Noname manuscript No.
(will be inserted by the editor)

Change in structural patterns in the time series of Onion prices in India

Saksham Bassi [1] · **Atharva Gomekar** [1] · **A. S. Vasudeva Murthy** [1][2]

Received: date / Accepted: date

Abstract The primary aim in this study is analysing the structural patterns of time series of Onion retail prices in the city of Mumbai. The retail prices of Onions for various cities were accessed through the Open Government Data (OGD) Platform India website. We implement various kinds of analysis techniques like statistics, curve fitting, mathematical functions to conclude about the patterns and periodicity in the time-series. We show there is a seasonality break in prices of Onion after 2010 indicating the role of unknown external factors not seen up till 2010.

Keywords Seasonality · Structural patterns · Statistics

1 Introduction

Time series of commodity prices have structures or patterns due to various influential factors. Fair amount of insights about these factors can then be drawn from the patterns found. Characteristics of commodity prices were done in the early past with simulation study (see [1]) to examine the problems faced by the properties of these time series. Generally, primary commodities have high volatility and this creates problems for participants involved in sales of these commodities. The existence of periodicity in the time-series of Onion retail prices in India was pointed out in [2], this will be useful in the analysis below. The cause of periodicity can be related to various factors like rainfall, seasons, temperature etc. [3] studied the causes of volatility in Onion prices by measuring variations in wholesale prices and finding seasonality in market arrivals. Despite no significant shocks, strong uncertainty is present in the Onion prices which shows strong influence of external causes. Seasonality in

All authors contributed equally

[1] Tata Institute of Fundamental Research, Centre For Applicable Mathematics, Bangalore. [2] Mahindra University, Hyderabad E-mail: sakshambassi@gmail.com, gomekaratharva@gmail.com, vasu@tifrbng.res.in

arrivals and prices for different markets of Onion in India is discussed in [4]. Retail markups showed how retailers sold Onions at higher rate in 2010 benefiting from unseasonal rains and low production. Onion prices for multiple cities including, analyzed production, import and exports of Onions are well presented in [5]. It further tried to explain the causality for spikes in prices by speculating political impacts on these prices.

In order to observe the possibility of periodicity in the time-series data of retail prices, the daily retail prices (accessed from [6]) were analyzed for multiple cities of India like Mumbai, Delhi, Kolkata, and Bengaluru. The structure and inferences made from the retail prices of Onions in Mumbai are similar and so we restrict ourselves to the city of Mumbai. The plot of retail prices of Onions in Mumbai with respect to the time is shown in Figure 1. The time period is found to be between April of the present year to March of the next year. It can also be deduced that, generally, the prices peak at around October of each year and prices drop near January next year as confirmed in [2]. Due to uncertain rains in 2010, the supply of Onions decreased, and the prices rose exceedingly. This caused a huge unusual behavior in the time series. A similar case happened in 1998, which caused a similar effect on retail prices. This was consistent in all the parts of India.

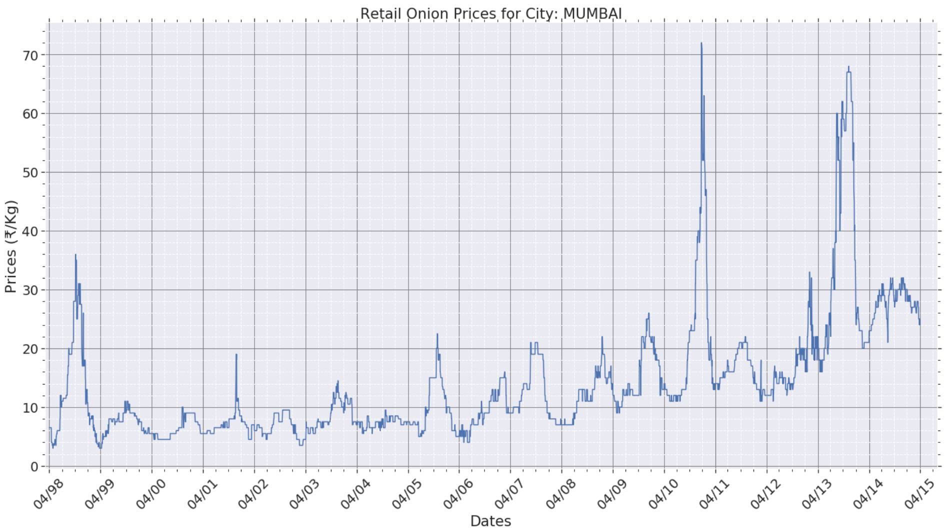

Fig. 1 Daily retail Onion prices of Mumbai from April 1998 to March 2015

2 Generalized Hurst exponent

The method of studying scaling attributes of data using qth-order moments can be done by calculating Generalized Hurst Exponent (GHE). This is generally used to measure the long-term effect in time series. The work in [7] demonstrates how GHE can be used to interpret various kinds of time series and provides an explanation of the formula. Consider a time-series $X : S \to R$ where $S := \{v, 2v, 3v, ..., T\}$. T is a multiplier of $v > 0$ and is v is defined as time resolution. For each $\tau \in (0, \tau_{max}] \cap S$ and $q \in N$, the scaling behaviour of $K_q(\tau)$ is defined as

$$K_q(\tau) \sim (\tau/\nu)^{qH(q)}$$

where

$$K_q(\tau) = \frac{\langle |X(t+\tau) - X(t)|^q \rangle}{\langle |X(t)|^q \rangle}$$

and the operator $\langle \cdot \rangle$ denotes averaging over all $t \in S$. The equation denotes the q^{th} order moment of the distribution of the increments of the time series $X(t)$.

The bar graph in Figure 2 shows the 1st order GHE values of Mumbai for two year periods. All the values look similar, thus it is difficult to conclude anything from this survey. We can't comment on the stability of the market, as it looks fluctuating in a 2-year span but the values are similar. All that we can construe is that there are a slight uneven decrease post-2006 in GHE values.

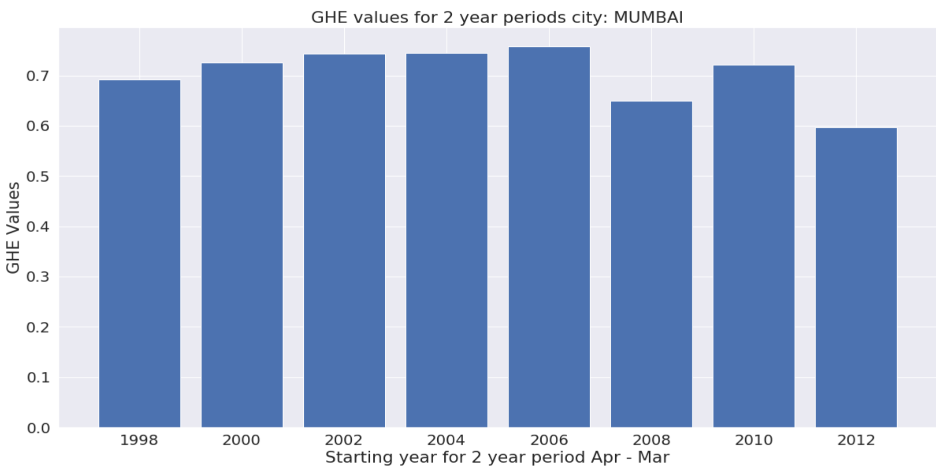

Fig. 2 Bar graph of GHE values of Mumbai for two year time periods

3 Change point detection

Change Point Detection helps us in finding the regimes present in the time series. Changes in the structural patterns of time series might suggest that external factors are responsible for it. There are different approaches for detecting change points [8]. First, the approach of binary change point detection works sequentially. A change point is found in the whole time series and two parts of the time series are formed. Then, change points in these two separate time series are found. This method is repeated for further steps. Second, bottom-up segmentation approach works on dynamic programming where initially many change points are considered. Following this, deletion of less important change points is performed. Third, properties of the time-series are calculated within each window and compared with the following properties of the window. This process is known as window-based change point detection and implemented in `ruptures.detection.Window` [8]. We used this process to detect change points in the retail prices Onion time-series.

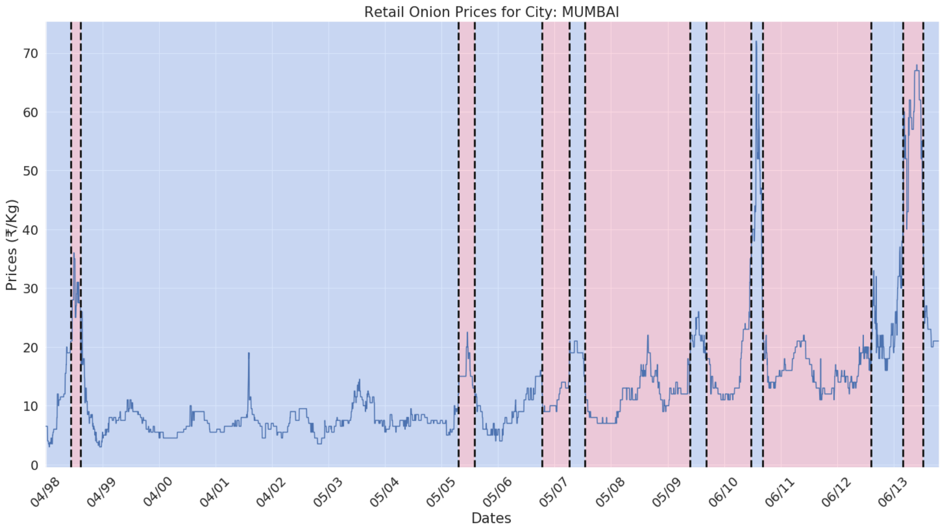

Fig. 3 Results of change point algorithm on retail prices of Onions in Mumbai

The results of the window-based change point detection are shown in Figure 3. The pink and blue portions are seen in either sides of the change points. It can also be observed through the results that around October 2010, a high rise in the retail price of Onion occurred. A very specific period of time series is created after performing Change point detection in October 2010 which is unusual. This could be possible due to various extrinsic factors like limited production, storing goods, or conflicts. For the period of the years of 1999 to 2005, there is no change point because the pattern mostly remains same (that generally varies based on the seasons) and we can infer that the external factors were absent.

4 Statistics

Statistics helps us in analyzing and finding insights in a given time series. We calculated statistics of the Onion retail prices year wise. We explored three basic statistics to find conclusions. The statistics we used are Skewness, Mean and Kurtosis.

4.1 Skewness

Skewness is the degree of asymmetry of a distribution of a random variable around its mean. In other words, it is a calculative measure of which side of the mean the heavy tail is present. Skewness is negative if the distribution is left-skewed and it is positive if the distribution is right-skewed. In Figure 4, all but two year retail prices are positively skewed. A majority of positive skewness deduces that there is a general trend in skewness in peak in first half of the considered year period (April to March of next year). A key insight that always highlights is the disturbance in the year of 2010. We explore the case of the year of 2001 in Subsection 4.3.

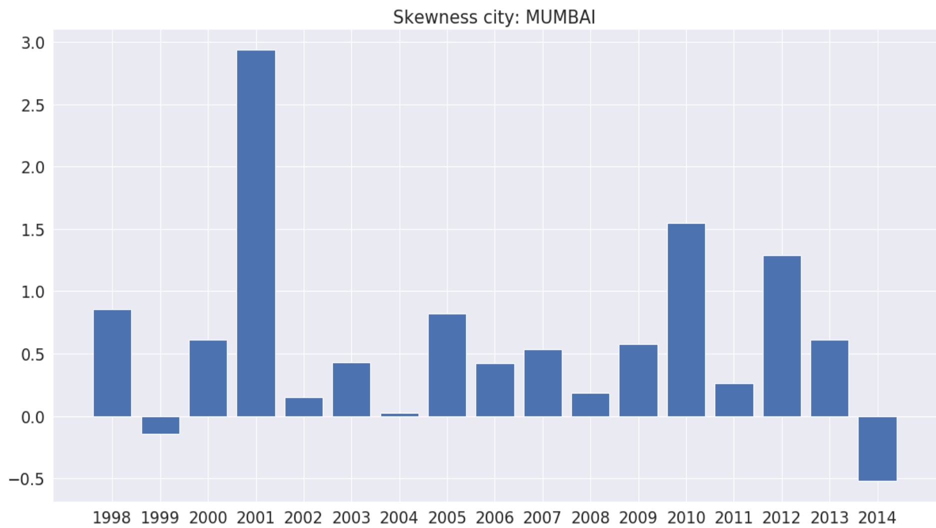

Fig. 4 Skewness values year-wise of Mumbai from 1998 to 2014

4.2 Mean

The average or mean of the Onion retail prices of each year is shown in Figure 5. The interpretations that can be made from this bar plot of means are consistent with what we interpreted in Section 1 from the retail prices time series.

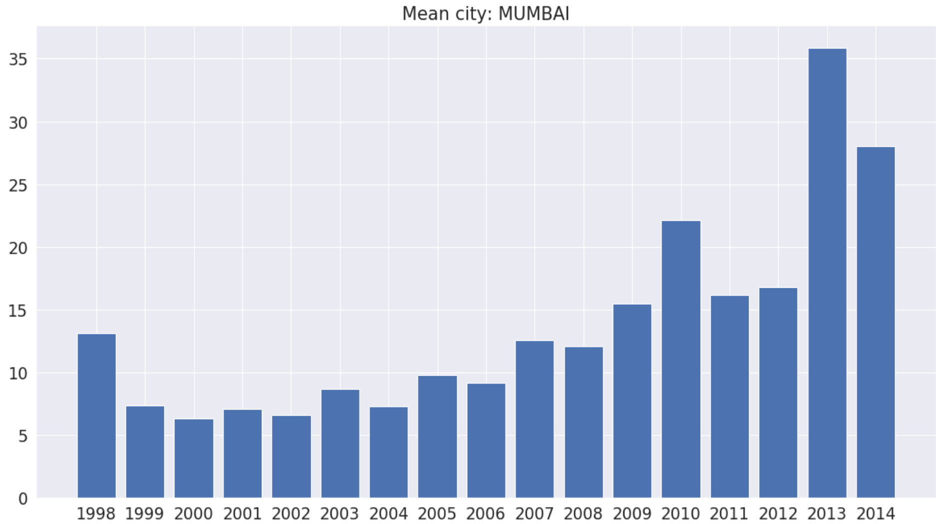

Fig. 5 Mean values year-wise of Mumbai from 1998 to 2014

4.3 Kurtosis

Kurtosis is the measure of the tail in the given time-series. We use Fisher's definition where 3.0 is subtracted from the result to give 0.0 for a normal distribution. In Figure 6, we can construe that the distribution of all but one year has similar values. The year 2001 experienced a long-tailed distribution in the time series. We can support this by referring to Figure 1 where we see a long spike in prices compared to the price values around the tail and in Figure 4, 2001 had the maximum value in skewness compared to other years. The quantitative measure of the tails of each year tends to remain slightly consistent.

5 Cosinor Function Fit

The general idea of Cosinor model is to fit a cosine function to identify seasonal patterns in non-stationary data. For more details we refer to [9]. Cosinor is a parametric model function that identifies seasonal patterns based on sinusoidal function using amplitude and phase concerning time. Amplitude provides the overall size of the pattern and phase delivers the peaks of the time-series patterns.

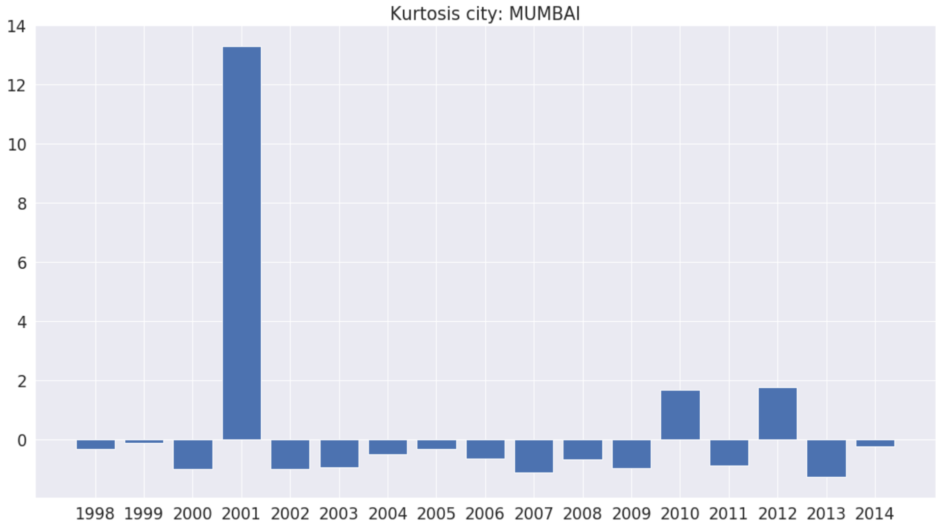

Fig. 6 Kurtosis values year-wise of Mumbai from 1998 to 2014

The formula to denote the Cosinor function is given by

$$s_t = A_t \cos\left(\frac{2\pi t}{c} - P_t\right)$$

where A_t is the amplitude of the sinusoidal form, P_t is the phase of signal, c is the span of the seasonal cycles in a pattern, t is the time of the observation and n is the total number of observations.

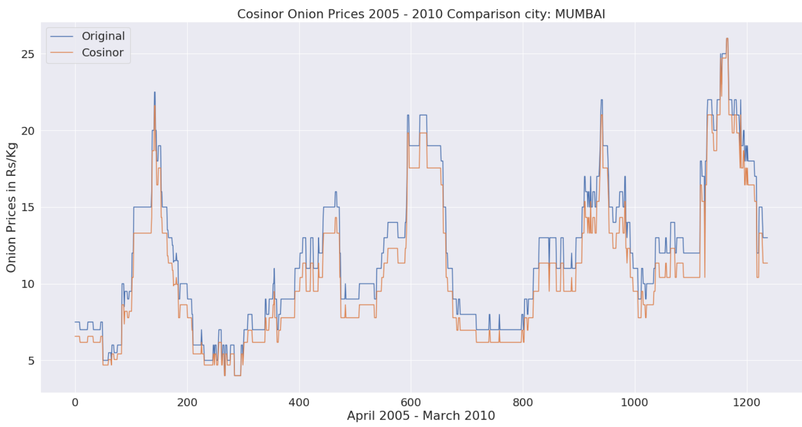

Fig. 7 Cosinor function fit to 2005 - 2010 time series data

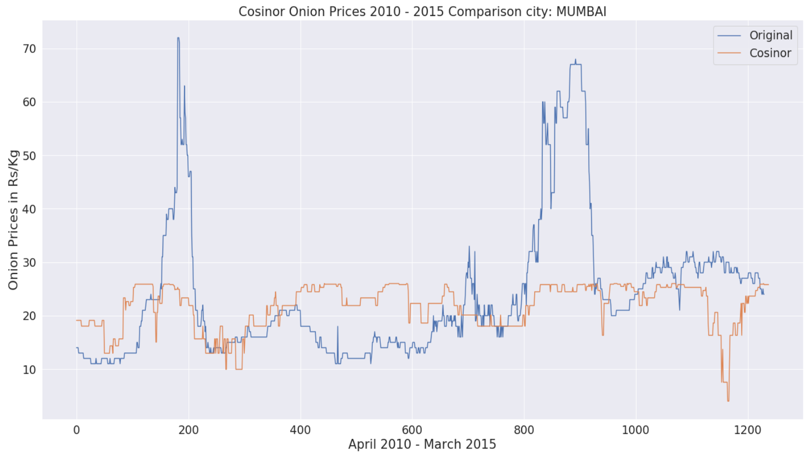

Fig. 8 Cosinor function fit to 2010 - 2015 time series data

We plot the original time series along with simulated time-series generated using the time-based coefficients of the Cosinor model. We used the R package Cosinor ([10]) to fit the cosinor model on the time series.To plot figures 7 and 8, we fitted the cosinor model to the time-series for the period of 2005-2010 and later to 2010-2015 for the city of Mumbai. For each fit, the model throws out values of A_t, P_t and c. We used this value to generate s_t for $t \in [0, n]$ where n is the total number of observations. These new values of s_t become the simulated time-series. This model helps us to analyze this fitting technique for the two periods. Figure 7 shows the plot for time series data between 2005 and 2010 of wholesale Onion prices alongwith the simulated time-series. Figure 8 shows a similar plot for the time series data from 2010 to 2015. The simulated time series for the period, 2005 - 2010 mapped closely to the original time series whereas, for the latter period, it failed to identify seasonal patterns. From these results, we construe that multiple factors are affecting the retail prices of Onions post-2010. Thus, it becomes difficult to find a recurring seasonal pattern in this period and what caused this anomaly is not clear.

6 Conclusions

We analyzed structural patterns and seasonality in time series of Onion prices using various techniques. We saw that each technique yielded unique insights. Generalized Hurst Exponent provided moments of time series to interpret the long-term properties of the time-series, but the values for different periods weren't distinct enough to draw an inference. Change Point Detection helped identify how often regimes changes in a time series by finding the change points. This also showed how fluctuations in the patterns of prices have been increased post-2005 and how frequent it was in the later years. The detection noticed regime changes near peaks specifying an action at the respective time. Basic three statistics were calculated to focus on the peak structure of the distribution of time series in individual years and we deduced insights about the years 2001 and 2010. The distinctiveness in the two years were highlighted by Skewness and Kurtosis. Last, we fitted Cosinor function on the time series which demonstrated contrasting results for the two time periods. From the techniques of Cosinor fit, peak statistics, and change points, we conclude that post-2010, external agents are influencing the retail prices fluctuations.

Acknowledgements

The authors would like to thank Prof. Amit Apte for useful discussions.

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. R. Myers, Time series econometrics and commodity price analysis: A review, Review of Marketing and Agricultural Economics, 62 (1994)
2. A. Desai, The Price of Onions, Ananya Publishing House, Pune (2012)
3. P. Birthal, A. Negi, and P. Joshi, Understanding causes of volatility in onion pricesin india, Journal of Agribusiness in 87Developing and Emerging Economies, 9 (2019)
4. P. Sharma, K. Gummagolmath, and R. Sharma, Prices of onion: An analysis, Economic and political weekly, 46, 22–25 (2011)
5. Gro-Intelligence, When Onion Prices Bring Indians to Tears, https://gro-intelligence.com/insights/onion-prices-in-india, [Online] (2016)
6. M. Chandra, Daily retail prices of Onion upto April - 2015, https://data.gov.in/catalog/dailyweekly-retail-prices-onion, [Online] (2015)
7. T. D. Matteo, Multi-scaling in finance, Quantitative Finance, 7, 21–36 (2007)
8. C. Truong, L. Oudre, and N. Vayatis, ruptures: change point detection in python, ArXiv e-prints, 1801.00826 [stat.CO] (2018)
9. A. Barnett, P. Baker, and A. Dobson, Analysing seasonal data, R Journal, 4 (2012)
10. M. Sachs. Cosinor. https://github.com/sachsmc/cosinor, [Online] (2014)