Group for Time Series Components (GFTSC) Identification of Gross Domestic Product (GDP) of United Kingdom (UK)

[1, 2] Ajare Emmanuel Oloruntoba; [1]Adefabi Adekunle
[1]School of Quantitative Sciences, College of Art and Sciences, University Utara Malaysia, Malaysia
[2]Department of Mathematical Sciences, Faculty of Sciences, Federal University Gusau, Gusau, Nigeria
[3]Department of Mathematics and Statistics, Austin Peay State University, Clarksville, Tennessee, USA

Abstract:- The main objective of this study is to use GFTSC (Group for Time Series Components) to identify the components of time series present in the seasonal data of (UK GDP). This data is the GDP yearly data of United Kingdom gross domestic product (UK GDP). The (UK GDP) data spanned for the period of twenty years. The GDP of UK is a secondary data obtained from the DataStream of Universiti Utara Malaysia Library. The weaknesses of BFAST (Break for Additive Seasonal and Trend) were corrected by the extension of BFAST to GFTSC which resulted into creation of a new technique named Group for Time Series Components. BFTSC was created to capture the cyclical and irregular components that was not captured by BFAST technique. BFTSC is designed to present the image of all the 4 time series components. BFAST only identifies trend and seasonal components only. Evaluation using simulation data was conducted to verify the accuracy of GFTSC using monthly simulated of 144, 000 data unit. This data contained 48 months small monthly sample size, 96 monthly medium sample size, 144 months large sample size. Each of the sample size was replicated 100 time each. GFTSC is effective and better than BFAST because it was able to identify approximately 100% of the data with the basic four time series components monthly. BFTSC detects 99.99% of the entire components in the time series monthly data that was tested. Empirical data were employed to BFTSC and subsequently determine the next forecasting technique after which one step forecast is made ahead. The simulated and real data findings suggested that GFTSC can provide a better alternative to BFAST technique, hence GFTSC is recommended.

Keywords:- *Group for Time Series Components, Seasonal Data, Gross, Cyclical , Irregular Components.*

I. INTRODUCTION

This study uses GFTSC (Group for time series components) to identify the components of time series present in the empirical data of quarterly seasonal data which is the GDP yearly data of United Kingdom gross domestic product (UK GDP). GFTSC is considered to be more efficient in identifying all the components of time series statistics better than BFAST. GFTSC is an improved BFAST. GFAST (Group for Additive Seasonal and Trend) is a technique used for identification of trend and seasonal components of time series observations, trend breaking was first suggested by Bai and Perron (2003). Jong, Verbesselt, Schaepman and Bruin (2012) recommended an approach of basic swing identification to spot time series component. This approach was also used by Zewdie, Csaplovics and Inostroza (2017) as the latest time series component recognition approach which is a technique that was first described and utilized by Verbesselt et al. (2010).

The technique BFAST was for recognizing breaking points with the help of seasonal and trend decomposition using loess (STL), it facilitates the detection of trend change in a given information. The elementary standard of the BFAST technique is the splitting of time series into seasonal, trend and also remnants element by the approach for breaks detecting software in R studio core 2012 **(**Cunha**,** 2013).

II. LITERATURE REVIEW

The technique BFAST had much lower RMSE and was more robust against noise, Hence BFAST is recommended as one of the best trend break detection. One of the limitation of CCDC with CV is that its algorithm was made complicated, unlike CCDC, CCDC with CV did not have a straightforward relationship between RMSE number of breaks and noise. CCDC with CV was also found to be less accurate (Zhu, Zhang, Yang, Aljaddani Cohen, Qiu & Zhou, 2020). Another limitation of this technique is also in terms of noise, with increased noise, the technique was less likely to detect correct results and the likelihood of detecting at least one false break remained constant. The unique pattern shown by CCDC with CV suggests that it must also detect more breaks if there is very little noise (Zhu, Zhang, Yang, Aljaddani Cohen, Qiu & Zhou, 2020).

EWMACD was built to focus on subtle changes, such as partial changes within pixels (Brooks, Wynne, Thomas, Blinn, Coulston, 2019). Just like CCDC and BFAST Monitor, EWMACD also detects condition (increasing/decreasing trend) the EWMA chart, to rapidly help in identification of time series component.

Zhu, Zhang, Yang, Aljaddani, Cohen, Qiu and Zhou (2020) developed a new univariate time series components identification method known as COntinuous Monitoring of Land Disturbance (COLD) using Landsat time series data. COLD can detect many time series component such as trend and seasonal. COLD can also detect land disturbance continuously as new pattern is collected and likewise provide historical land disturbance history. Evaluation of the trend detection ability and land disturbance, different kinds of data are utilized. The COLD algorithm was developed and calibrated based on all the lessons learned. The accuracy assessment shows that COLD results were accurate for detecting trend and seasonal as land disturbance with an omission error of 27% and a commission error of 28%. The limitation of COLD was inability to detect time series components accurately with large noise.

Zewdie et al. (2017) argues that the technique of BFAST can predict and analyse a topographical forest movement with the help of normalized difference vegetation index's branded as (NDVI). This was done by detecting and determining factors of arid area changes using (NDIV) data to monitor the variations (Cesta, Cortellessa, Pecora, & Rasconi, 2005 ; Buhalau, 2016 ; DeVries, 2013). Many scholars employ the use of BFAST in identifying trend in topographical data (Porter & Zhang, 2018).

The extension of BFAST is an improved technique that identifies all-time series components. This new technique is known as GFTSC (Group for time series components). Many of the automated techniques of pattern detection are computer oriented. GFTSC is one of the first extension of BFAST in history which also focus more on computer approach strategy rather than theoretical approach strategy GFTSC technique considers every vital component of time series statistics. BFAST is known to be weak in identifying and breaking random variations, also very weak in applicability to other types of empirical data (Flicek & Birney 2009). The technique considers the extension and improvement of the BFAST to GFTSC.

GFTSC is made available into computer R package and can be used by anyone who wishes to be a beneficiary, for better identification and diagrammatic representation of the time series data to bridge the gap of time series components identification (Flicek & Birney 2009). GFTSC followed similar derivative steps like BFAST but in addition of cyclical and irregular components. GFTSC is the technique used in analyzing the generality of time series data by extracting the trend components and seasonal components, cyclical components and irregular components during time series decomposition. Given the general time series additive model as in equation (1.1) of the form:

$$Y_P = T_p + S_p + C_p + \mathrm{I_p} \tag{1.1}$$

For identification of Y_P , S_p , C_p , and I_p (See the paper: Box, Jenkins, Reinsel, & Ljung, 2015 :Maggi, 2018; Cleveland & Tiao, 1976; Caiado, 2009; Bohn, 1995; Cipra, & Romera, 1997).

GFTSC takes all the important components relatively trend, seasonal, cyclical and irregular components to be important. The residual component in BFAST now converted to contained cyclical and irregular component in GFTSC. In BFAST only random component can be observed but in GFTSC the cyclical and irregular components is included (Zdravevski, Lameski, Mingov, Kulakov & Gjorgjevikj, 2015).

III. MATERIAL AND METHODS

BFAST is the technique used in analyzing the generality of time series data by extracting the trend and seasonal pattern during time series decomposition. Given the general time series additive model of the form of equation 1.1 (Maggi, 2018 ; Zhao, Li, Mu, Wen, Rayburg, & Tian, 2015).

From equation (1.2) BFAST takes all other components relatively trend and seasonal component to be randomised (R_p) and the equation was expressed as

$$Y_P = T_p + S_p + R_p \tag{1.2}$$

The residual random consist of cyclical and irregular component (Zdravevski, Lameski, Mingov, Kulakov & Gjorgjevikj, 2015).

To generate trend components using BFAST, we need a piecewise linear model approach. Suppose T_p is a piecewise linear model with an actual slope and intercept on q+1 segments broken with q breakpoints and P period; $p_1^{\neq}$,………., $p_q^{\neq}$ then T_p can takes the form

$T_p = \alpha_k + \beta_k \mathrm{P}$

Where $p_{k-1}^{\neq} < \mathrm{p} \leq p_k^{\neq}$

And If k = 1,…..,q then $p_0^{\neq} = 0$ and $p_{q+1}^{\neq} = \mathrm{n}$.

The slope of the change before the breakpoints while β_{k-1} and the slope of the breaks after the change breakpoints are β_k. The intercept and the slop of the linear model α_k and β_k with time period p and it will be used to derive the magnitude and direction of change.

To generate seasonal components using BFAST, we need a simple harmonic model.

Thus, S_p can be represented by a simple harmonic model with j terms; j = 12….J and time t.

$$S_p = \sum_{j=1}^{J} \omega_{k,j} \ \mathrm{Sin}\left(\frac{2\pi jt}{F} + \sigma_{K,j}\right) \tag{1.3}$$

Where k = 1… q, $p_{k-1}^{\neq} < \mathrm{p} \leq p_k^{\neq}$ and also $\omega_{k,j}$, $\sigma_{K,j}$ are the segment amplitude and F is the frequency (Zeileis, Kleiber, Krämer & Hornik, 2003).

To generate random components, any data that does not belong to trend nor seasonal is classified random R_p.

$$Y_P = \underbrace{\{\alpha_k + \beta_k P\}}_{T_p} + \underbrace{\left\{\sum_{j=1}^{J} \omega_{k,j} \quad \text{Sin}\left(\frac{2\pi jt}{F} + \sigma_{K,j}\right)\right\}}_{S_p} + \underbrace{R_p}_{R_p} \quad (1.4)$$

$$Y_P = T_p + S_p + R_p$$

The new technique called GFTSC considered splitting the random into cyclical components and irregular components which is an extension of BFAST. This was done through the inclusion of two new components.

To calculate cyclical components, center moving average is involved (Bornhorst, Dobrescu, Fedelino, Gottschalk & Nakata, 2011).

Derivation of cyclical code, let CMA be the center moving average of t objects, then CMA can be computed as follow

$$\text{CMA} = \sum_{t}^{n} \frac{Yt}{nt} \quad (1.5)$$

$$C_p = \frac{CMA}{\overset{\wedge}{CMA}} \quad (1.7)$$

After extracting the trend, seasonal and cyclical components, the left out components is called irregular components, the new equation becomes

$$Y_P = \underbrace{\{\alpha_k + \beta_k P\}}_{T_p} + \underbrace{\{\sum_{j=1}^{J} \omega_{k,j} \quad \text{Sin}\,(\frac{2\pi jt}{F} + \sigma_{K,j})\}}_{S_p} + \underbrace{\{\frac{CMA}{\overset{\wedge}{CMA}}\}}_{C_p} + \{\underbrace{I_p}_{I_p}\} \quad (1.8)$$

$$Y_P = T_p + S_p + C_p + I_p$$

For identification of Y_P , S_p , C_p , and I_p (See the paper: Box, Jenkins, Reinsel, & Ljung, 2015 :Maggi, 2018; Cleveland & Tiao, 1976; Caiado, 2009; Bohn, 1995; Cipra, & Romera, 1997).

The first stage in forecasting is to view the data and to examine all the components of time series present in that data in order to select the most appropriate forecasting technique. The UK yearly quarterly GDP data components identification was carried out with the help of the new technique called BFTSC. This new technique helps to have a clear image of the entire variations presents in the time series data.

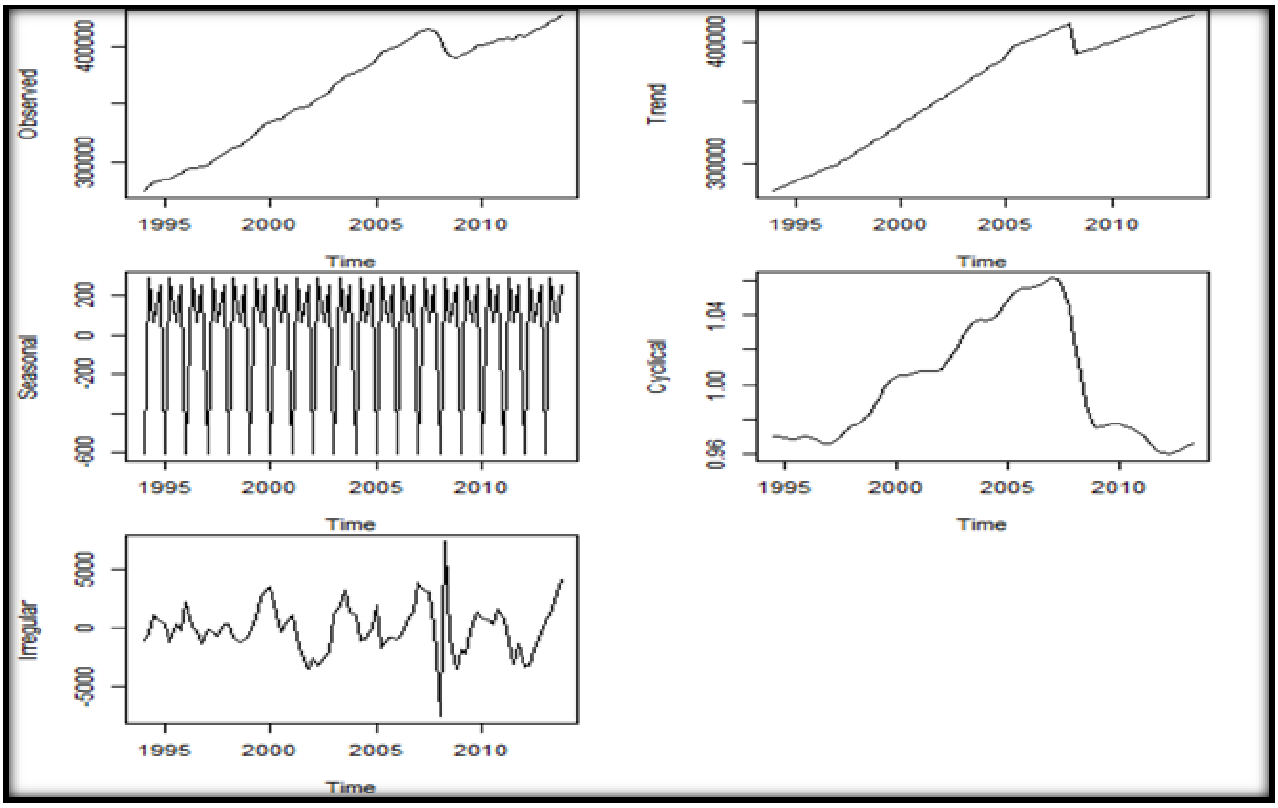

Fig 1 GFTSC for Quarterly UK GDP Data for 20 Years

ISSN No:-2456-2165

Figure 1 reveals all the time series components hidden in the quarterly UK GDP data for 20 Years, the image in the figure above indicate the presence of trend, seasonal, cyclical and irregular components, Hence the most appropriate techniques for analyzing such data is ARIMA.

Ten ARIMA models were fitted and the best model was selected based on the ARIMA with the smallest AIC (Akaike's Information Criterion). Based on the AIC models, the ARIMA(1,2,3) is the best model to be used in fitting the UK quarterly GDP . ARIMA(1,2,3) is selected and used for fitting the model .

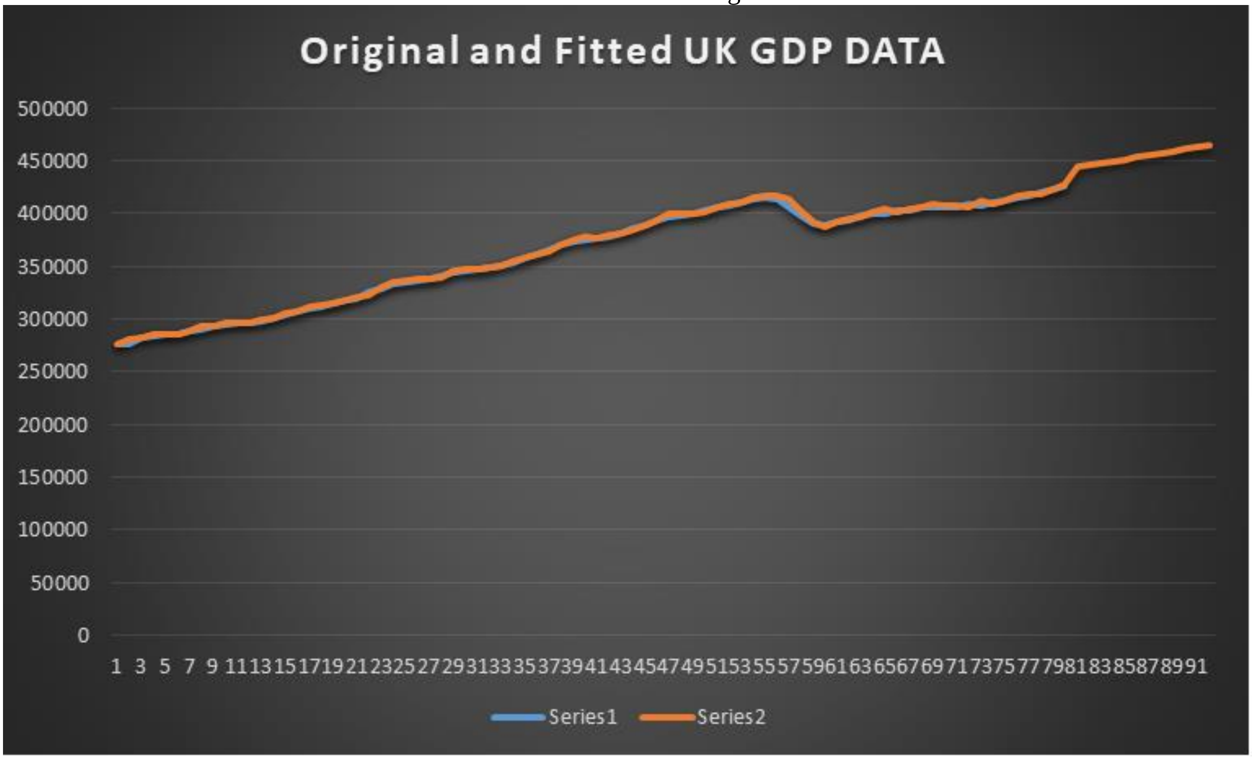

Fig 2 Original UK GDP Data and the Fitted Value

Figure 2. The fitted value and the real data of the yearly quarter GDP data of United Kingdom gross domestic product (UK GDP). This reveal that for the next two years period I the United State, The United State GDP show no evidence of decline and the fitted value fit well and match intact to the original UK GDP data so the model can be applied for prediction of more quarterly years GDP of UK.

IV. DISCUSSION/CONCLUSION

Verbesselt, Hyndman, Newnham, and Culvenor (2010). The technique was for recognising Breaks for Additive Seasonal and Trend (BFAST). This technique helps to recognise trend breaks enclosed by the series. The essential guide of the BFAST technique is the decomposition of time series component into seasonal, trends and miscellany elements with the technique for recognising structural similarity and difference. Verbesselt et al. (2010) recommended that the technique of BFAST is for identifying topographical pattern and also for improvement to be applied in other related disciplines.

Jamali, Jönsson, Eklundh, Ardö, and Seaquist (2015) describe BFAST as not being capable of identifying topographical vegetation basic component perfectly, though satellite sensor image have made topographical vegetation data available for so many years but yet the detection of topographic trend and variation is not yet clearly defined. Chen (2006) suggested that, this may be due to the limited number of available trend and change detection techniques accessible, algorithm suitable in identifying and characterizing abrupt changes without sacrificing accuracy and efficiency.

Based on previous studies, BFAST is used for topographical green forest picture data at certain specific time. Introducing BFAST to time series data and how to implement BFAST on time series data which contain only one variable for each time is another form of challenge. BFAST is a technique that take in data and processed to extract each component point of the data, it would be reasonable to use BFAST for time series components identifications (Rikus, 2018; Gorelick, 2017; Zhu, 2017).

BFAST approach give a very considerable outcome and was recommend as a modern instrument for statistics information decomposition and detections but could not separate random noise and is a customized additive decomposition method, from all indication observed so far, it reveal that BFAST need to be extended for the purpose of coping with other varieties of uses (Tolsheden, 2018; Mok et al., 2017; Maus, Câmara, Appel & Pebesma, 2017).

Based on the result in the simulated data and the empirical analysis, GFTSC is the most appropriate for time series components identification. GFTSC is recommended as a good alternative to BFAST. This is because GFTSC identifies the four components of time series statistics which

ISSN No:-2456-2165

is one of the basic limitations of BFAST. Based on the forecast value for 2019 and 2020 , it reveal no scientific evidence of drop and crash in UK GDP so improvement can be establish to improve on the yearly quarterly UK GDP.

The contribution of this study to the scientific community is that the GFTSC gives good results that improve the weaknesses of the existing GFAST. GFTSC forecast output is more reasonable for effective policy making.

- *Note: The data, BFTSC and GFTSC can be made available based on request from the original author of this paper Dr. Ajare Emmanuel*

ACKNOWLEDGMENT

The authors thank the Universiti Utara Malaysia to perfect this article.

- *Authors Contributions*

Ajare Emmanuel Oloruntoba: Analyzing, producing the results and writing the paper.

- *Ethics*

This is the original manuscript; there will be no expectation of any ethical problems.

REFERENCES

[1]. Abbes, A. B., & Farah, I. R. (2017). Prediction Changes for Nonstationary Multi-Temporal Satellite Images using HMM. In *Handbook of Research on Geographic Information Systems Applications and Advancements* (pp. 387-406). IGI Global.

[2]. Adewoye, R., & Chapman, h. (2018) testing spectral variation hypothesis on the afromontane forest ecosystem of ngelnyaki, north eastern nigeria with landsat 8 (oli) and macro-ecological data.

[3]. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

[4]. Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day. Oakland, California.

[5]. Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of applied econometrics*, *18*(1), 1-22.

[6]. Bornhorst, F., Dobrescu, G., Fedelino, A., Gottschalk, J., & Nakata, T. (2011). When and how to adjust beyond the business cycle? A guide to structural fiscal balances. *IMF Technical Notes and Manuals*, *11*(02).

[7]. Bohn, H. (1995). The sustainability of budget deficits in a stochastic economy. *Journal of Money, Credit and Banking*, *27*(1), 257-271.

[8]. Buhalau, T. (2016). Detecting clear-cut deforestation using Landsat data: A time series analysis of remote sensing data in Covasna County, Romania between 2005 and 2015. *Student thesis series INES*.

[9]. Cesta, A., Cortellessa, G., Pecora, F., & Rasconi, R. (2005, May). Monitoring Domestic Activities with Scheduling Techniques. In *Proceedings of the 2nd*.

[10]. Cleveland, W. P., & Tiao, G. C. (1976). Decomposition of seasonal time series: A model for the Census X-11 program. *Journal of the American statistical Association*, *71*(355), 581-587.

[11]. Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for information Science and Technology*, *57*(3), 359-377.

[12]. Caiado, J. (2009). Performance of combined double seasonal univariate time series models for forecasting water demand. *Journal of Hydrologic Engineering*, *15*(3), 215-222. Cipra, T., & Romera, R. (1997). Kalman filter with outliers and missing observations. *Test*, *6*(2), 379-395.

[13]. Cipra, T., & Romera, R. (1997). Kalman filter with outliers and missing observations. *Test*, *6*(2), 379-395.

[14]. DeVries, B., Pratihast, A. K., Verbesselt, J., Kooistra, L., de Bruin, S., & Herold, M. (2013, June). Near real-time tropical forest disturbance monitoring using Landsat time series and local expert monitoring data. In *Analysis of Multi-temporal Remote Sensing Images, MultiTemp 2013: 7th International Workshop on the* (pp. 1-4). IEEE

[15]. Ewing, B. T., & Malik, F. (2013). Volatility transmission between gold and oil futures under structural breaks. *International Review of Economics & Finance*, *25*, 113-121.journal homepage: www.elsevier.com/locate/iref

[16]. Flicek, P., & Birney, E. (2009). Sense from sequence reads: methods for alignment and assembly. *Nature methods*, *6*(11s), S6.

[17]. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18-27

[18]. Jong, R., Verbesselt, J., Schaepman, M. E., & Bruin, S. (2012). Trend changes in global greening and browning: contribution of short-term trends to longer-term change. *Global Change Biology*, *18*(2), 642-655. DOI: 10.1111/j.1365-2486.2011.02578.x

[19]. Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *22*(1), 4-37. **DOI:** 10.1109/34.824819

[20]. Jamali, S., Jönsson, P., Eklundh, L., Ardö, J., & Seaquist, J. (2015). Detecting changes in vegetation trends using time series segmentation. *Remote Sensing of Environment*, *156*, 182-195.

[21]. doi.org/10.1016/j.rse.2014.09.010

[22]. Porter, J., & Zhang, L. (2018). BisPin and BFAST-Gap: Mapping bisulfite-treated reads. *bioRxiv*, 284596.

[23]. Tolsheden, J. (2018). Detecting and testing for Seasonal breaks in Quarterly National accounts: Based on X-12-ARIMA and BFAST Methods.

[24]. Maggi, L. M. B. (2018). Times Series Analysis. In *Multiscale Forecasting Models* (pp. 1-29). Springer,

[25]. Mok, T. S., Wu, Y. L., Ahn, M. J., Garassino, M. C., Kim, H. R., Ramalingam, S. S., ... & Lee, C. K. (2017). Osimertinib or platinum–pemetrexed in EGFR T790M–positive lung cancer. *New England Journal of Medicine*, *376*(7), 629-640.

[26]. Maus, V., Câmara, G., Appel, M., & Pebesma, E. (2017). dtwSat: Time-Weighted Dynamic Time Warping for Satellite Image Time Series Analysis in R. *Journal of Statistical Software*.

[27]. Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. *Remote sensing of Environment*, *114*(1), 106-115. doi.org/10.1016/j.rse.2009.08.014

[28]. Rikus, L. (2018). A simple climatology of westerly jet streams in global reanalysis datasets part 1: mid-latitude upper tropospheric jets. *Climate Dynamics*, *50*(7-8), 2285-2310.

[29]. Verbesselt, J., Zeileis, A., Hyndman, R., & Verbesselt, M. J. (2012). Package 'bfast'

[30]. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint*.

[31]. Zewdie, W., Csaplovics, E., & Inostroza, L. (2017). Monitoring ecosystem dynamics in northwestern Ethiopia using NDVI and climate variables to assess long term trends in dry land vegetation variability. *Applied Geography*, *79*, 167-178.

[32]. Zdravevski, E., Lameski, P., Mingov, R., Kulakov, A., & Gjorgjevikj, D. (2015, September). Robust histogram-based feature engineering of time series data. In *Computer Science and Information Systems (FedCSIS), 2015 Federated Conference on* (pp. 381-388). IEEE.

[33]. Zhao, G., Li, E., Mu, X., Wen, Z., Rayburg, S., & Tian, P. (2015). Changing trends and regime shift of streamflow in the Yellow River basin. *Stochastic environmental research and risk assessment*, *29*(5), 1331-1343.

[34]. Zeileis, A., Kleiber, C., Krämer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, *44*(1-2), 109-123.