Journal of Statistics and Computer Science
Vol. 2, Issue 2, 2023, pp. 103-110
https://doi.org/10.47509/JSCS.2023.v02i02.02

De-Noising Time Scale Decomposition Graph Metrics of S&P BSE Sensex: Analyze through MODWT with Daubechies Filter

Rahul Kumar Si

Statistical Assistant, Tobacco Board, Ministry of Commerce & Industry, Government of India, Guntur, Andhra Pradesh – 522004, E-mail:rahulstat.mbj@gmail.com.

Abstract

STOCK MARKETs are important aspect of fiscal statistics, which spores the delight over the years to originate better apocalyptic models. Due to nonstationarity and noise in stock market data, financial model produces unreliable and spurious results and leads to poor understanding and forecasting. To erase these problems, log-transformation was used to decrease the variability and then, Maximal Overlap Discrete Wavelet Transform (MODWT) was used on the log transform data for construction of time scale decomposition graphs and de-noises the stock market data with Daubechies filter. From the above study, we identify that, on the estimated values of graph metrics, the length of the wavelet filter chosen has a greater effect when compared with log-transformation and row data.

KEY WORDS: De-Noising, Wavelet, Daubechies and Filter
MSC 2020 Subject Classification: 62J05, 62P20, 65T60

1 Introduction

From the past few decades, stock price forecasting has drawn substantial surveillance among researchers as well as policy makers because of the crucial role in economy. From macro-economic prospective, it provides basic information for government decision making, whereas, from micro-economic prospective, it gives idea about investor's decision. In simple words, the principal target behind the stock price prediction is to realize best outcome with minimum risk.

Basically, stock market data are non-stationary in nature and heteroskedastic processes with random noise and exhibit changing frequencies over time (Luo et al., 2016). Wavelet transform has become very popular in the field of stock market analysis due to its pliability to handling non-stationary data (Kumar et al., 2011). It effectively eliminates the unnecessary signal in wavelet domain and recovers the coveted signal from an inverse wavelet transform with small waste of main details (Bemporad et al., 2002).

A useful application of wavelet transforms in investment theory was advocated by Chakrabarty et al., (2015). Among two types of wavelet transforms, various researchers did a sensational work to rate the performance of MODWT over DWT (Bolzan et al., 2009; Stolojescu-Crisan et al., 2010; Dghais and Ismail, 2013; Follis and Lai, 2020).

In order to study the strong to moderate cointegration, Pinho and Madaleno (2009), used the MODWT, cross wavelet technique and regression to study the world's major eleven stock markets. Similar work was also done by Ranta (2013), to examine the correlation composition of the major world markets by using the wavelet coherency. Similarly, Rua and Nunes (2009), used measures of dispersion to get figures for the major developed economies and then, wavelet squared coherency were applied to the time and frequency varying co-movement.

The accessible belles-lettres on implementation of MODWT on financial statistics is divided into four main categories: Transform, Variance Decomposition, Outlier Detection and De-Noising. This paper is basically concentrated on the study of the above four

Article History
Received : 26 June 2023; Revised : 20 October 2023, Accepted : 28 November 2023; Published : 30 December 2023

To cite this paper

Rahul Kumar Si (2023). De-Noising Time Scale Decomposition Graph Metrics of S&P BSE Sensex: Analyze through MODWT with Daubechies Filter. *Journal of Statistics and Computer Science*. 2(2), 103-110.

categories due to the complexity of heterogeneous financial time series, which results in risks and correlation.

The remaining structure of the article proceeds as follows. Section 2 describes the data used in this article and their sources. Section 3 overview the methodology used. In Result and Discussion: Normality test and log transformation, MODWT based multiresolution analysis, Variance Decomposition, De-noising with MODWT transformation and Stationarity test are analysed in Section 4. Finally, section 5 concludes the paper.

2 Data and their Sources

The daily historical S&P BSE Sensex data from 6th April, 2015 to 31st March, 2022 having a total number of 1732 observation used in this research was taken from the official website of Bombay Stock Exchange and analyzed with the help of SPSS and E Views - 12 Software. In order to achieve the reliable results, following methods are adopted and before log transformation results are compared to after log transformation results.

3 Research Design and Methodology

The prime intention of the research is to propose a pre-sample forecast of S&P BSE Sensex, so that a reliable and sensitive forecasting model should be applied with minimum error with following methodologies.

- ❖ Firstly, K-S test was used to check whether the observed S&P BSE Sensex data are normal or not, because significantly skewed data grossly underestimates the forecast values and produce unreliable and spurious results.
- ❖ After studying the skewed behaviour, log transformation was done in S&P BSE Sensex data to escalate its illustrability and afterwards for the statistical analysis.
- ❖ Then, MODWT with Daubechies filter was used on both the log-transformation and row data for construction of time scale decomposition graphs.
- ❖ Variance decomposition was analysed with the help of time scale decomposition graphs and scale-by-scale decomposition of variance produces by spectrum table.
- ❖ Persistent behaviour of both the log-transformation and row data were analysed by Variance Spectrum Distribution and Cumulative Variance for feasibly affirmation of a unit root.
- ❖ Then, a unit root test on the series will confirm the present / absent of stationarity.

4 Results and Discussion

4.1 Normality Test and Log Transformation

During the study period, Skewness of S&P BSE Sensex, before log transformation is moderately skewed. To overcome this problem, log-transformation was used on S&P BSE Sensex data for approximate confirmation to normality and increase the interpretability and subsequently for the statistical analysis (Lütkepohl and Xu, 2010).

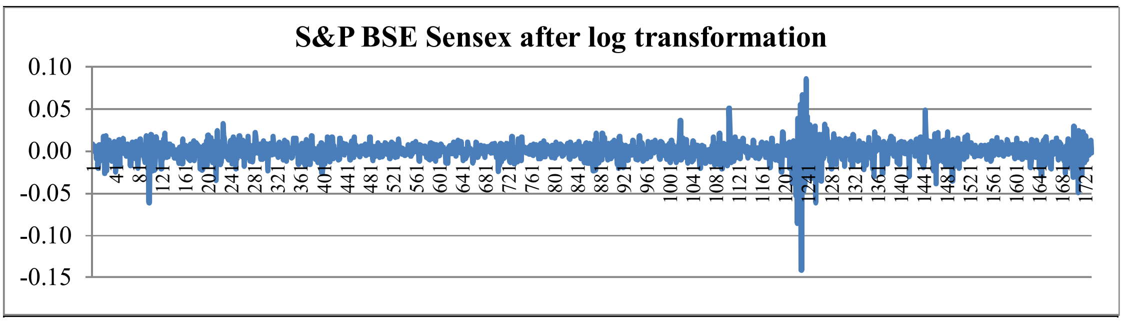


Figure 1: S&P BSE Sensex after log transformation.

For testing the normality, the Kolmogorov - Smirnov (K-S) statistic after log transformation was 0.086 (Asymp. Sig. > 0.05) and leading to the conclusion that daily observed S&P BSE Sensex is normally distributed and feasible to forecast after log transformation (Simard and L'Ecuyer, 2011).

4.2 MODWT Based Multi–Resolution Analysis

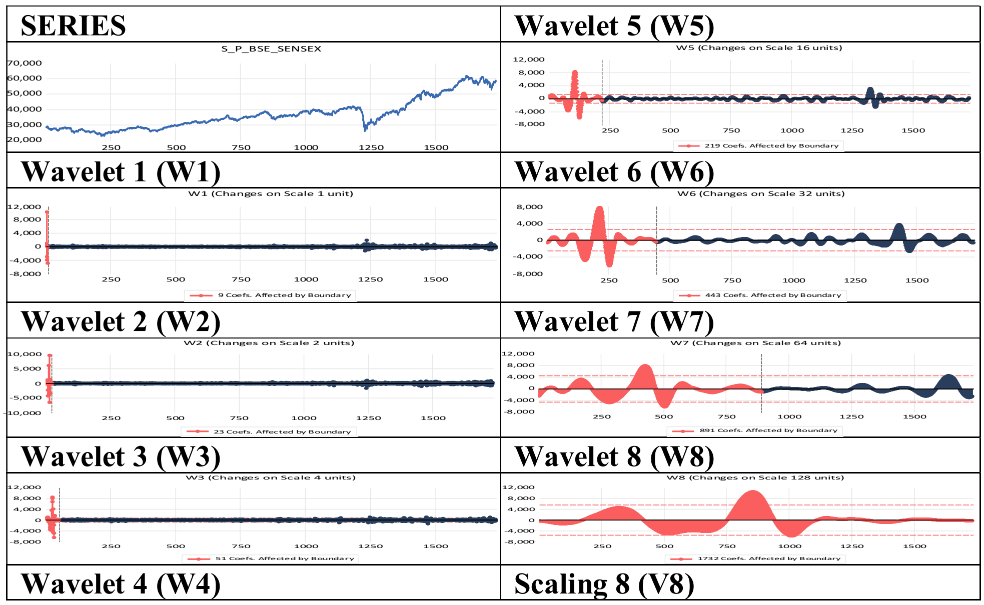
At the starting point, we assumed that, log returns of S&P BSE Sensex were non-stationary in nature. To extract seasonal trend and abrupt components and allow for a better explication, MODWT is used to decompose the data into a non-orthogonal set of components which uses all of the available observations with different frequencies (Polanco Martínez et al., 2018).

For estimating the scaling and wavelet coefficients, MODWT does not suffer from level of sample size to an integer multiple of 2^t (Percival and Walden, 2000), where t is the level of decomposition and do not require any length adjustments**.** It is inspecting that, estimators based on MODWT are asymptomatically more efficient as compared to DWT (Percival and Mofjeld, 1997). Since the MODWT is a non-orthonormal transform and does not require any length adjustments, the S&P BSE Sensex data decomposed into nine time-scale components with Daubechies filter, which is smoother than Haar wavelet filters and provides better un-correlated-ness across scales (Cornis et al., 2006).

In other words, the Daubechies wavelet filter with periodic boundary conditions of length L = 8, decomposes level the S&P BSE Sensex data with maximum decomposes level t = 8 and produces nine wavelet and scaling filter sets of coefficients (Daubechies, 1992).

At the first scale, some wavelet coefficients exceed the threshold bounds in both the series: normal and log transform of S&P BSE Sensex data and the transient features keep on at scale two and three onward, but after that, don't seem to contribute much. Higher frequencies are dominated by lower frequency forces and likely to be stationary after log transformation, whereas, reverse is seen in before log transformation series**.**

Finally, for each scale, the coefficient affected by the boundary are displayed in red, their count in blue and a vertical dashed black line shows the region upto which the boundary conditions persist for each scale. Boundary coefficients are important for longer filters and higher scales.



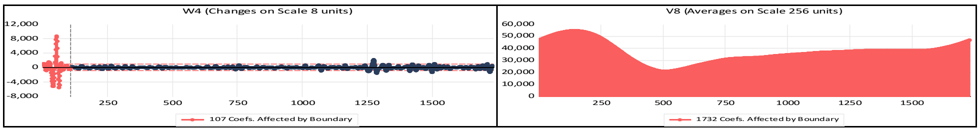


Figure 2: MODWT based time scale decomposition of S&P BSE Sensex before log transformation.

.

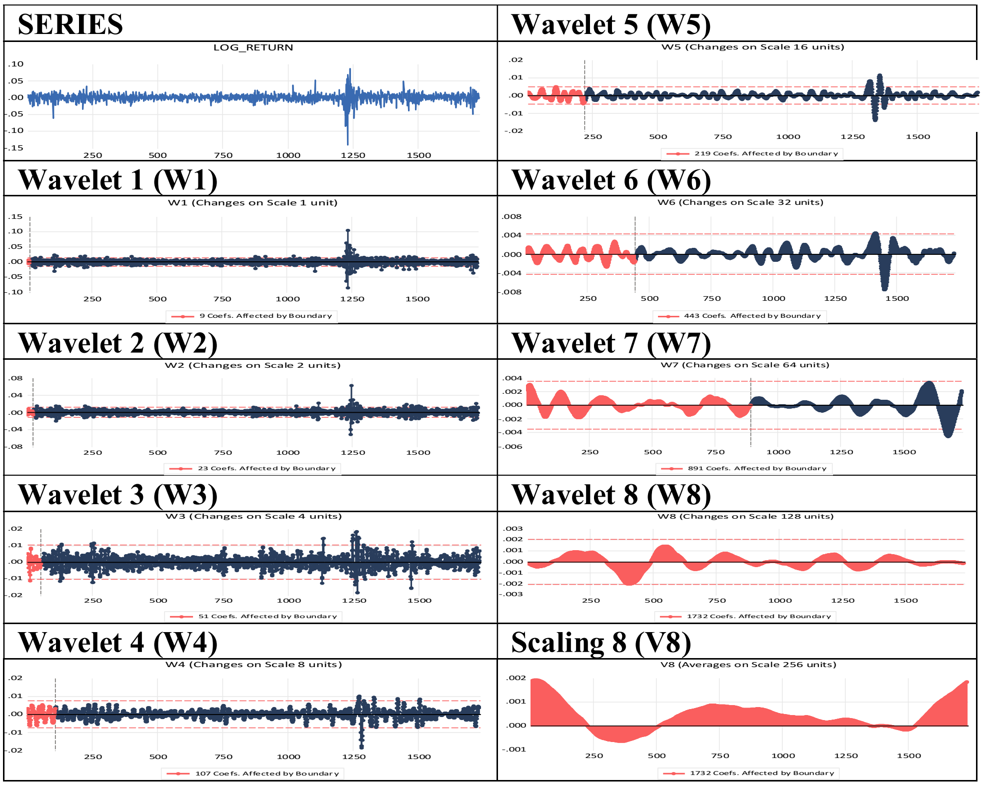


Figure 3: MODWT based time scale decomposition of S&P BSE Sensex after log transformation

4.3 Variance Decomposition

The contribution of variance by wavelet coefficient at each scale is demonstrated in spectrum table. The column titled Variance, Relative Proportion and Cumulative Proportion respectively is the variance contributed to the total at a given scale, the proportion of overall variance contributing to the total at a given scale and its cumulative total.

Table 1: Spectrum Table of S&P BSE Sensex before log transformation.

Scale	Variance	Relative Proportion	Cumulative Proportion	95% Conf. Interval	
				Lower	Upper
W1	29235.23	0.0088	0.0088	26206.85	32263.61
W2	41588.35	0.0125	0.0213	35331.76	47844.95
W3	67893.06	0.0205	0.0418	53116.42	82669.69
W4	134736.70	0.0406	0.0824	96075.59	173397.90
W5	273617.60	0.0824	0.1648	163368.3	383867.00
W6	757450.00	0.2282	0.3930	241820.8	1273079.00
W7	2014661.00	0.6070	1.0000	54596.82	3974726.00

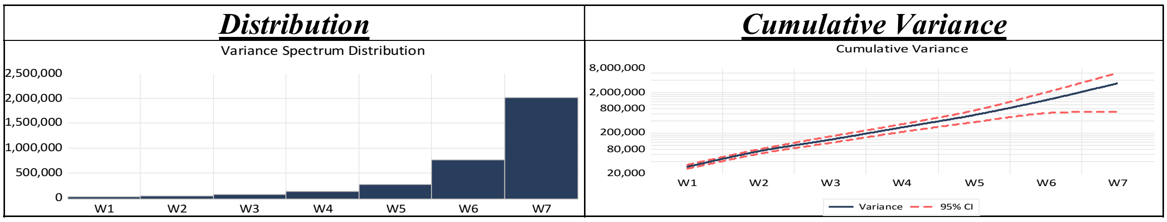


Figure 4: Variance Spectrum Distribution and Cumulative Variance before log transformation.

Majority of the variation before log transformation (Figure 4) in the S&P BSE Sensex series comes from higher scales or lower frequencies and have an indication of scattered behaviour in the original data and possibly evidenced that data are non stationary and not free from unit root, whereas, the variation after log transformation (Figure 5) are stationary and free from unit root.

Table 2: Spectrum Table of S&P BSE Sensex after log transformation.

Scale	Variance	Relative Proportion	Cumulative Proportion	95% Conf. Interval	
				Lower	Upper
W1	6.73e-05	0.5205	0.5205	5.94e-05	5.71e-05
W2	3.16e-05	0.2444	0.7649	2.72e-05	3.60e-05
W3	1.46e-05	0.1131	0.8780	1.14e-05	1.79e-05
W4	7.26e-06	0.0562	0.9342	5.23e-06	9.30e-06
W5	4.37e-06	0.0338	0.9681	2.74e-06	6.01e-06
W6	2.29e-06	0.0177	0.9858	9.50e-07	3.63e-06
W7	1.83e-06	0.0142	1.0000	4.69e-09	3.66e-06

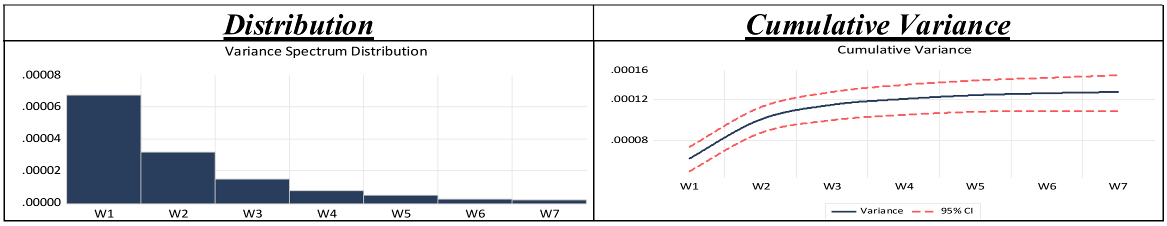


Figure 5: Variance Spectrum Distribution and Cumulative Variance after log transformation.

4.4 Denoising with MODWT Transform

After completing the MODWT-based denoising process, the stable denoised data are obtained with smoother continuous trend. A closer look reveals that, the heterogeneous nature occurred in denoised data of S&P BSE Sensex before log transformation is an indication of abnormal fluctuation caused by random chances.

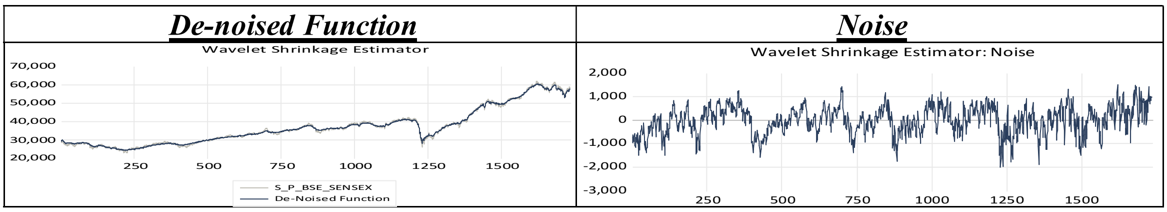


Figure 6: De-Noised Function and Noise of S&P BSE Sensex before log transformation.

From Figure 6, the irregular variations in S&P BSE Sensex may be caused by mishaps often reflect in short-term flow and deviate from normal trend. Such types of abnormal component are filtered and perpetuate the trend with more sensitive to normal points and ameliorate the continuity of the trend (Figure 7).

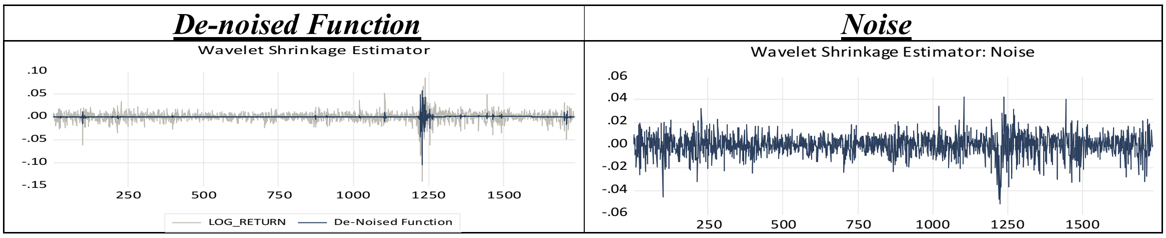


Figure 7: De-Noised Function and Noise of S&P BSE Sensex after log transformation.

4.5 Stationarity Test

In order to ascertain the series stationarity, ADF and PP test are conducted and according to their t-statistic, the stationarity could be identified (Said and Dickey, 1984; Phillips and Perron, 1988).

Table – 3: Summary of ADF and PP test.

Variables	ADF Test					PP Test				
	t-stat.	Prob*	Level of Significance			t-stat.	Prob*	Level of Significance		
			1%	5%	10%			1%	5%	10%
Before log transformation	0.16	0.97	-3.43	-2.86	-2.57	0.27	0.98	-3.43	-2.86	-2.57
After log transformation	-11.35	0.00	-3.43	-2.86	-2.57	-42.93	0.00	-3.43	-2.86	-2.57

The optimal lag for ADF test is selected based on the AIC criteria and Bartlett Kernel method and Newey - West method is used for fixing the truncation lag and for spectral estimation in PP test (Mishra, Arunachalam and Patnaik, 2018). The calculated values < tabulated t-statistic at 1%, 5% and 10% level of significance for both the ADF and PP test and hence the null hypothesis is rejected and confirms the existence of stationarity after log transformation of Sensex.

5 Conclusion

De-noising analysis through MODWT is a powerful mathematical contrivance for stock market forecasting. It is simple to work out and does not depend on certain parameter choices and model preference criteria. It decomposes the stock market data into a non orthogonal set of units with distinct prevalence and estimates the scaling and wavelet coefficients. While the main purpose is to introduce a wavelet methodology in our study for decomposing the S&P BSE Sensex data based on MODWT to examine the graph matrices before log transformation and after log transformation. The variances before log transformation and after log transformation yield a specific difference after examining the variance, relative proportion and cumulative proportion in spectrum table using the LA 8 wavelet. After getting the stationarity in log transformation data, a robust real time stock prediction technique should be used with minimum risk. Basically, here we concentrate on to bespeaking how MODWT model can be used as a fact-finding instrument for stationarity instead of a conventional test.

Acknowledgement

I gratefully extend my sincere thanks to the reviewers for their precious time in reviewing my paper and providing valuable comments, which inspired me to make possible improvements to the current version. I also thank Dr. S. K. Satpathy, Centurion University of Technology and Management, Bhubaneswar, Odisha, for his kind cooperation in completing the manuscript on time. I would like to express my sincere gratitude to the editor of the Journal of Statistics and Computer Science for his valuable guidance and moral support throughout the process.

REFERENCES

Bemporad, A., Borrelli, F. and Morari, M. (2002). Model Predictive Control Based on Linear Programming: The Explicit Solution. *IEEE-Transaction on Automatic Control*, 47(12), 1974-1985.

Bolzan, M. J. A., Guarnieri, F. L. and Vieira, P. C. (2009). Comparisons between two Wavelet Functions in Extracting Coherent Structures from Solar Wind Time Series. *Brazilian Journal of Physics*, 39(1), 12-17.

Chakrabarty, A., De, A., Gunasekaran, A. and Dubey, R. (2015). Investment Horizon Heterogeneity and Wavelet: Overview and Further Research Directions. *Physica A: Statistical Mechanics and its Applications*, 429, 45-61.

Cornish, C. R., Bretherton, C. S. and Percival, D. B. (2006). Maximal Overlap Wavelet Statistical Analysis with Application to Atmospheric Turbulence. *Boundary Layer Meteorology*, 119(2), 339-374.

Daubechies, I. (1992). *Ten Lectures of Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics Analysis, pages 1544-1576.

Dghais, A. A. A., and Ismail, M. T. (2013). A Comparative Study between Discrete Wavelet Transform and Maximal Overlap Discrete Wavelet Transform for Testing Stationarity. *International Journal of Mathematical Science and Engineering*, 7(12), 471-475.

Follis, J. L., and Lai, D. (2020). Variability Analysis of Epileptic EEG using the Maximal Overlap Discrete Wavelet Transform. *Health Information Science and Systems*, 8(26), 1-12.

Kumar, A., Joshi, L. K., Pal, A. K. and Shukla, A. K. (2011). MODWT Based Time Scale Decomposition Analysis of BSE and NSE Indexes Financial Time Series. *International Journal of Mathematical Analysis*, 5(27), 1343-1352.

Luo, S., Yan, F., Lai, D., Wu, W. and Lu, F. (2016). Using ARIMA Model to Fit and Predict Index of Stock Price Based on Wavelet De-Noising. *International Journal of u- and e-Service: Science and Technology*, 9(12), 317-326.

Lütkepohl, H., and Xu, F. (2010). The Role of the Log Transformation in Forecasting Economic Variables. *Empirical Economics*, 42, 619-638.

Mishra, A. K., Arunachalam, V. and Patnaik, D. (2018). Current Issues in the Economy and Finance of India. *IEEF Springer Proceedings in Business and Economics.*

Polanco Martínez, Josué M., Abadie, Luis M. and Fernandez-Macho, J. (2018). A Multi Resolution and Multivariate Analysis of the Dynamic Relationships between Crude Oil and Petroleum Product Prices. *Applied Energy*, 228, 1550-1560.

Percival, D. B., and Mofjeld, H. O. (1997). Analysis of Subtidal Coastal Sea Level Fluctuations using Wavelets *Journal of American Statistical Association*, 92(439), 868-880.

Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge.

Phillips, P.C.B., and Perron, P. (1988). Testing for Unit Roots in Time Series Regression. *Biometrika*, 75, 335-346.

Pinho, C., and Madaleno, M. (2009). Time Frequency Effects on Market Indices: World Co-movements. *Proceedings of the Finance International Meeting*, AFFI-EUROFIDAI, Paris, France, 17–18 Dec. 2009.

Ranta, M. (2013). Contagion among Major World Markets: A Wavelet Approach. *International Journal of Managerial Finance*, 9(2), 133-149.

Rua, A., and Nunes, L. C. (2009). International Co-movement of Stock Market Returns: A Wavelet Analysis. *Journal of Empirical Finance*, 16, 632-639.

Said, S.E., and Dickey, D. (1984). Testing for Unit Roots in Autoregressive Moving Average Models with Unknown Order. *Biometrika*, 71, 599-607.

Simard, R., and L'Ecuyer, P. (2011). Computing the Two-Sided Kolmogorov–Smirnov Distribution. *Journal of Statistical Software*, 39(11), 1-18.

Stolojescu, C., Railean, I., Moga, S. and Isar, A. (2010). Comparison of Wavelet Families with Application to Wi-MAX Traffic Forecasting. *12th International Conference on Optimization of Electrical and Electronic Equipment (IEEE)*, 932-937. OPTIM 2010.