turned to be somewhat confusing. In an annual report of 2011-2012[8], a brief discussion over agricultural and rural labor was done and the Labor Bureau has been compiling CPI numbers to get a separated series of CPI numbers of 20 states based on weighted diagrams. Therefore, a statistical approach was not so used for predicting the CPI in future reference based on past data.

Kirti et al. [9] reviews the paper for the calculation of wholesaler and consumer price indexes and also tells that a country like India where inflation expectations demands are less driven compared to more developed countries. They also came to a suggestion that the efficiency of the discovery mechanism which directly correlates to the financial sector should be improved. Akash et al. [10] in their research stated that India for a very long time is struggling to keep food inflation within political and sustainable levels. As India is a developing country and with increasing, population and income food demand will increase soon. Food expenditure constitutes 40% of total household expenditure, food inflation thus having a cumulative effect on the Economy. There are determinant factors for food inflation, so in this research, a machine learning technique or approach i.e. gradient boost regression trees were used which to a point predicted the significant change in food prices, but overall consumer price index was still a handy task.

Himanshu et al. [11] in their collective research came across that higher food inflation can be broadly divided into two factors i.e. supply and demand factors. For years spatial distribution of monsoon rains has affected agriculture on a large scale plus in the long run shifting of land for commercial crops had a great effect too. Increasing demand particularly arising from rural wages, rising agricultural cost production, and many more factors led to higher food inflation. But in the long-run increase in real wages have the largest influence in food inflation which resulted in a reason to increase agricultural productivity with raise in real wages.

In this paper, we used a statistical model for dealing with numerical data i.e. linear regression, we first divide the dataset in training and testing in the ratio 80:20 respectively. Then made use of root mean squared error to reduce the cost function. It's time to fit the curve and then finding the intercept and slope of the curve to CPI for the near future.

## III. METHODOLOGY

### A. Linear Regression

Linear regression can be understood as a linear model that

is based on two variables i.e. one dependent and another single independent predictor variable. The output behavior of independent predictor variable changes as the input behavior of dependent variable changes. It can be easily understood by the simple eq.1.

$$y = b_0 + b_1 * x \quad (1)$$

where x is an independent variable, y is a dependent variable,  $b_0$  is constant term or independent and  $b_1$  is the coefficient of relationship between x and y. The equation can also be understood in the form of a slope-intercept form.

### B. Cost Function

Cost function is the difference between the actual value and predicted value. The cost function is defined by eq. 2

$$J(\theta) = \frac{1}{2m} \sum (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (2)$$

where m is the number of states.

Sigma ( $\sum$ ) means the sum and the actual calculation is the square of the hypothesis value of  $h(x)$ , minus the actual value of y. The final result is a single number.

### C. Algorithm

Linear regression is a supervised learning approach, finding a relationship between the independent and dependent variable that gives the best fit line. It performs a task or its prediction to gives the value of the dependent variable (y) based on the independent variable (x). So, we start by giving an input variable or input data to x and labels of data to y. Then, first, we find the hypothesis function by calculating values of

$$\begin{aligned} \theta_1 &= \text{intercept} \\ \theta_2 &= \text{coefficient of } x \end{aligned}$$

which will give the best fit line and once we reach there, our model will predict y for x. The model aims to find or analyze the minimum error between the predicted value and the true value so it is important to update the values of  $\theta_1, \theta_2$  where cost function comes in play i.e. cost function of linear regression is the root mean squared error and then at the end of the algorithm, our goal is to minimize a cost function that uses gradient descent method so the idea starts with random  $\theta_1, \theta_2$  values, then updating the values and reaching to a minimum cost. The algorithm steps are:

Step 1	Input Data
Step 2	Data Splitting as 80% train and rest test
Step 3	Use of classifier model (to minimize cost function)
Step 4	Plot data in the y-axis (Rate of CPI) and x-axis (Year)
Step 5	Calculate slope and intercept from the plot
Step 6	Predict CPI for agriculture and rural laborers for future reference.

#### IV. DATASET

Dataset used in this work has been taken from an open government data platform where data is regulated and monitored by various national agencies. Fig 1 represents the increase in CPI for agricultural laborers over a certain period taking 2003 as base year. Fig 2 represents top wise CPI for different states as per data collected until August 2016. We can see how CPI rises or falls with time showing the annual growth per month in each state giving a concise annual growth for a year. Sikkim is ruling the CPI column.

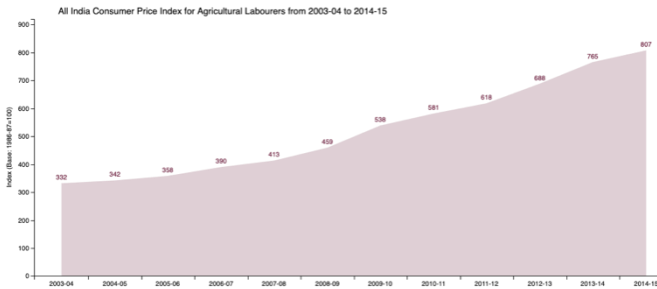


Fig. 1 Agricultural CPI over period of time

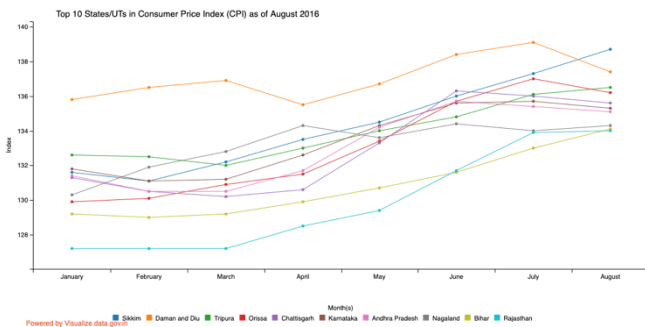


Fig. 2 State-wise CPI

#### V. RESULT ANALYSIS AND FUTURE PREDICTION

For taking consideration of the economy, it is vital for taking care of Inflation as it is a measure of CPI, and an increase or decrease of CPI graph annually depends on how commodities take a rise or decrease with comparison to base value as 100.

As per algorithm prediction based on previous data sets the agriculture CPI for 2020 will be 1052.2 and 2025 will be 1280.4

Fig 3 shows how agricultural CPI for rural laborers vary over a certain period of time.

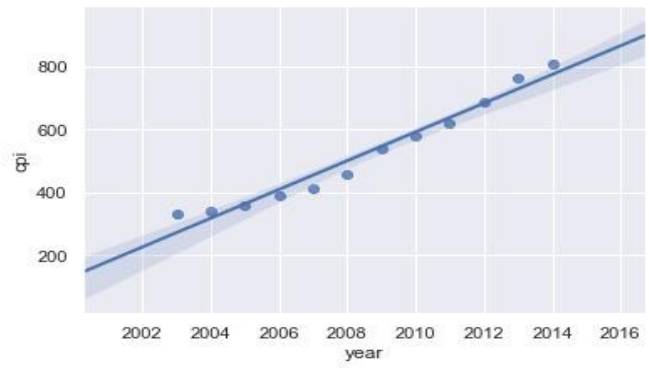


Fig. 3 Agricultural CPI year wise

#### VI. CONCLUSION

This paper deals with the fundamental process of determining CPI using a statistical approach using linear regression for predicting future values, as it would benefit rural laborers in determining their wages taking into consideration each state individually. As per capital, CPI is a determining factor for a country's economy.

By using our proposed algorithm, we predicted the values of agricultural CPI for rural laborers for the years 2020 and 2025. So, it can be easily cleared how the slope will take rise or fall and how the government should act in the field of agriculture in maintaining the needs of customer to handling the state economy contributing to the outlook of the country.

#### REFERENCES

- Arora, Yash & Kumar, Santosh. (2020). Statistical Approach to Predict Road Accidents in India. 10.1007/978-981-32-9515-5\_18
- Dennis R. Starleaf, The Impact of Inflation on farmers and agriculture, IOWA State University, ISU Economic Report Series.
- Goel, D. (2018). Measures of Inflation in India. Journal of Business Thought, 9, 24–45. <https://doi.org/10.18311/jbt/2018/21171>
- Aheibam, M., & Singh, R. (2017). Inflationary Trends and Agricultural Production in India: An Empirical Analysis. Indian Journal of Economics and Development, 13(3), 533. <https://doi.org/10.5958/2322-0430.2017.00212.8>
- IIBM Management review, measuring the contribution of mark-up shock in Food Inflation in India, Volume 31, June 2019
- Sasmal, J. (2015). Food price inflation in India: The growing economy with sluggish agriculture. Journal of Economics, Finance and Administrative Science, 20(38), 30–40. <https://doi.org/10.1016/j.jefas.2015.01.005>
- Pollak, R. A. (1998). The Consumer Price Index: A Research Agenda and Three Proposals. Journal of Economic Perspectives, 12(1), 69–78. <https://doi.org/10.1257/jep.12.1.69>
- Labor Bureau, Annual Report, Consumer Price Index Numbers for Agricultural and Rural Laborers, 2011-2012
- Gupta K and Siddiqui F (2014), "Salient Features of Measuring, Interpreting and Addressing Indian Inflation", Indian Council for Research on International Economic Relations, Working Paper No. 279.
- Malhotra, Akash & Maloo, Mayank. (2017). Understanding Food Inflation in India: A Machine Learning Approach. SSRN Electronic Journal. 10.2139/ssrn.2908354.
- Sonna, T., H. Joshi, A. Sebastian, and U. Sharma. 2014. Analytics of food inflation in India. RBI Working Paper, October, Reserve Bank of India, Mumbai.

#### AUTHORS PROFILE



**Yash Kumar Arora** am pursuing BTech in Computer science with specialisation in Big Data & Analytics from Graphic Era deemed to be University, Dehradun.



## Statistical Method to predict Agricultural CPI

With, interest in research, I want to research more and more to gain knowledge that can be used efficiently. Currently, I am an intern at IBM India Software Labs, Bengaluru.



**Dr. Santosh Kumar** had received his Ph.D. from IIT Roorkee (India) in 2012, M. Tech. (CSE) from Aligarh Muslim University, Aligarh (India) in 2007 and B.E. (IT) from C.C.S. University, Meerut (India) in 2003. He has more than 13 years of experience in teaching/research of UG (B. Tech.) and PG (M.Tech.) level courses as a Lecturer/Assistant Professor/

Associate Professor in various academic /research organizations. He has supervised 01 Ph.D. Thesis, 20 M.Tech. Thesis, 18 B.Tech projects and presently mentoring 06 Ph.D students, 03 M.Tech students and 04 B.Tech. students. He has also completed a consultancy project titled “MANET Architecture Design for Tactical Radios” of DRDO, Dehradun in between 2009-2011. He is an active reviewer board member in various national/International Journals and Conferences. He has memberships of ACM (Senior Member), IEEE, IAENG, ACEEE, ISOC (USA) and contributed more than 46 research papers in National and International Journals/conferences in the field of Wireless Communication Networks, Mobile Computing and Grid Computing and software Engineering. Currently holding position of Associate professor in the Graphic Era Deemed to be University, Dehradun (India). His research interest includes Wireless Networks, MANET, WSN, IoT, and Software Engineering.



**Aditi Goswami** am pursuing Btech - Computer science with specialisation in Cloud computing from Graphic Era Deemed to be University Dehradun.