

# MODELING AND FORECASTING CUMULATIVE EVI ANOMALIES USING SARIMA FOR BIOPHYSICAL MONITORING: A CASE STUDY IN THE PHILIPPINES

A. J. L. Diccion, J. Z. Duran

School of Advanced Studies- Environmental Engineering Program, Saint Louis University, Baguio City, Philippines – (2130446, jzduran@slu.edu.ph)

**KEY WORDS:** cumulative EVI anomalies, modeling, forecasting, SARIMA, biophysical monitoring, remote sensing

## ABSTRACT:

Understanding changes in vegetation cover that affect the biophysical conditions of a region can help in formulating policies to address current and future problems of terrestrial ecosystems such as deforestation and environmental degradation. This study focuses on developing a model that forecasts the cumulative Enhanced Vegetation Index (EVI) anomalies as a tool for biophysical conditions monitoring in the Philippines. Satellite data from MODIS MYD13Q1 V6, which contains vegetation index per pixel at 16-day intervals with a resolution of 250 meters, were utilized. The cumulative EVI anomalies per instant were calculated in Google Earth Engine by aggregating the difference of a specific data point in 2011-2020 to a reference EVI mean computed from 2001-2010. The Error-Trend-Seasonality model shows that the cumulative EVI anomalies graph is non-stationary with an upward trend and seasonality. The upward trend of the cumulative EVI anomalies indicates the improvement of vegetation in the Philippines. To check the stationarity of the cumulative EVI anomalies data, the Augmented Dickey-Fuller test was utilized and the model was generated using Seasonal Autoregressive Integrated Moving Average model. Based on the analysis, the best-fit model for the cumulative EVI anomalies is SARIMA (1,1,0)(1,1,1)<sub>12</sub> with a mean absolute percentage error (MAPE) of 13.26%. Thus, the proposed model can be used as a tool for biophysical assessment by monitoring and forecasting changes in vegetation and contribute to attaining the UN Sustainable Development Goals 2 and 15 – ‘Eliminating Hunger’ and ‘Life on Land’.

## 1. INTRODUCTION

Satellite remote sensing is an effective tool to gain understanding of environmental impacts and challenges that acquires a collection of images over wide areas taking quick measurements. This method has the advantage of assessing large areas, where access in a forested area and collection of field measurements are limited. Sensors installed at the satellite collect information at a certain distance from the earth using the absorbed, reflected, and emitted electromagnetic radiation at different wavelengths of the areas under observation (Vorovencii, 2011). There has been a growing number of studies conducted utilizing geospatial technologies such as satellite remote sensing that has significantly improved the understanding of biophysical conditions such as biodiversity conservation, agriculture, forestry (Xue, and Su, 2017), hydrological modeling, watershed mapping, urban modeling, predicting droughts (Melesse, 2007) and mapping natural hazards and disasters (Joyce, 2009). It is vital to ensure the synergy of ecology and remote sensing to ensure effective environmental management (Pettorelli., Safi, and Turner, 2014).

Vegetation cover plays a critical role in sustaining life on earth by acting as a shelter for fauna, and minimizing air and water pollution (Abujayyab, Karaş, 2019). However, terrestrial ecosystems face environmental deforestation, destruction, and degradation, which are the main sources of greenhouse gas emissions (Poortinga et al., 2018) resulting from harmful anthropogenic activities (Šalić, A., and Zelić, 2018). For the past decade, there has been a reduction in the rate of deforestation in tropical regions. Strikingly, forest areas in Asia have significantly increased which indicates a thriving ecosystem (UN, 2021).

Several challenges for biophysical conditions monitoring include limited accessibility and the unavailability of in situ measurements of the environment (Pastor-Guzman, Dash, and Atkinson, 2018). The biosphere consisting of flora and fauna may experience long-haul effects due to changes in vegetation. The local climate of an area varies with the changes and transition of vegetation cover which depends on several factors such as timing and location (Duveiller et al., 2018). The prediction of changes in vegetation cover will help to provide timely insights for decision-makers to effectively manage vegetation in a region (Abujayyab, Karaş, 2019).

The commonly used metrics for vegetation phenology and dynamics fall in to three categories of indices: ratio and combined, physically based and adjusted vegetation, which undergoes further smoothing that includes i. empirical method which upholds assumptions within a time frame, ii. curve fitting method which utilizes mathematical mapping, and iii. data transformation which utilizes mathematical transformations and manipulations (Zeng et al., 2020).

One of the established approaches to quantify changes in vegetation using Satellite remote sensing is computing cumulative EVI anomalies. The Ecological Monitoring Application developed by Poortinga et al. (2018) used the Before-After, Control-Impact (BACI) method which utilizes cumulative EVI anomalies to quantify vegetation changes. However, calculating cumulative anomaly with satellite data is computationally expensive.

This paper focuses on developing a spatiotemporal model of the cumulative EVI anomalies in the Philippines to fully understand how the vegetation cover changes over time as a tool for

biophysical assessment to give timely insights to determine the suitable vegetation management. The model is useful for short-term forecasting and the data must be updated to ensure the accuracy of the forecasted values. Furthermore, the biophysical factors that influence changes in vegetation are not incorporated in the model.

## 2. METHODOLOGY

### 2.1 Data

Satellite images were taken from the enhanced vegetation index (EVI) vegetation layer of the Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) Vegetation Indices (MYD13Q1 V6), which contains vegetation index per pixel at 16-day intervals with a resolution of 250 meters (Didan, 2015). The EVI vegetation layer has improved sensitivity over high biomass regions and minimizes variations in canopy background (Poortinga et al., 2018) which is suitable for the Philippine Archipelago.

### 2.2 Cumulative EVI Anomalies

The EVI anomalies can be interpreted as a deviation from the mean (Xiao et al., 2003) or reference which is a climatological change approach that follows BACI method to quantify changes in vegetation (Poortinga et al., 2018). The baseline or reference EVI ( $EVI_{ref}$ ) for the Philippines was computed by taking the mean of the EVI data from 2001-2010 using Equation (1).

$$EVI_{ref} = 1/N (EVI_{d1} + EVI_{d2} + \dots + EVI_{dN}) \quad (1)$$

where  $EVI_{ref}$  = EVI reference  
 $N$  = total number of EVI observations from 2001-2010  
 $EVI_d$  = 16-day EVI observations

Afterwards, the EVI anomalies ( $EVI_a$ ) are derived by getting the difference of EVI from 2011 to 2020 and the reference EVI using Equation (2).

$$EVI_a = EVI_d - EVI_{ref} \quad (2)$$

where  $EVI_a$  = EVI anomalies  
 $EVI_d$  = 16-day EVI observations  
 $EVI_{ref}$  = EVI reference

Then, the cumulative EVI ( $EVI_c$ ) anomalies will be the aggregate of all the EVI anomalies in the country shown by Equation (3).

$$EVI_c = \sum EVI_a \quad (3)$$

where  $EVI_c$  = cumulative EVI anomalies  
 $EVI_a$  = EVI anomalies

The positive (+) value of cumulative EVI anomalies indicates an overall improvement of a certain region. A negative (-) value indicates an overall reduction of vegetation in a region. The map of the cumulative anomalies from 2010-2020 was generated by aggregating all cumulative anomalies (Poortinga et al., 2018).

The cumulative EVI anomalies of the MODIS MYD13Q1 were computed using the Google Earth Engine (GEE) platform. The GEE is a platform deployed in the cloud that is capable of running high computing of geospatial analysis on a global scale which can be used for applications such as mapping of malaria risk,

flooding and rice paddy, and monitoring land use, land cover, climate and changes in planetary surface water and forest. The platform utilizes a user-friendly web interface which can directly access and work with data either privately or publicly available (Gorelick et al., 2017).

### 2.3 SARIMA modeling

A supplement of the ARIMA model is the seasonal autoregressive integrated moving average (SARIMA) model (Mwanga, Ong'ala, and Orwa, 2017), which can be applied to time series having trend, seasonality, and periodicity (Lou et al., 2013) denoted as

$$ARIMA(p, d, q)(P, D, Q)_s \quad (4)$$

where

$p$  = the non-seasonal Autoregressive (AR) order;  
 $d$  = the non-seasonal differencing order;  
 $q$  = the non-seasonal moving average (MA) order;  
 $P$  = the seasonal AR order;  
 $D$  = the seasonal differencing order;  
 $Q$  = the seasonal MA order; and  
 $S$  = the number of periods in a season

The steps to determine the appropriate SARIMA model (Nai et al., 2017) are the following: (i) test for stationarity, (ii) time series differencing for data that are not stationary, (iii) evaluate the data after differencing using the autocorrelation function (ACF) and partial autocorrelation function (PACF), and (iv) choose the best model using Akaike information criterion (AIC) and Bayesian information criterion (BIC) which optimizes the model parameters Kullback-Leibler divergence and the posterior model probability respectively (Ding, Tarokh, and Yang, 2017).

The cumulative EVI anomalies plot was further analyzed by performing time series decomposition to identify the trend or pattern, seasonality, and residual of the data using the Error-Trend-Seasonality (ETS) model (Mahajan, Rastogi, and Sharma, 2020). The stationarity of the data was assessed using the Augmented Dickey-Fuller (ADF) test and the non-stationary data was transformed to stationary data using Differencing technique (Mohamed, 2008). The mean absolute percentage error (MAPE) and root mean squared error (RMSE) were computed to validate the best-fit model to ensure its accuracy.

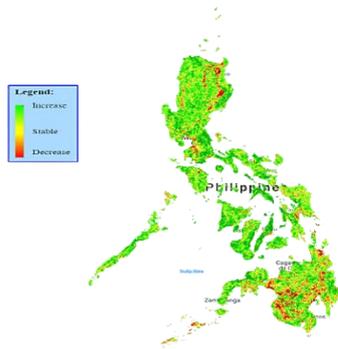
Time series decomposition, stationarity test, differencing and SARIMA modeling were done in Google Collaboratory or simply Google Colab which is a free service from Google operating as cloud-based Jupyter Notebooks and offering computational resources that have Artificial Intelligence libraries loaded in advance like matplotlib, NumPy, pandas, Torch, TensorFlow, and other libraries (Nelson and Hover, 2020).

## 3. RESULTS AND DISCUSSION

### 3.1 EVI Anomalies in the Philippines

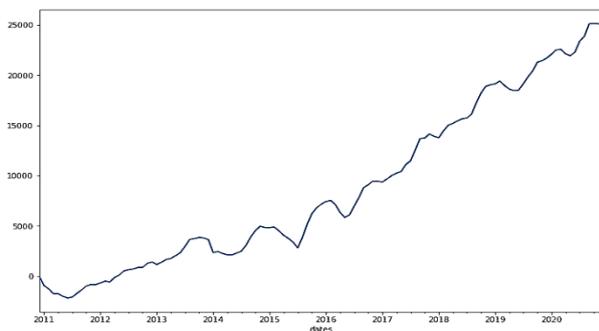
The cumulative EVI Anomalies for 2011-2020 is shown in Figure 1 which is generated from Google Earth Engine. Certain parts of the country experience a decrease in vegetation as shown by the hue of red in the map, which may be attributed to anthropogenic disturbances that includes urbanization, cutting of

trees, changes in land use, and climate change and global warming (Liu et al., 2019).



**Figure 1.** Cumulative EVI anomalies in the Philippines.

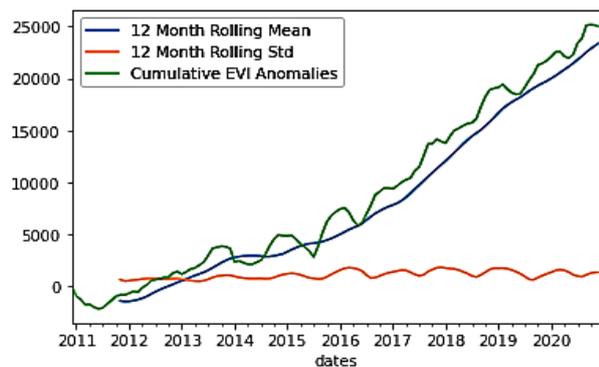
The cumulative EVI anomalies plot is shown in Figure 2. There are negative cumulative EVI anomalies values from 2011-2012 which means that during the period, there was an overall reduction of vegetation in the country. After the 1<sup>st</sup> quarter of 2012, the overall health of the vegetation in the country slowly increased as shown by the positive cumulative EVI anomalies values computed. The increasing value of the plot supports that there is an increase in vegetation in the Philippines (UN, 2021).



**Figure 2.** Cumulative EVI anomalies plot of the Philippines for 2011-2020.

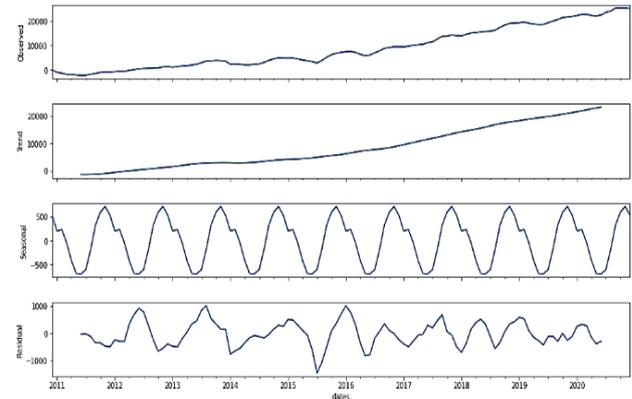
### 3.2 Model Construction and Analysis

The EVI cumulative plot has been transformed from 16-day observations to monthly time series data. As indicated in Figure 3, the mean of the cumulative EVI anomalies plot is trending upward and the standard deviation has minimal variability.



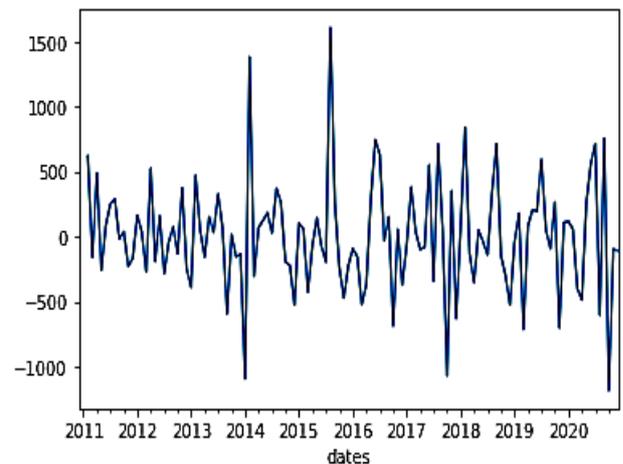
**Figure 3.** Cumulative EVI anomalies rolling statistics plot.

The ETS decomposition in Figure 4 shows that the cumulative EVI anomalies plot has an uptrend and seasonal component. A negative seasonal component can be observed in the dry season and a positive seasonal component in the rainy season. Moreover, the average of the residual component is computed to be zero.



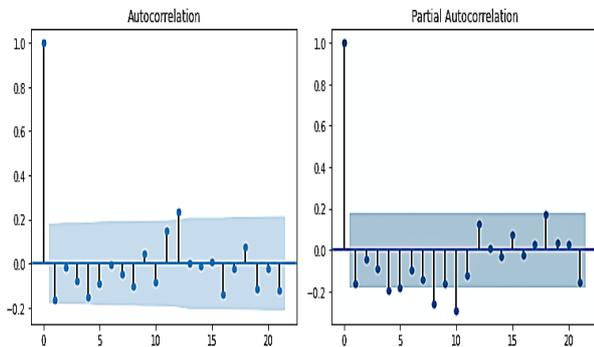
**Figure 4.** ETS decomposition of the cumulative EVI anomalies.

After the first differencing, the Augmented Dickey-Fuller test was performed, and the ADF statistic computed is 2.5240 with a p-value of 0.99906. Since the computed p-value is greater than 0.05, the data is non-stationary. Second differencing was performed to make the data stationary. The plot of the cumulative EVI anomalies after second differencing is shown in Figure 5.



**Figure 5.** Cumulative EVI anomalies plot after second differencing.

Figure 6 shows the ACF and PACF plots after second differencing. Since stationarity is achieved in second differencing, the researchers considered the parameter  $d = D = 1$ . With both ACF and PACF plots showing a positive peak at lag 0 and diminishes significantly at lag 1,  $P = Q = 1$  will be used for this case. In addition,  $p$  may take either a value of 2 or 1 while  $q = 0$ . Through the analysis, the suitable models for the cumulative EVI anomalies are SARIMA(2,1,0)(1,1,1)<sub>12</sub> and SARIMA(1,1,0)(1,1,1)<sub>12</sub>.



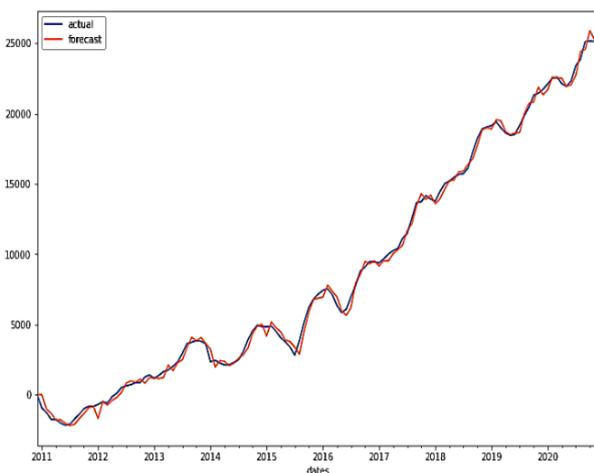
**Figure 6.** ACF and PCF plots.

The AIC and BIC of each of the suggested models for the cumulative EVI Anomalies are presented in Table 1. The best-fit model for the cumulative EVI anomalies is SARIMA(1,1,0)(1,1,1)<sub>12</sub>, since it has the lowest AIC and BIC values between the two models.

MODEL	AIC	BIC
SARIMA(2,1,0)(1,1,1) <sub>12</sub>	1589.833	1603.243
SARIMA(1,1,0)(1,1,1) <sub>12</sub>	1587.859	1598.588

**Table 1.** Comparison of SARIMA models.

Figure 7 presents the forecasted data generated by the model. The computed MAPE and RMSE are 13.26% and 361.3796, respectively which tells us that SARIMA(1,1,0)(1,1,1)<sub>12</sub> is a good model for the cumulative EVI anomalies in the Philippines.



**Figure 7.** Forecasted and actual cumulative EVI anomalies.

It should be noted that biophysical factors that influence changes in vegetation are not incorporated in the model. The model is useful for short-term forecasting and the data must be updated to ensure the accuracy of the forecasted values.

### 3.3 Implications on Sustainable Development

The United Nations has set the Sustainable Development Goals (UN SDGs) aiming for a nation’s economic growth, protect the

health of the inhabitants and the environment, alleviate poverty, eliminate hunger, and promote social equality and justice, among others. Furthermore, Holloway and Mengersen (2018) have pointed out the SDG targets with the corresponding remote sensing applications and indicators. SDG 2 and 15, pertaining to ‘eliminating hunger’ and ‘life on land’, respectively, points out achieving food security and improved nutrition through resilient agricultural practices, adaptation to climate change, and progressively improving land and soil quality, among others. This accounts the protection, restoration, and responsible usage of ecosystems, while conforming to both national and international laws, as well the sustainable development goal indicators.

The vegetation growth or degradation may be correlated with socio-economic and environmental factors, such as the production and consumption, affluence, climate change, among others. The Philippines is a tropical country situated in Southeast Asia, endowed with wide array of natural resources. Given the Philippines is considered as an agricultural country, remotely-sensed data may be beneficial in optimizing the growth and harvest of crops, together with past data generated coming from government agencies. Similarly, both developed and developing economies may utilize statistical models to map out natural phenomenon, such as floods, droughts, wildfires, to name a few. This will help to plan when and where to plant and harvest. Paddy rice growth and development in South Korea was monitored using high-temporal-resolution Geostationary Ocean Colour Imagery (GOCI) (Yeom & Kim, 2015). In Arid regions, droughts are more apparent due to lack of precipitation. Land management in the said areas may be well supported using remote sensing through examining water supply pattern throughout the years, in consideration with the climate variability, crop management, and livestock production (Stanimirova et al., 2019). This said, food security may be addressed with forecasting through agricultural activities patterned using generated models (de Azevedo Silva et al., 2021).

In terms of plastic waste, build-up of plastic barges on the coastlines of countries may also be monitored. This is of concern because of the high probability of plastic degradation that may affect the marine life through microplastics. The minute polymeric pollutants may be ingested by fishes and other marine species, that may lead to eventual consumption of humans. Over time, the accumulation of such microplastics in living organisms may cause irreversible effects (Bucol et al., 2020; Davaasuren et al., 2018; Walker, 2021).

Health-related issues may also be monitored and mitigated using SARIMA. Areas with disease and virus outbreaks may be analyzed using remotely sensed environmental indicators, such as that in Ethiopia, where relationships among malaria cases, rainfall, vegetation, and ambient temperature were quantified over time (Midekisa et al., 2012). In conjunction with the usage of statistical and machine learning algorithms, epidemiological models pertaining to mosquito-borne diseases may be replicated with ease in different geographic contexts (Parselia et al., 2019).

In general, socio-economic and environmental effects that have relationships with global climate change may be monitored and examined to reduce or eliminate the impacts by creating policies and guidelines pertaining to land usage, with the utilization of satellite remote sensing data as a tool for decision-making.

#### 4. CONCLUSION

The cumulative EVI anomalies can be used as a tool for the overall health assessment of a certain region. In this paper, the researchers proposed the SARIMA model for cumulative EVI anomalies in the Philippines as a tool for monitoring and forecasting changes in vegetation. The results of this study can help policymakers better understand changes in vegetation and formulate proper management and interventions. In addition, other socio-economic and environmental issues, such as food security, risk reduction, disease outbreaks, and waste build-up, among others, may be monitored and further examined. Thus, the cumulative EVI anomalies information can significantly contribute to the growth of a nation and conform to the UN Sustainable Development Goals, particularly goals 2 and 15 – ‘Eliminating Hunger’ and ‘Life on Land’ through monitoring terrestrial ecosystems.

#### ACKNOWLEDGEMENTS

We thank the School of Advanced Studies- Engineering Program of Saint Louis University for supporting this paper.

#### REFERENCES

- Abujayyab, S.K. and Karaş, İ.R., 2019. Automated Prediction System For Vegetation Cover Based On MODIS-NDVI Satellite Data And Neural Networks. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, pp.9-15.
- Bucol, L. A., Romano, E. F., Cabcaban, S. M., Siplon, L. M. D., Madrid, G. C., Bucol, A. A., & Polidoro, B., 2020. Microplastics in marine sediments and rabbitfish (*Siganus fuscus*) from selected coastal areas of Negros Oriental, Philippines. *Marine Pollution Bulletin*, 150, 110685.
- Davaasuren, N., Marino, A., Boardman, C., Alparone, M., Nunziata, F., Ackermann, N., & Hajnsek, I., 2018. Detecting microplastics pollution in world oceans using SAR remote sensing. 938–941.
- De Azevedo Silva, P. A., de Carvalho Alves, M., Sáfadi, T., & Pozza, E. A., 2021. Time series analysis of the enhanced vegetation index to detect coffee crop development under different irrigation systems. *Journal of Applied Remote Sensing*, 15(1), 014511.
- Didan, K. 2015. MYD13Q1 MODIS/Aqua Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2021-09-08 from <https://doi.org/10.5067/MODIS/MYD13Q1.006>.
- Ding, J., Tarokh, V. and Yang, Y., 2017. Bridging AIC and BIC: a new criterion for autoregression. *IEEE Transactions on Information Theory*, 64(6), pp.4024-4043.
- Duveiller, G., Forzieri, G., Robertson, E., Li, W., Georgievski, G., Lawrence, P., Wiltshire, A., Ciaia, P., Pongratz, J., Sitch, S. and Arneth, A. 2018. Biophysics and vegetation cover change: a process-based evaluation framework for confronting land surface models with satellite observations. *Earth System Science Data*, 10(3), pp.1265-1279.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, pp.18-27.
- Holloway, J., & Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sensing*, 10(9), 1365.
- Joyce, K.E., Belliss, S.E., Samsonov, S.V., McNeill, S.J. and Glassey, P.J., 2009. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Progress in physical geography*, 33(2), pp.183-207.
- Liu, S., Huang, S., Xie, Y., Wang, H., Huang, Q., Leng, G., Li, P. and Wang, L., 2019. Spatial-temporal changes in vegetation cover in a typical semi-humid and semi-arid region in China: Changing patterns, causes and implications. *Ecological indicators*, 98, pp.462-475.
- Mahajan, A., Rastogi, A. and Sharma, N., 2020. Annual Rainfall Prediction Using Time Series Forecasting. In *Soft Computing: Theories and Applications* (pp. 69-79). Springer, Singapore.
- Melesse, A.M., Weng, Q., Thenkabail, P.S. and Senay, G.B., 2007. Remote sensing sensors and applications in environmental resources mapping and modelling. *Sensors*, 7(12), pp.3209-3241.
- Midekisa, A., Senay, G., Henebry, G. M., Semuniguse, P., & Wimberly, M. C., 2012. Remote sensing-based time series models for malaria early warning in the highlands of Ethiopia. *Malaria Journal*, 11(1), 1–10.
- Mohamed, I.E., 2008, September. Time series analysis using SAS-Part I-The Augmented Dickey-Fuller (ADF) Test. In *SAS Conference Proceedings*, Pittsburgh, PA.
- Mwanga, D., Ong’ala, J. and Orwa, G., 2017. Modeling sugarcane yields in the Kenya sugar industry: A SARIMA model forecasting approach. *International Journal of Statistics and Applications*, 7(6), pp.280-288.
- Nai, W., Liu, L., Wang, S. and Dong, D., 2017. An EMD–SARIMA-based modeling approach for air traffic forecasting. *Algorithms*, 10(4), p.139.
- Nelson, M.J. and Hoover, A.K., 2020. Notes on using Google Colaboratory in AI education. In *Proceedings of the 2020 ACM conference on innovation and Technology in Computer Science Education* (pp. 533-534)
- Parselia, E., Kontoes, C., Tsouni, A., Hadjichristodoulou, C., Kioutsoukis, I., Magiorkinis, G., & Stilianakis, N. I., 2019. Satellite earth observation data in epidemiological modeling of malaria, dengue and West Nile virus: A scoping review. *Remote Sensing*, 11(16), 1862.
- Pastor-Guzman, J., Dash, J. and Atkinson, P.M., 2018. Remote sensing of mangrove forest phenology and its environmental drivers. *Remote sensing of environment*, 205, pp.71-84.
- Pettorelli, N., Safi, K. and Turner, W., 2014. Satellite remote sensing, biodiversity research and conservation of the future.
- Poortinga, A., Clinton, N., Saah, D., Cutter, P., Chishtie, F., Markert, K.N., Anderson, E.R., Troy, A., Fenn, M., Tran, L.H.

and Bean, B., 2018. An Operational Before-after-control-impact (BACI) Designed Platform For Vegetation Monitoring At Planetary Scale. *Remote Sensing*, 10(5), p.760.

Šalić, A., and Zelić, B. 2018. Introduction to environmental engineering. *Environmental Engineering*: 1-16.

Stanimirova, R., Arévalo, P., Kaufmann, R. K., Maus, V., Lesiv, M., Havlík, P., & Friedl, M. A., 2019. Sensitivity of global pasturelands to climate variation. *Earth's Future*, 7(12), 1353–1366.

Walker, T. R., 2021. (Micro) plastics and the UN sustainable development goals. *Current Opinion in Green and Sustainable Chemistry*, 100497.

United Nations. 2021. Progress and Info. Department of Economic and Social Affairs. Accessed 09 Sep 2021 <https://sdgs.un.org/goals/goal15>

Vorovencii, I., 2011. Satellite remote sensing in environmental impact assessment: an overview. *Bulletin of the Transilvania University of Brasov. Forestry, Wood Industry, Agricultural Food Engineering. Series II*, 4(1), p.73.

Xiao, X., Braswell, B., Zhang, Q., Boles, S., Frolking, S. and Moore III, B., 2003. Sensitivity of vegetation indices to atmospheric aerosols: continental-scale observations in Northern Asia. *Remote sensing of Environment*, 84(3), pp.385-392.

Xue, J. and Su, B., 2017. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of sensors*, 2017.

Yeom, J.-M., & Kim, H.-O. (2015). Comparison of NDVIs from GOCI and MODIS data towards improved assessment of crop temporal dynamics in the case of paddy rice. *Remote Sensing*, 7(9), 11326–11343.

Zeng, L., Wardlow, B.D., Xiang, D., Hu, S. and Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, 237, p.111511.

## APPENDIX

The source code for generating the cumulative EVI anomalies and SARIMA modeling is available at: [https://github.com/tonydicc/Cumulative\\_EVI\\_Anomalies\\_Philippines](https://github.com/tonydicc/Cumulative_EVI_Anomalies_Philippines)