Understanding tree phenology using Sentinel-2A data

Introduction

Annually recurring biological timing in plants like flowering, bud-bursting, and leaf fall are known as tree phenology. Tree phenology is highly sensitive so they act as a visual indicator of climate change. Evidence of the early start of the growing season in temperate regions induced by climate change has been reported in many studies (Inouye, 2008; Roshbakh et. al, 2021). The biological events that occur early or late can also influence gaseous exchange, succession, carbon cycling, and ecosystem productivity (Kosugi, 2013; Keeling 1996; Richardson et. al, 2010). Also, the net primary production of forest and its spatial pattern are increased by long growing seasons (Nemani et. al, 2003).

Traditionally, phenological studies rely on human observation and noting down the data based on their subjective judgement, and were confined to a limited geographical extent. Satellite-based remote sensing can be a reliable tool to study and monitor the phenology of trees through continuous observation (Studer et. al, 2007). These satellites collect the data in the form of light waves reflected from earth's surface, and are recorded in digital numbers. Those digital numbers are then converted to unique spectral signatures and help in the identification of different objects on the earth’s surface (Karle et. al, 2004). Moreover, spectral reflectance is generated due to the ability of the target object to reflect electromagnetic spectrum under different wavelengths making it easy to detect. The reflectance is then used to generate different indices that best help to discern the target object. (Zhu et. al, 2018).

Although satellite remote sensing datasets are available in different spatial and temporal resolutions, datasets with high temporal resolution are preferred for phenological studies. Coarseresolution remote sensing datasets from Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High-Resolution Radiometer (AVHRR) are used to study phenological conditions over a landscape level (Ahl et. al, 2006, Heumann et. al, 2007, Maigan et. al, 2008, Zhang et. al, 2003). For a smaller geographical extent, medium-resolution datasets from Landsat or Sentinel images are popular for phenological study in a higher spatial and temporal resolution because of their smaller uncertainty and free availability (Younes et. al, 2021). Finally, this study aims to combine the phenological dynamics with remote sensing dataset and dig deeper to understand the impact of fertilizer on forest trees species and their capacity to retain leaf.

Research Questions

1. Is the spatiotemporal resolution of Sentinel 2 NDVI sufficient to accurately determine SOS and EOS?

2. Do we observe significant differences in Sentinel-2-based SOS and EOS dates between plots with different nutrient regimes?

Hypothesis

1. Remote sensing data helps to identify the day of SOS and EOS with accuracy.
2. We can detect the difference between SOS and EOS among plots under different nutrient regimes.

Data and Methods

Site description

For the ongoing Multiple Element Limitation in Northern Hardwood Ecosystems (MELNHE) project, full factorial Nitrogen (N) and Phosphorus (P) are added to nine stands of Bartlett Experimental Forest (BEF) in White Mountain National Forest, New Hampshire. At BEF, we have three replicates for three age classes: young (clearcut 1982-1990), mid-aged (clearcut 1970-1979), and mature (cut 1883-1915).

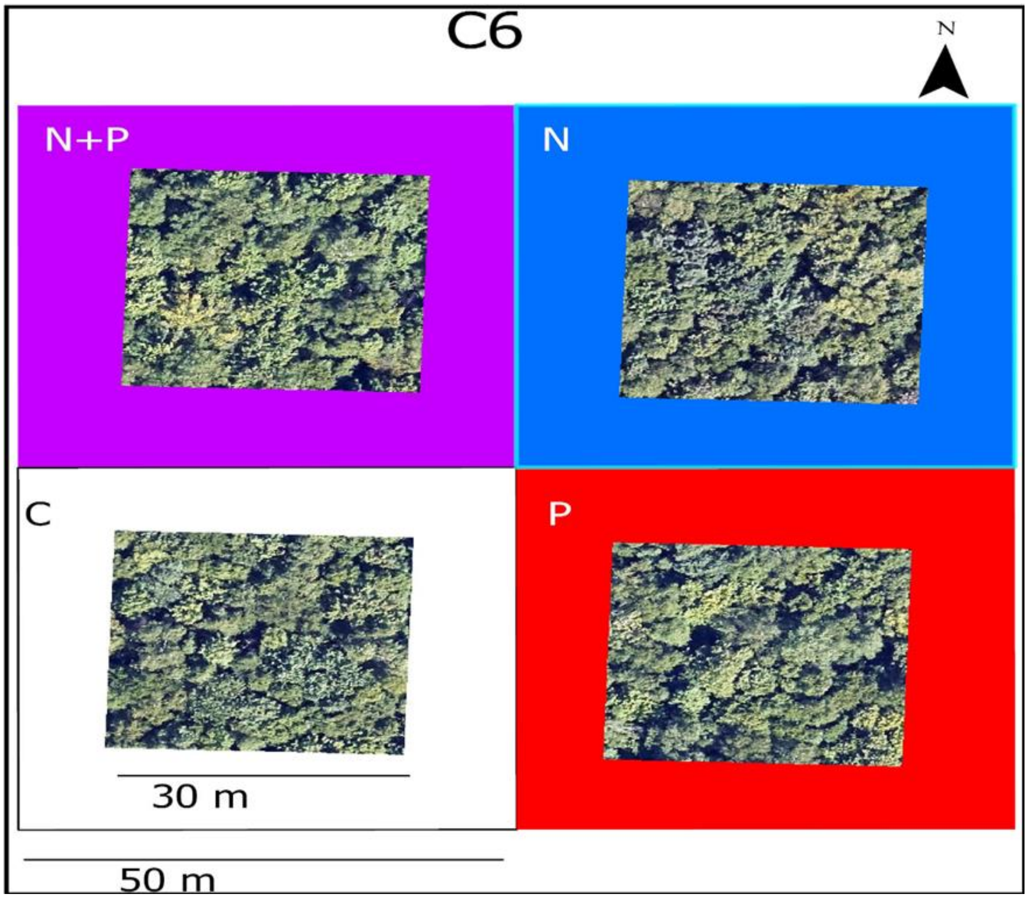

Figure 1: Map of study area showing one stand with all treatments

Each of the 9 stands has four treatment plots (50m x 50m) with N (30 kg N/ha/yr as NH_4NO_3), P (10 kg P/ha/yr as NaH_2PO_4), N+P at the same rates, and an untreated control since 2011. Generally, inner plots are 30m x 30m out of 50m x 50m allowing for 10m of buffer on each side of the plot. The area is mostly dominated by American beech (*Fagus grandifolia*), sugar maple (*Acer saccharum*), red maple (*Acer rubrum*), yellow birch (*Betula alleghaniensis*), white birch

(*Betula papyrifera*), and pin cherry (*Prunus pensylvanica*). The annual temperature is 5.60C and precipitation is 1400mm for this study site (Bailey et. al, 2003).

Methodology

For this study, sentinel images will be used because of their higher spatial (10 meters) and temporal resolution (5 days) compared to Landsat. European Space Agency (ESA) started a mission to develop next-generation earth observation under the collaboration of the European commission known as Copernicus. Under the program- Sentinel, satellites Sentinel-2A and 2B were launched in 2015 and 2017. The Copernicus Sentinel-2 mission has twin sun-synchronous satellites phased at 180^0 to each other. They are polar-orbiting satellites and have a swath width of 290 km. Each satellite has a revisit time of 10 days and when data are acquired from both satellites, we can acquire data from the region of interest every 5 days (https://scihub.copernicus.eu/). I will use Sentinel-2 MSI: MultiSpectral Instrument, Level-2A image collection by Sentinel.

It is common to use optical datasets and generate Normalized Difference Vegetation Index (NDVI) to study phenology (Xiao et. al, 2009). The mathematical relationship between Near Infrared (NI) and Red (R) through l bands of images generates because NDVI which depends on the chlorophyll content of the leaf resulting in higher values for leaves and lower for other earth surface. Continuous collection of data from satellites and the ability to manipulate those datasets makes it possible to conduct these kinds of studies. Many studies have investigated the impact of nutrient supply on the phenology of agricultural vegetation. Gungula et. al, (2003) study in Nigeria reported a delay in phenological activity under nitrogen stress in maize crops. Similarly, Moyo et. al, (2015) found that increased water supply and nutrient addition helps to extend the growing period of Terminalia sericea in South Africa. In contrast, there have been limited studies that examine phenological responses in forest ecosystems to varying nutrient conditions, and remote sensing technology has not been extensively used for this purpose. The existing studies are concentrated on using phenological properties for the classification of species (Mondensela et. a, 2017, Li et. al 2019) or changes in phenological trends due to climate change (Reed et al. 2009). Thus, established in the northern hardwood forest in the White Mountains of New Hampshire, U.S., the full factorial nutrient experimental plots can give a unique opportunity to compare the effect on tree phenology across different nutrient regimes in a temperate hardwood forest ecosystem.

Field Work

For this study, I collected corner stake locations for all MELNHE plots using Trimble Geo 7x Global Positioning System (GPS) during summer 2022. This device collects point locations under 1-2 meters of accuracy.

Lab Work

After precise point locations are collected, I prepared the boundary map for all the plots. The shapefile for the plot will be uploaded to Google Earth Engine so time-series satellite images can be collected from the delineated areas of interest. Most importantly, mean value for each plot will be used for the analysis. Data will be collected for a period of two-year (2019-2020) time range collected by Sentinel 2A and 2B from the start of the growing season to the end of the season.

Data analysis

I will use Google Earth Engine to select the images from the region of interest for the time interval from January 2019 to December 2021. Google Earth Engine (Gorelick et. al, 2017) has the feature of selecting any bands from the specified date, manipulating them as per our need and generating the graph. Google Earth Engine is a cloud computing platform which helps to provide access to super-computers and easily handles large datasets for analysis. For this study, NDVI from

Sentinel-2 data was calculated using band 8 (NIR) and band 4 (Red))values. These NDVI values act as an index for phenological cycle because they are higher during summer with all the leaf on it and minimum during winter because of leaf fall. Cloud cover was masked before calculating the vegetation indices as they can alter the result. Data were visualized and downloaded for further statistical analysis in R-studio 4.3.1 (R core Team, 2013).

After filtering the cloud cover data, the data was downloaded for each treatment plot. The data was imputed for each day using the stine interpolation method (Stineman, 1980). This method is considered better because it produces values that do not have more inflection points required by dataset. First, there were noise and outliers in the dataset. Remote sensing data have such issues because data are weather dependent. Smoothing the data helps to remove noise or outliers on the dataset. Savitsky-Golay filter for smoothening time-series data is considered best among others (Press and Teukolsky, 1990).

After smoothening the data, I tried to fit the model for the data. Further, the analysis will focus on characterizing the start of the season (SOS) and end of the season (EOS) by detecting the abrupt increase and decrease in slope of NDVI curve. For this study, we consider Start of Season (SOS) as the onset of the leaves in the trees. For the vegetation index, it will be the onset of the rising of the values after the constant values occurring due to leaf fall. As leaves start budding out, they will reach the maximum and will eventually fall out during winter. The vegetation index will have a peak followed by downfall and eventually a constant minimum value because of no leaves on the trees. This phase of decreasing VI followed by a constant minimum represents EOS. Break for Additive Season and Trends (BFAST) algorithm will be used to detect breakpoints for SOS and EOS (Verbesselt et. al, 2010). So, we will understand how precisely Bfast algorithm is able to SOS and EOS. Also, these breakpoints for the data across the plot will be compared to detect if there is any temporal difference subjected to the effect of fertilization.

Results

We have generated a plot from all the datasets acquired for one of our plot to observe the pattern of NDVI through the span of 3 years.

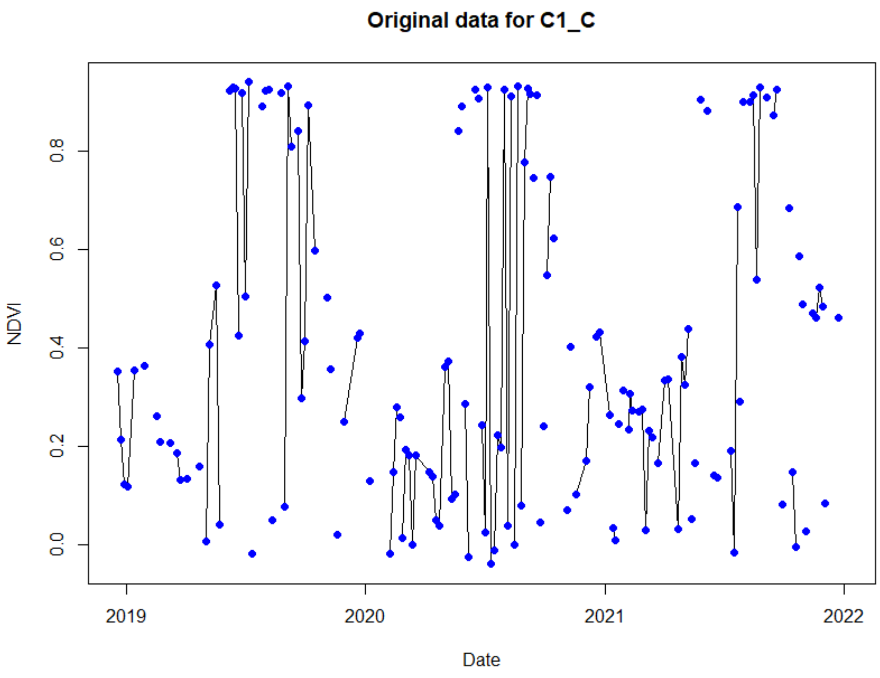

Figure 2: Original dataset acquired from Sentinel from 2019 till the end of 2021

After this, we interpolated the data to fill the gaps between the observed dataset. As Sentinel-2 is collecting data in 6 days intervals, still there could be a cloud on the day of data collection. These issues lead to an extended gap between two-time intervals for data collection. Thus, we interpolated the data for each day using stine method for all 3 years.

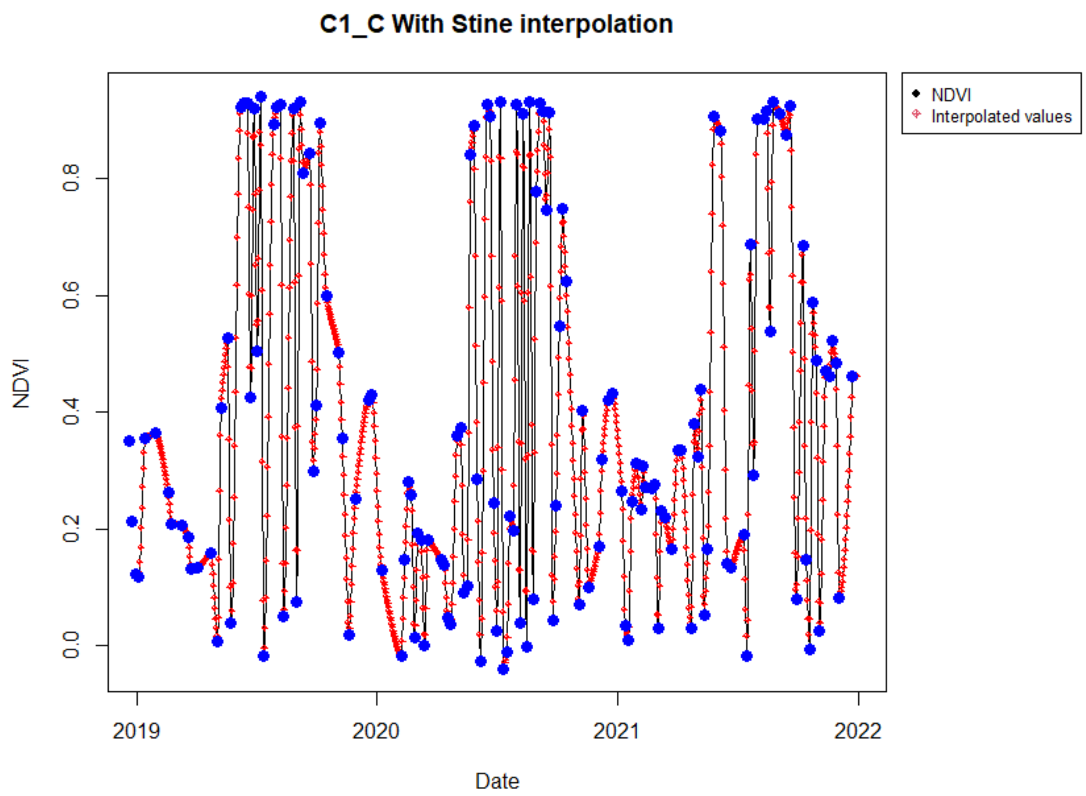

Figure 3: Interpolated NDVI value for daily data using stine method

For the analysis of time series data, noise can be an issue to interpret the results. They could be misleading and needs to be minimized. For noise reduction, we used Savitzky Golay smoothing method which is considered the best till date for time series data analysis. With hit and trial along with visual observation, two windows with 133 observations are best suited for smoothing. Two things are considered during data smoothing, the important information on patterns and trends should not be ignored or nullified while excluding as much noise as possible.

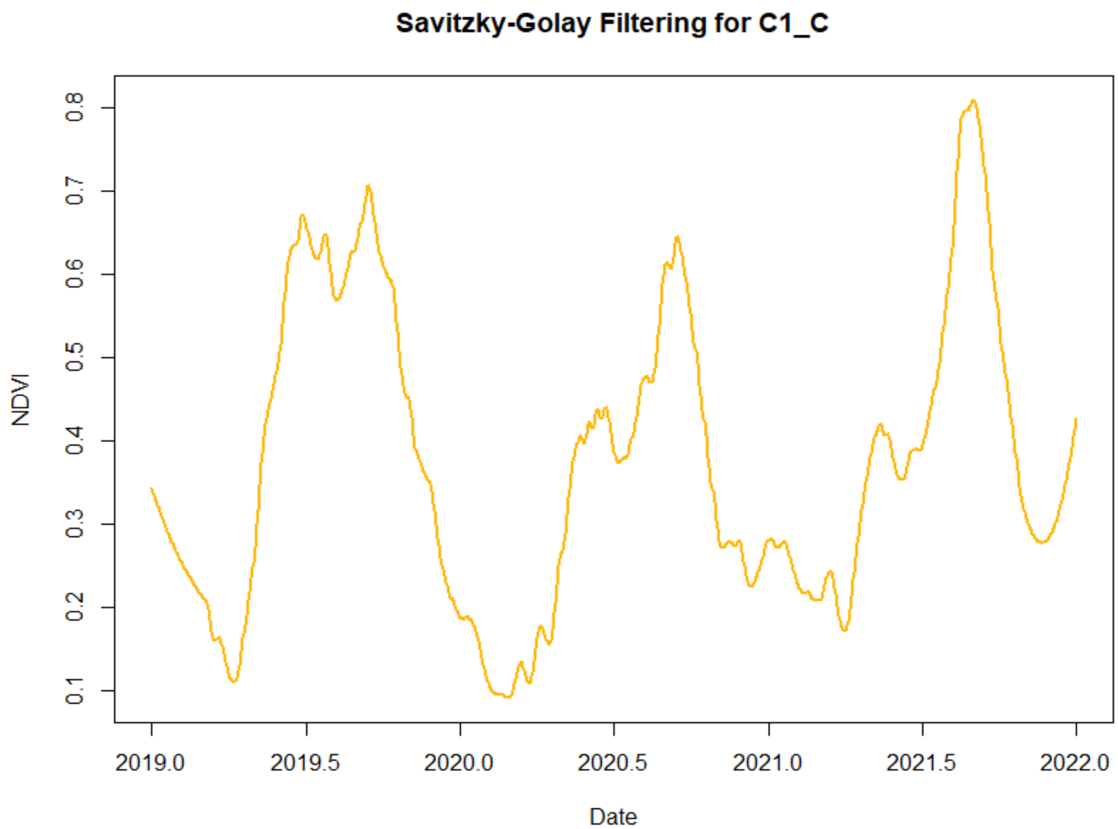

Figure 4: Savitzky-Golay Smoothing

After smoothing, Bfast model was used which helps to detect the breakpoints. We check through several parameters of the model so it could be able to detect the breakpoints.

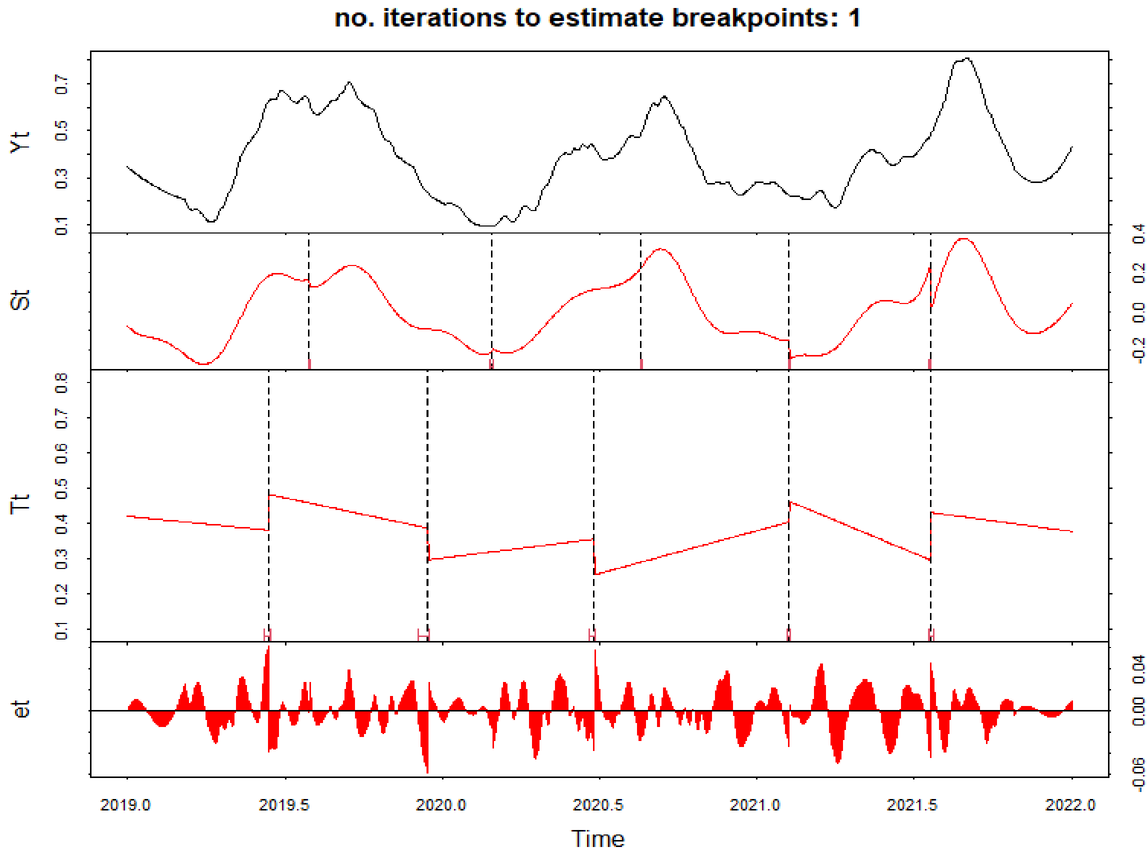

Figure 5: Plot from Bfast algorithm. The first pane shows the observed data. The second pane shows a seasonal pattern with breakpoints. The third pane shows a trend line which has not changed across the time span. And the last pane is for error terms while fitting the model.

If we watch carefully at the figure, we got five breakpoints from the given data of three years. Considering the period of leaf up and leaf fall in the trees, there should be six breakpoints for seasonality. Further, this model detected the start of the season at peak of the first year. As we know, this model is able to detect the change in gradient in a non-linear curve but it fails to meet our purpose. The clouds reduced the signal during summer which results in all the depression at that time which is misinterpreted as a change point by the algorithm. Even though smoothing techniques were used to eliminate these errors, that was not enough for the dataset we used.

For the comparison of breakpoints, it was not possible with this model. For each plot parameter was required to be tweaked so it could detect the breakpoints. Using different parameters sets different thresholds for the breakpoints which detect different numbers of breakpoints and unusual time of the year.

Further, we tried using Rbeast algorithm which is considered superior to Bfast for breakpoint detection in time series datasets.

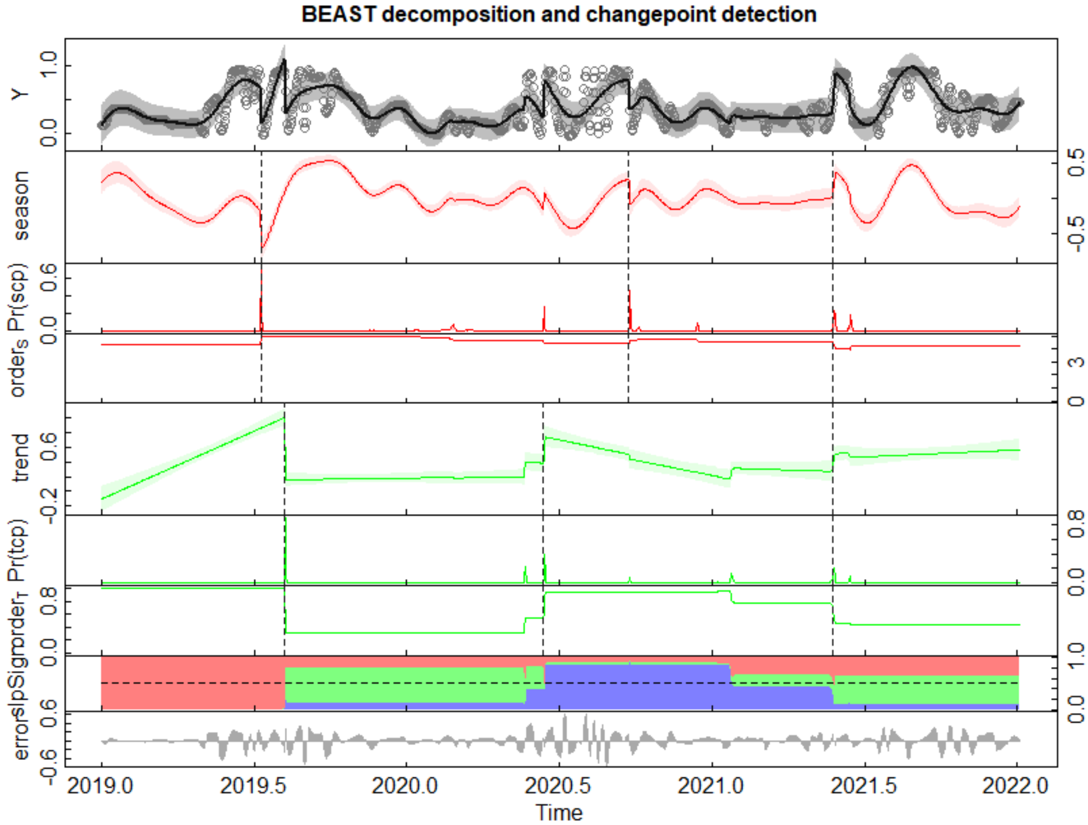

Figure 6: Plot from Rbeast algorithm.

From the second pane of figure 6, we can see in seasonal patterns, this algorithm was also able to detect only three breakpoints. Moreover, they seem to be affected by the cloud component during the summer. Thus, the study of phenology using remote sensing data for small patch of land where area comes under a single pixel; which will be easily affected by the cloud is not the best choice for this kind of analysis.

Discussion

We were not able to detect the breakpoints in the seasonal cycle with consistency using Bfast algorithm. The results do not seem convincing for detecting the start and end of the growing season in the northern temperate hardwood forest of New Hampshire. Phenological study are always a challenge for location which are supposed to have more cloudy days. Optical datasets are easily affected by weather conditions thus, these data may not reflect the trend and condition of region of interest across the timespan. Considering this, Synthetic Aperture Radar (SAR) data can be better utilized for phenological study because these data are not affected by clouds. The utilisation of SAR data can give better understanding of phenological conditions.

Verbesselt et. al, 2010 suggested taking care of noise which can influence the result. Masking does not fully eliminate the cloud component resulting in inconsistency in datasets. Although Bfast has two ways of addressing outliers, still they are affected by noise. Most importantly, signal-to-noise ratio is vital; the amplitude of signal fluctuates and makes the phenological study possible (Verbesselt et al., 2010). Similarly, Calders et al., (2015) used terrestrial LiDAR data to study phenology using Plant Area Index (PAI) and got better prediction of leaf up and leaf fall using a sigmoidal model. PAI is the sum of both the Leaf Area Index and other vegetative parts like woods, fruits, etc which can reflect back the LiDAR signal. The use of LiDAR data had better prediction than MODIS data because of the lower standard deviation (Calders et al. 2015).

Zhao et al. (2019) introduced Bayesian Estimator of Abrupt change, Seasonal change, and Trend (BEAST) model for detecting changepoints, seasonality, and trends. Although, it helps to alleviate model misspecification, chances of over-fitting and algorithmic uncertainty, it was not helpful for phenological study using Sentinel data.

Although, ground data are reliable for phenological study but aligning the remote sensing data with field data will help to cover larger spatial extent. It will be super helpful for not only observing the effect of climate change but also quantify the effect of management and fertilization with precise information.

Conclusions

This study aimed to relate remote sensing dataset with phenological cycle of trees (leaf on and leaf fall) of the trees in northern hardwood forest. Cloud masked data acquired from Sentinel-2A with higher spatial and temporal resolution was used. The timeseries data was interpolated to avoid any gap and smoothened to minimize noise. Finally, two algorithm Bfast and Rbeast were used to detect the breakpoints in timeseries data analysis. Although, they detected the breakpoints, they are not the relevant time for Start of Season (SOS) or End of Season (EOS). Finally, not totally relying on optical sensors but fusing other data source like SAR or LiDAR can be promising approach for this kind of study.

References

Bailey A, Hornbeck JW, Campbell JL, Eagar C. 2003. Hydrometeorological database for Hubbard Brook Experimental Forest: 1955–2000. Volume 305. Delaware: USDA Forest Service, Northeastern Research Station.

Calders, K., Schenkels, T., Bartholomeus, H., Armston, J., Verbesselt, J., & Herold, M. (2015). Monitoring spring phenology with high temporal resolution terrestrial LiDAR measurements. Agricultural and Forest Meteorology, 203, 158-168.

Copernicus Sentinel data, 2019- 2020. Retrieved from Google Earth Engine 08 October 2022, processed by ESA.

Copernicus, 2019. contains modified Copernicus Sentinel data. Retrieved from Google Earth Engine 6 October 2022, processed by ESA.

Gungula, D.T., Kling, J.G. and Togun, A.O., 2003. CERES-Maize predictions of maize phenology under nitrogen-stressed conditions in Nigeria. Agronomy Journal, 95(4), pp.892-899.

Li, H., Jia, M., Zhang, R., Ren, Y. and Wen, X., 2019. Incorporating the plant phenological trajectory into mangrove species mapping with dense time-series Sentinel-2 imagery and the Google Earth Engine platform. Remote Sensing, 11(21), p.2479.

Madonsela, S., Cho, M.A., Mathieu, R., Mutanga, O., Ramoelo, A., Kaszta, Ż., Van De Kerchove, R. and Wolff, E., 2017. Multi-phenology WorldView-2 imagery improves remote sensing of savannah tree species. International journal of applied earth observation and geoinformation, 58, pp.65-73.

Moyo, H., Scholes, M.C. and Twine, W., 2015. Effects of water and nutrient additions on the timing and duration of phenological stages of resprouting Terminalia sericea. South African Journal of Botany, 96, pp.85-90.

Press, W. H., & Teukolsky, S. A. (1990). Savitzky-Golay smoothing filters. Computers in Physics, 4(6), 669-672.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reed, B.C., Schwartz, M.D. and Xiao, X., 2009. Remote sensing phenology. In Phenology of ecosystem processes (pp. 231-246). Springer, New York, NY.

Sentinel, E.S.A., 1. Missions Sentinel Online. https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Stineman, R. W. (1980). A consistently well-behaved method of interpolation. Creative Computing, 6(7), 54-57.

Studer, S., Stöckli, R., Appenzeller, C. and Vidale, P.L., 2007. A comparative study of satellite and ground-based phenology. International Journal of Biometeorology, 51(5), pp.405

Verbesselt J, Hyndman R, Zeileis A, Culvenor D (2010). "Phenological Change Detection while Accounting for Abrupt and Gradual Trends in Satellite Image Time Series." Remote Sensing of Environment, 114(12), 2970–2980. doi: 10.1016/j.rse.2010.08.003.

Xiao, X., Zhang, J., Yan, H., Wu, W. and Biradar, C., 2009. Land surface phenology. In Phenology of ecosystem processes (pp. 247-270). Springer, New York, NY.

Younes, N., Joyce, K.E. and Maier, S.W., 2021. All models of satellite-derived phenology are wrong, but some are useful: A case study from northern Australia. International Journal of Applied Earth Observation and Geoinformation, 97, p.102285.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C., Gao, F., Reed, B.C. and Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote sensing of environment, 84(3), pp.471-475.

Zhao, K., Wulder, M. A., Hu, T., Bright, R., Wu, Q., Qin, H., ... & Brown, M. (2019). Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote sensing of Environment*, *232*, 111181.

Zhu, L., Suomalainen, J., Liu, J., Hyyppä, J., Kaartinen, H. and Haggren, H., 2018. A review: Remote sensing sensors. Multi-purposeful application of geospatial data, pp.19-28.