DOI: https://doi.org/10.61841/2mj8h704 Publication URL: https://nnpub.org/index.php/MS/article/view/1495

DETECTION OF ABERRANT OBSERVATIONS AND MODELLING OF CONSUMER PRICE INDEX AND INFLATION RATE IN NIGERIA FROM 1960 - 2014: THE WAVELET APPROACH

Efuwape, B.T [1]*, Aideyan, D.O[2], Efuwape, T.O[3], Solanke, O.O[4]

[1,] Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Nigeria.Email: efuwapebt@yahoo.com, solankeolakunle@gmail.com*
[2]Department of Mathematical Sciences, Kogi State University, Ayingba, Nigeria. Email: daideyan2006@yahoo.com
[3]Academic Planning Unit, Bells University of Technology, Ota, Nigeria, Email: toefuwape@bellsuniversity.edu.ng

***Corresponding author:-**
efuwapebt@yahoo.com

Abstract

Aberrant observations (AO) are data usually inconsistent with the rest of the series and have the tendency to render statistical inference invalid. The spectral method is widely used for detecting AOs but this is restricted to when series is stationary and periodic. Wavelet shrinkage is a mutually exclusive choice to the spectral method which shrinks the size of the series into multi-resolutions without losing the properties of the series. In this research paper, we present aberrant observation detection and modelling approach based on wavelet analysis in the detection of aberrant observations and modelling using Gaussian and Non-Gaussian distributions in the frequency domain with promising results, on inflation rate in Nigeria between 1960 to 2014. In the method of detecting AO, the inflation rate was highest in 1995 followed by 1993, 1994, 1988, 1989 and 1992. These inflation rates were all observed during the military era. Normal distribution has highest values in all the resolutions from the Loglikelihood calculated, hence, it is the best of the three for detecting AO from the data. Akaike Information Criteria (AIC) confirms that at higher resolutions, Cauchy distribution has lowest AIC which confirms it as the best method of modelling data at such resolutions.

Keywords: *Wavelets, Aberrant observation, Resolution, Modelling, Distributions*

INTRODUCTION

Consumer Price Index (CPI) is average change overtime in the prices of consumer goods and services that people buy for day-to-day living. The CPI uses information from survey of consumption pattern of households to produce a timely and precise mean price variation for the consumption sector of any economy like the Nigeria effective management of the resources. The most important task in the production of CPI is to determine inflation of the market basket of goods and services whose current price is usually compared with its base year price to measure changes in price. Annual CPI has pattern in the forms of up and down movement and circular movements that have been recognized over the years. These movements in CPI sometimes continue for a very long time and also tend to move back and forth as the case may be. As a result, world economies witness periods of resonant and reduced economic activity. In order to understand what causes this, local up and down movement and circular movements can be modelled in a variety of ways in accordance with Money. Notwithstanding, this work was on Nigeria All Items Consumer Price Index (NAICPI) data, the comparisons are also directly related to other state of affairs where data contain a period, or other circular, pattern, for example in geology (Nason 2002), biology (Abraham 1998). The approach to modeling and forecasting time series such as the NAICPI data has been developed over several decades (Dainotti et al. 2006, Nason and Sapatinas 2005). The modelling and forecasting methods for both stationary and non-stationary time series have been applied to many different fields and many successful results have been obtained in different areas. The method of Wavelet Analysis approach in the frequency domain, which assumes that the time series data is best regarded as a sum or linear superposition of periodic sine and cosine waves of different periods or frequencies. (Polikar 2001) and (Graps 1995) will be used to detect aberrant observations and in the model of the data with some distributions in this paper.

Various factors affect the equilibrium of an economy; these factors can either be minute (microeconomic factors) or big (macroeconomic factors). Every player (e.g. the common man, policy makers, government and its agencies etc.) has an important role to play. It is noteworthy that the direct and undulatory effect of the macroeconomic factors or variables is very crucial in the stability of an economy as they deal with summation and not just unit unlike the microeconomic. Among the macroeconomic variables that ensure economic equilibrium, (e.g. exchange rate, interest rate, GDP etc.) inflation is one that constitute a very big imminent danger to economies all around the world, as a matter of fact, it is no respecter of any economy-developed, developing or under developed. It is seen to affect developing economies like Brazil, Venezuela, India, Pakistan, Germany, Israel, Bolivia, Zimbabwe and so on even though they employ strict limit to curb it and its effect, it still goes on to affect them seriously. This is a more reason why constant forecasting has to be done so that policy makers can know what end they should work towards when it comes to inflation as it has proven to avoid limit given its dynamic nature.

Inflation refers to a broad rise in prices across many goods and services in an economy over a steady period of time. Reversed relationship, inflation can also be thought of as the erosion in value of an economy's currency. It is also the rate at which the general level of prices for goods and services is increasing and subsequently, purchasing power is decreasing. Inflation is define as a persistent rise in the level of consumer prices or a continuous decrease in the purchasing power of money caused by a rise in available currency and credit beyond the proportion of available goods and services of a particular country.

Consumer price index

The consumer price index (CPI) is a limit of the total cost of the goods and services bought by a normal consumer. The goal of the consumer price index is to limit changes in the cost of living. This it does by taking into consideration when calculating, various goods a normal consumer will buy and their prices and calculation is done with respect to a certain year called the base year.

Methodology

Wavelet analysis

Wavelet analysis as a mathematical Statistics tool that can be used to collect data from any kind of information and are generally needed to resolve data fully at different multi-resolution (scale) and location (time) whether the data is stationary or not. It is also used to resolve localized difference and allows us to partition (decompose) the data in a time series into pieces that are associated with different multi-resolution (scale) and location (time).

Wavelets are fairly new family of basic functions that are used to depict and proximate other functions. Wavelet coefficients are capable of revealing aspects of the information that other techniques might miss aspects such as changes in variance, level changes, and discontinuities in functions. Thus, due to the essentialities of detecting AOs in signals (or a sequence of data), wavelet analysis is well suited and comes in handy for AO detection. It is also a waveform of limited duration that has a mean value of zero. Wavelet analysis allows data resolve with different resolution match to its scale. It is used to analyse aspects like trends, break points, discontinuity at higher derivatives and self–symmetry, compression or denoising of the data without appreciable degradation.

It is used to refer to a set of orthonormal basis functions generated by dilation and translation of a compactly supported scaling function (or father wavelet), Ø, and a mother wavelet, ψ, associated with an r-regular multi-resolution analysis of $L^2(\mathbb{R})$ (the space of square integrable function). A variety of different wavelet families now exist that combine compact support with various degrees of smoothness and numbers of vanishing moments, and these are now the most intensively used wavelet families in practical applications in statistics. Hence, many types of functions encountered in practice can be sparsely (i.e. parsimoniously) and uniquely represented in terms of a wavelet series. Wavelet bases are therefore not only useful by virtue of their special structure, but they may also be (and have been!) applied in a wide variety of contexts.

Multi-resolution analysis (MRA)

Multi-resolution analysis, as implied by its name, analyses the data at different rate of occurrence with different resolutions. Multi-resolution analysis is designed to give good time resolution and poor frequency resolution at high rate of occurrences and good frequency resolution and poor time resolution at low frequencies. This approach makes sense especially when the data at hand has high frequency components for short durations and low rate of occurrence components for long durations.

"Wavelet analysis investigates the data at different frequencies with different resolutions. That is to say, it uses wider window for low frequency and uses narrower window for high frequency analysis. This feature especially works well for data whose high frequency components have short durations and low frequency components have long durations" (Eckley and Nason 2005).

Thresholding

Let Z_i be independently, identically distributed (iid) standard normal random variables. Define

A_n=$\{max_{i=1,n}|Z_i|\leq \sqrt{2logn}\}$. (1)

Then,

π_n=P(A_n)→0, n→ ∞

In addition, if $B_n(t)$= $\{max_{i=1,n}|Z_i|> t+\sqrt{2logn}\}$ (2)

Then P($B_n(t)$)<$e^{-\frac{t^2}{2}}$. That motivates the following threshold:

τ^u= $\hat{\sigma}\sqrt{2logn}$, (3)

Which is call *universal*. This threshold is one of the first intended and provides an easy, fast, and automatic thresholding. The rationale is to remove all wavelet coefficients that are smaller than the expected maximum of an assumed independently and identically distributed normal noise sequence of given size. There are several possibilities for the estimator, $\hat{\sigma}$. Almost all methods involve the wavelet coefficients of the finest scale. The signal-to-noise ratio is smallest at high resolutions in wavelet decomposition for almost all reasonably behaved signals.

Some standard estimators are:

$\hat{\sigma}^2$= $\frac{1}{\frac{N}{2}-1}\sum_{i=1}^{N/2}(d_{n-1,i}-\bar{d})^2$, or a more robust (5)

$\hat{\sigma}^2$= $\frac{1}{0.6745}$MAD ($\{d_{n-1,i}, i=1, N/2\}$), where n-1 is the highest level. (6)

Proposed wavelet threshold method

τ = $\sigma(2\log(n))^{1/2}$ is the threshold originally suggested by some authors whereas the proposed threshold is

Z=$\frac{X_j-\bar{X}_j}{\sigma_j\sqrt{2logn_j}}$ which can be obtained as follows (7)

Given the data series $\{X_t\}$,

Calculate the mean of X_t.

Calculate the standard deviation $\sigma$of X_t.

Calculate Z as above and j =1, 2, 3…J representing the level of wavelet decomposition.

Perform step 1 to 3 on the original series and on different level of wavelet coefficient J of interest.

It should be noted that Z is assumed to follow a normal distribution and can have possible values in the range [-1, 1]. Any observation outside this range is considered as an AO. Observations lesser than -1 are said to be AOs with lower values and observations greater than 1 are said to be AOs with higher values when compared with the threshold value. This is also true for the compressed data by wavelet analysis.

The key features of wavelet methods

- Sparsity of representation for a wide range of information as resolution decreases including those with discontinuities guaranties the presence of needed statistics.
- The ability to analyse information at a certain number of resolutions and also to work with data at such resolutions.
- Its ability to detect aberrant observations and represent neighbourhood features and also to create localized characteristics on synthesis (the process of combining differences into a new whole)
- Efficiency in terms of compilation speed and storage.

A text statistic was developed to detect these aberrant observations in the non-parametric approach.

Results and Discussion

The AOS detected with their years of occurrence from the original data are as follows:

Table 1: Detected aberrant observations

YEAR	AO
1988	1.456783860
1989	1.311709734
1992	1.087463235
1993	1.561889502
1994	1.555887799
1995	2.164837999

Table 2: Loglikelihood figures

Resolution	Normal Distribution	Laplace Distribution	Cauchy Distribution
64	1008.593	952.237	492.912
32	539.213	205.765	207.434
16	253.416	105.685	105.800
8	116.032	51.362	51.648

Table 3: Akaike Information Criteria

Resolution	Normal Distribution	Laplace Distribution	Cauchy Distribution
64	1012.593	956.237	496.912
32	543.213	209.765	211.434
16	257.416	109.685	109.800
8	120.032	55.362	55.648

From table 1 it was observed using the developed threshold that the inflation rates were observed between 1988 to 1995 except in 1990 and 1991. The rate of inflation were observed in 1995 with (2.1648), 1993 (1.5619), 1992 (1.5559), 1988 (1.4567), 1989 (1.3117) and 1992 (1.0875).

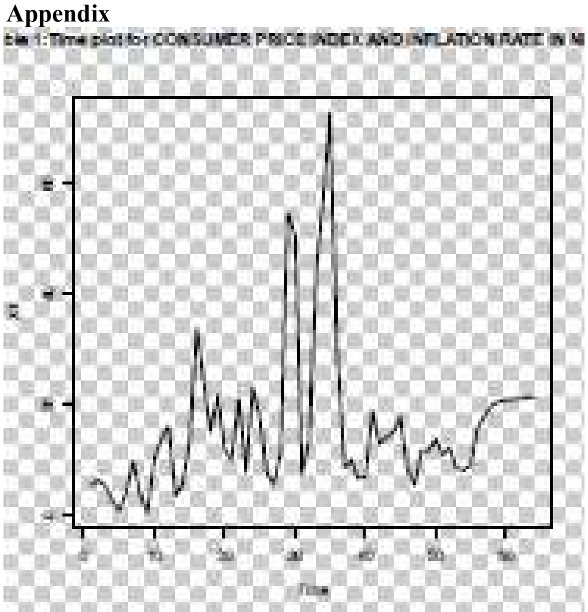
From table 2 and table 3, with the use of some distributions, it was observed that at different resolutions, Normal distribution has higher values followed by Laplace distribution at higher resolutions while Cauchy distribution has lowest values at higher resolutions but higher values at lower resolutions. These findings are also confirmed in their graphs in the appendix (Figures 3 and 4).

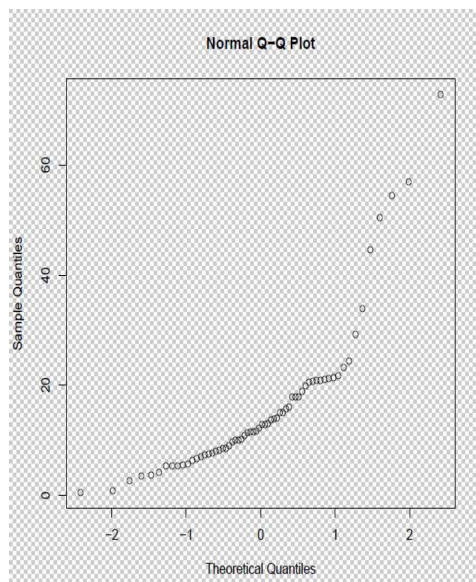
Conclusion

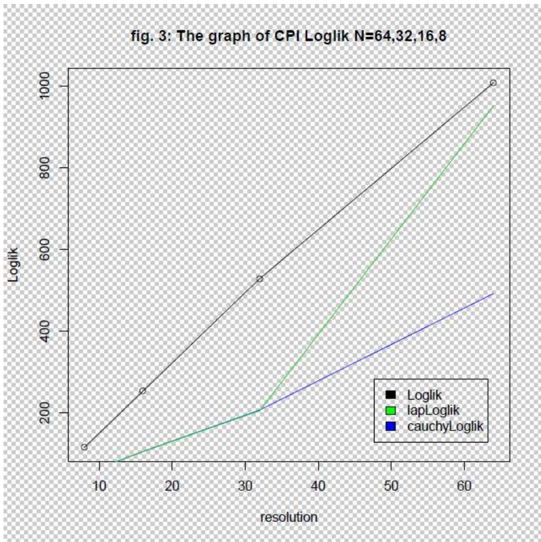
In the method of detecting AO, the inflation rate was highest in 1995 followed by 1993, 1994, 1988, 1989 and 1992. These inflation rates were all observed during the military era

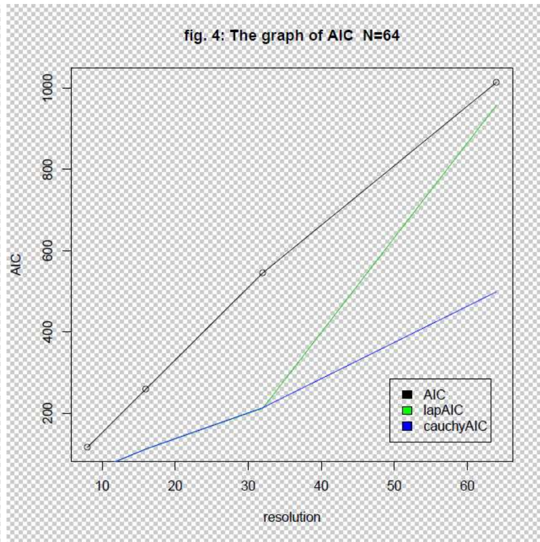
From tables 2 and 3, since normal distribution has highest values in all the resolutions from the Loglikelihood calculated, we conclude that of the three distributions, normal distribution is the best of the three for detecting AO from the data. The AIC confirms that at higher resolutions, Cauchy distribution has lowest AIC which confirms it as the best method of modelling the data; while at lower resolutions, the Laplace distributions is best in modelling the data.

Appendix









Conflict of interest

No conflict of interest has been declared by the authors.

REFERENCES

[1].Nason GP 2002 Choice of wavelet smoothness, primary resolution and threshold in wavelet shrinkage. Statistics and Computing, 12(3), 219-227

[2].Abraham M 1998 Histogram Smoothing Via the Wavelet Transform. Journal of Computational and Graphic Statistics, 7(4).

[3].Dainotti A, Pescape A and Viorgio G 2006 Wavelet-based detection of DOS attack. IEEE Globecom, 1-6.

[4].Nason G and Sapatinas T 2005 Wavelet packet transfer function modeling of nonstationary time series. Statistics and Computing, 12(1), 45-56.

[5].Polikar R 2001 The Wavelet Tutorial: Fundamental concepts and an overview of the wavelet theory. Rowan University, Glassboro, NJ: Dept of Electrical and Computer Engineering,

[6].Graps A 1995 An Introduction to Wavelets. IEEE Computational Science and Engineering, 2(2), 50-61.

[7].Eckley I and Nason GP 2005 Efficient computation of the discrete autocorrelation wavelet inner product matrix. Statistics and Computing, 15, 83-92.