

NOISE REDUCTION OF TIME-SERIES SATELLITE DATA USING VARIOUS DE- NOISING ALGORITHMS

Sharmeen Khan¹, M. Afshar Alam², N.R. Shankar Ram³, Khushboo Mirza³, V.M. Chowdary³
E-Mail Id: ¹sharmeen1812@gmail.com, ²alam@jamiyahamdard.ac.in, ³shankarramnr@gmail.com,
³khushboo_m@nrsc.gov.in, ³chowdary_isro@yahoo.com

^{1,2}SEST, Jamia Hamdard, New Delhi, India

³Regional Remote Sensing Centre (RRSC)–North, National Remote Sensing Centre (NRSC), New Delhi, India

Abstract—Remotely sensed time series data is increasingly gaining attention in research due to its wide coverage and high temporal resolution. The vegetation indices (NDVI and EVI) products are mainly obtained from medium resolution satellite sensors such as NOAA/AVHRR, SPOT/VGT, Landsat TM, etc. These products are highly used in various remote sensing applications for detecting long term LULC changes, phenological crop identification, environmental dynamics and climate change, modelling terrestrial ecosystems on global, continental and regional levels. Although the time series datasets of vegetation indices obtained from different satellite sensors have been proved to manage the global environmental change, residual noise due to poor atmospheric conditions and instrumental errors generate unfavorable/erroneous results which impedes further analysis. This study focuses on reducing noise from the daily Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI datasets for the period 2016-2019 using three different de-noising techniques namely Fast Fourier Transform (FFT), Moving Average filter (MA), Savitzky-Golay technique (SG) for noise reduction and reconstruction of time series MODIS NDVI data. The evaluation of the de-noising algorithms was first carried out on synthetically generated signals (sine, square and triangular) to which Gaussian noise with multiple variance levels was introduced. Since smoothing temporal data with ideal noise free data is not possible, therefore, a two-step validation process is implemented which first measures the performance of the techniques on simulated signals using Mean Square Error (MSE) and Peak signal-to-noise ratio (PSNR) and then evaluated the de-noising algorithms both spatially and temporally. For temporal evaluation, NDVI datasets for five land cover classes namely forest, built-up, agriculture, water-body and wasteland from four different locations were selected and reconstructed using the three de-noising algorithms. The filtering techniques were able to reduce the noise from the temporal profiles of all the land cover types by removing the sudden peaks and drops from the NDVI time-series. Performance of temporal analysis was further validated using statistical parameters such as RMSE, PSNR and Pearson correlation coefficient. Hence, the de-noising algorithms were able to reconstruct the NDVI data while preserving the spatial and temporal pattern for better analysis and interpretation.

Keywords: De-Noising algorithms, FFT filter, Moving Average filter, Savitzky-Golay, Time Series.

1. INTRODUCTION

Remote sensing is a one of the best tool for monitoring the earth's resources and environment. Being cost effective, the satellite remote sensing data provides essential information that helps in monitoring and modelling various applications such as image fusion, change detection and land cover classification. Imaging sensor MODIS (Moderate Resolution Imaging Spectroradiometer) aboards the Terra and Aqua satellites, launched in 1999 and 2002 respectively, provides low spatial resolution but high temporal resolution data which is useful to track changes in the landscape over time due to its wide swath area and high revisit capability [1]. Time series data is concerned with applications such as analysis of long term land use/land cover (LULC) dynamics, variability in phenological parameters, climate change, modelling terrestrial ecosystems [2] [3] [4] [5] [6] [1]. One of the major applications of time series data is to monitor the vegetation growth at a regional level by means of such temporal data with vegetation indices.

The time-series datasets computed for vegetation indices like the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI) have been derived from the satellite sensors with medium resolution such as NOAA/AVHRR (Advanced very high-resolution radiometer), SPOT/VEGETATION, TERRA or AQUA/MODIS, ENVISAT/MERIS and LANDSAT TM (Thematic Mapper). NDVI is computed using the reflectance in the red and near-infrared wavelengths and is an indicator of photosynthetically active biomass. The variations in NDVI time series indicate the changes in vegetation health [7]. In order to prepare time-series of NDVI, data has to be passed through several steps of processing. One such important step is noise reduction. Remote sensing datasets frequently suffer from inherent noise caused by cloud cover effects, atmospheric variability and sensor and viewing geometry effects [8] [9] [10] [1] [11] [12] [13]. Although the time-series NDVI data is obtained using Maximum Value Composite (MVC) method, it still includes some residual noise which affects the accuracy of the data [14]. These disturbances degrade the data quality and affects the monitoring of land cover and terrestrial ecosystems. Therefore, it is necessary to remove such uncertainties from the temporal datasets and obtain high-quality [15], filtered data to carry out further analysis and interpretation.

For this purpose, a number of de-noising techniques have been developed in the last two decades to reconstruct high-quality NDVI time-series. These techniques can be broadly classified into five types: (1) Fourier-based fitting methods, including the Fast Fourier Transform (FFT) and the Harmonic analysis of time-series (HANTS) [16] [17]; (2) Threshold-based methods, such as the best index slope extraction technique (BISE) and the modified best index slope extraction technique (MBISE) [18] [19]; (3) Function fitting methods, such as double logistic function fitting (DL) [2] and the Asymmetric Gaussian function fitting (AG) [20]; (4) Filter-based methods, including the Savitzky-Golay [4], the mean value iteration filter (MVI) [21], the ARMD3-ARMA5 filter [22] and phenology-preserving filtering (PP) [23]; and (5) other methods such as wavelet transform (WT) [1], data assimilation (DA) [10], the Whittaker smoother (WS) [24] and iterative interpolation for data reconstruction (IDR) [25].

The techniques mentioned above have both advantages and limitations over each other [12]. The parameters used in some of the techniques are required to be adjusted for setting threshold values, changing the size of the moving window or temporal neighborhood and deciding the number of harmonics. Jonsson & Eklundh [20] showed that AG is superior to BISE and FFT but failed to identify the maxima/minima to which local functions can be fitted especially for noisy data and for areas where there is no seasonality [4]. Chen et al. [4] highlighted that SG, BISE and FFT are effective techniques for reconstructing NDVI time series and SG is the best of the three. Ma and Veroustraete [21] explained MVI performed better than MBISE and FFT. However, Beck et al. [2] demonstrated that new DL is a better technique than FFT and AG because it can handle outliers effectively and estimate phenological parameters. Lu et al. [1] developed the WT technique and compared it with BISE, FFT and SG and found that WT performed best in removing noise. Hird and McDermid [12] empirically evaluated the AG, DL, SG, 4253H, MVI, ARMD3-ARMA5 techniques demonstrated the general superiority of function-fitting methods over four other alternative filters. Julien and Sobrino [25] showed that IDR gives best reconstruction for moist land cover classes compared to HANTS and DL. Atkinson et al. [26] compared FFT, AG, DL and WS and found that the de-noising techniques should be adjusted according to the annual growing seasons to produce reliable fits.

This study is carried out with the specific objectives of 1) Reconstruction and analysis of high-quality NDVI time series datasets from raw daily MODIS datasets, utilizing de-noising techniques, viz. Fast Fourier Transform (FFT), Savitzky-Golay (SG) and Moving-Average filter (MA). 2) Statistical performance evaluation of the filters using Root Mean Square Error (RMSE), Peak signal-to-noise ratio (PSNR) and Pearson correlation coefficient(r), to identify the best filtering technique.

2. STUDY AREA AND DATASETS

2.1 Study Area

Uttar Pradesh located in northern India, cover the total geographical area of 240,928 km². It is situated between 23.52° N - 30.24° N latitude and 77.05°E - 84.38°E longitude. It is bordered by nine states and is internationally bordered by Nepal. Physiographically, the state is divided into two regions namely, the Gangetic Plain, within which the vast majority of Uttar Pradesh lies and, the southern uplands forming part of the highly dissected and rugged Vindhya Range, which gradually rises toward the southeast. It is characterized by a hard rock strata and a varied topography of hills, plains, valleys and plateaus. The elevation of the Gangetic plain varies from about 1,000 feet in the northwest to about 190 feet in the extreme east, while, the elevation of the plateau region rarely exceeds 1,000 feet. The alluvial or Gangetic plain is divided into three sub-regions. The eastern tract is subjected to periodical floods and droughts. The central and the western regions, are fairly arid and have well-developed irrigation system. However, they suffer from water-logging and large-scale user tracts. A deep layer of alluvium which is spread by the number of tributaries of the Ganges system has covered most of the state. The valley areas have fertile and rich soil. The terraced hill slopes have intensive cultivation, while irrigation facilities are deficient. Tropical and dry deciduous forests are found in all parts of the plains. The climate of Uttar Pradesh is humid subtropical with dry winters. The Gangetic plain varies from semiarid to sub-humid whereas climate of the Vindhya Range and plateau is subtropical. Annual rainfall in the state ranges from (1,000–2,000 mm) – (600–1,000 mm) in the west. Floods, fatalities and heavy damage to crops and property, particularly in the eastern part of the state is caused by heavy rainfall during the southwest monsoon while drought conditions likely to occur due to failure of monsoons. Population of Uttar Pradesh is 199.58 million (Census 2011) making it as most populous state in the country. A number of rivers originating from the Himalayas and Vindhya Range provide a well drainage system for the state. Providing perennial source of irrigation, the Ganges basin and its tributaries, with its fertile soil is instrumental to the agricultural economies of the state. Chief crops cultivated in the area include rice, sugarcane, lentils, oil seeds, potatoes, and wheat. The river mainly Ganges is vulnerable to human induced pollution due to the millions of people residing in the basins in the eastern side.

2.2 Satellite Datasets Used

Daily MODIS data products (MOD09GQ), location tile h25v06, for the period 2016-2019 were downloaded from the Atmosphere Archive & Distribution System (LAADS) Distributed Active Archive Center (DAAC) for this study. The data set covered the most of the part of north India and some part of Nepal. The MOD09GQ product has spatial resolution of 250m and is of processing level L2G. It provides daily surface reflectance, corrected for atmospheric conditions such as gasses, aerosols, and Rayleigh scattering, for band 1 (620-670 nm) and band 2 (841-876 nm) on a sinusoidal gridded projection. Along with surface reflectance bands, there is the Quality Assurance (QA) layer and

five other observation layers. For this work, points/coordinates from the central region of Uttar Pradesh were selected to perform the temporal analysis.

3. METHODOLOGY

3.1 Generation of time-series NDVI data at daily scale

The sinusoidal gridded MOD09GQ Version 6 product provides an estimate of the surface spectral reflectance of Terra Moderate Resolution Imaging Spectroradiometer (MODIS) 250 meter bands 1 and 2, corrected for atmospheric conditions such as gasses, aerosols, and Rayleigh scattering. Along with the 250 m spatial resolution, surface reflectance bands are the Quality Assurance (QA) layer and five observation layers. For this analysis, the daily MODIS images are downloaded for the period 2015-2019. The $\rho_{\text{refl_b01_1}}$ (620-670 nm) and $\rho_{\text{refl_b02_2}}$ (841-876 nm) corresponding to red and (Near Infra-Red) NIR band respectively, are extracted in order to calculate the Normalized Difference Vegetation Index (NDVI) and then the NDVI time series is generated for the entire dataset for further analysis. NDVI is sensitive to changes in both temperature and precipitation [5]. Because the NDVI correlates directly with vegetation productivity and with absorbed photosynthetically active radiation (APAR), there are numerous possible applications of this index for ecological purposes.

The NDVI is a normalized transform of the NIR to red reflectance ratio, $\rho_{\text{NIR}}/\rho_{\text{red}}$, designed to standardize values between -1 and +1. It is commonly expressed as:

$$\text{NDVI} = \frac{\rho_{\text{nir}} - \rho_{\text{red}}}{\rho_{\text{nir}} + \rho_{\text{red}}}$$

Where, ρ_{nir} = wavelength of near-infrared band and ρ_{red} = wavelength of red band

NDVI is especially useful for vegetation monitoring because it can compensate for changing illumination conditions, surface slope, and viewing angle. It does tend to saturate over dense vegetation and is sensitive to underlying soil color.

3.2 Reduction of noise in NDVI time-series datasets

The undesirable noise and disturbances in the remotely sensed time series data has to be minimized in order to preserve the original trends. For this purpose, a number of time-series filtering methods have been developed and implemented in the past. Noise reduction in this study is incorporated using three techniques, such as Savitzky-Golay filter (SG), Moving Average filter (MA) and Fast Fourier transform filter (FFT).

These algorithms are tested on pixels of different land cover types, which have been selected from the LULC Thematic maps, which are freely available on Bhuvan (a utility software from ISRO) as a web service. Test locations and their corresponding coordinates for each of the five basic land-cover types viz. forest, built-up, agriculture, water-body and wasteland, have been extracted as shown in table 1.

The temporal profiles for all these 20 locations were generated and were further subjected to de-noising algorithms which have been outlined as follows:

3.2.1 Savitzky-Golay Filter

Savitzky-Golay filter is a digital smoothing filter which uses least-squares polynomial fitting for reduction of noise from NDVI time-series [27]. Savitzky-Golay (SG) is particular type of low-pass filter replaces each value of the series with a new value obtained from polynomial fit to $2n+1$ neighboring points including the point to be smoothed, with n being equal to greater than the order of the polynomial. General expression for the filter can be given as:

$$g_i = \sum_{i=-m}^{+m} c_i f_{i+1} / N$$

Where, f_i is the original time series, g_i is the smoothed value which is a linear combination of c_i and f_i , the filter coefficient c_i is given as the polynomial of certain degree maintains higher values and N is the convoluting integer and is equal to the smoothing window size consisting of $(2m+1)$ points.

The algorithm requires the optimal selection of the window size and the order of polynomial based on NDVI observations. Larger the window size and lower the polynomial order, more smoothing occurs. The polynomial is designed to preserve dataset features such as peak height and width and reduce the bias introduced by the filter [4]. In this method, fifth order polynomial coefficient for 55 window length was computed in order to remove peaks and spikes.

3.2.2 Moving Average filter

Moving Average is a naïve and effective technique in time-series analysis or forecasting. It replaces a data point with the average or weighted form of its neighbors, creating a series of the averages of different subsets of the full dataset. The window width is slid along the time series to calculate the average values in the new series [27]. General form to denote moving average is:

$$y(t) = \sum_{i=-k}^{+k} W_i x(t + 1)$$

Where $2k+1$ is the width of the moving window, W_i is the weights and $x(t)$ being the data values at time t . Moving average is closely related to Savitzky-Golay filter and FIR (Finite impulse response) filter. Since, it is a type of convolution so it can also be viewed as a low pass filter. This study used window width of length 8 to calculate weights and obtain the filtered time series.

3.2.3 Fast Fourier Transform Filter

Fourier Transform (FT) is probably the most widely used signal analysis method which decomposes the signal present in the complex space into the sum of sinusoidal functions [28]. The Fast Fourier Transform (FFT) provides an efficient algorithm for converting the waveform from the time domain into the frequency domain and there is one coefficient per frequency present in the signal.

The Fourier transform can be defined as:

$$x[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{j2\pi kn}{N}}$$

The de-noising of the time-series using Fast Fourier transform is done as follows: First FFT filter from a python module is applied to the signal and frequencies associated with each coefficient are computed. Then, the coefficients which have a low frequencies (in absolute) are kept and inverse of FFT is computed which generated the noise-free NDVI time-series. This work incorporated three parameters namely filter order, cutoff frequency and sampling frequency in order to filter the time-series using FFT. Hence, filter order of size 2, cutoff frequency of 50 and sampling frequency of 2000 was taken to get the smoothed time-series.

3.3 Generation of Synthetic Signals

The evaluation of the de-noising algorithms was first carried out on the synthetically generated signals (sine, square, sawtooth/triangular) to which Gaussian noise at a different variance level (0.01, 0.02, 0.03) was introduced. These signals were reconstructed again using the de-noising algorithms. The varying parameters used in the algorithms are tuned until satisfactory results were obtained as shown in Table 2. The performance and reconstruction results of the filters are directly affected by the parameters used. Hence, these identified parameters have been used for the temporal analysis of the NDVI time-series datasets. The efficiency of the filtering techniques was evaluated using the statistical parameters namely Mean Square Error (MSE) and Peak signal-to-noise ratio (PSNR) (Table 2). Thus, the evaluation of the de-noising algorithms can be considered as a two-step validation process; first, checking their effectiveness on the simulated signals and then carrying out the analysis on the temporal NDVI datasets.

3.4 Performance Evaluation of Noise Reduction Algorithms

The performance of noise reduction is greatly affected by the varying parameters used in the three techniques and the smoothing parameters are tuned until satisfactory results are obtained. To further analyze the effectiveness of each filtering technique, the results are validated using three statistical parameters such as Root Mean Square Error (RMSE), Peak signal-to-noise ratio (PSNR) and Pearson Correlation Coefficient (r).

3.4.1. Mean Square Error (MSE)

MSE value denotes the average difference of the pixels all over the image. A higher value of MSE means there is a large difference between the original image and processed image as data values are widely dispersed around its central mean [29].

$$MSE = \frac{1}{mn} \left(\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i, j) - R(i, j))^2 \right)$$

Where I and R can be interpreted as input and reconstructed images respectively. m defines the number of pixel in vertical dimension of images I and R and n defines the number of pixel in horizontal dimension of images I and R .

3.4.2. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) also known as Root Mean Square Deviation (RMSD) is a quadratic scoring rule that measures the average magnitude of the error. It is calculated as the square root of the average of squared differences between predicted and observed values. The errors are squared before they are averaged, therefore, a relatively high weight is given to large errors by RMSE. Therefore, RMSE is more useful when large errors are particularly undesirable [30]. The smaller an RMSE value indicates a better fit.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (p_i - a_i)^2}{n}}$$

Where p_i is the predicted NDVI values and a_i is the NDVI actual observed values and n is the number of samples.

3.4.3. Peak Signal-to-Noise Ratio (PSNR)

PSNR is an expression which gives the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation.

$$\text{PSNR} = 10 \cdot \log_{10} (\text{MAX}_I^2 / \text{MSE}) = 20 \cdot \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right)$$

Where 'MAX_I' is the maximum possible pixel value of the image [31].

The higher value of PSNR indicates the better the quality of the reconstructed image.

3.4.4. Pearson correlation coefficient (r)

The Pearson correlation coefficient (r) is a measure of strength of the statistical relationship or association between the two linear variables. The Pearson correlation coefficient is calculated as the covariance of the two variables divided by the product of the standard deviation of each data sample. It is the normalization of the covariance between the two variables to give an interpretable score [32]. The Pearson correlation coefficient, r, can take a range of values from +1 to -1 with 0 implying no correlation.

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

Where x is the original NDVI values and y is the corrected NDVI values.

Performance analysis of the de-noising algorithms was carried out temporally. Time-series NDVI temporal profile data have inherent noise which is reconstructed using the three de-noising algorithms and then further evaluated with respect to original one using RMSE, PSNR and Pearson coefficient.

4. RESULTS AND DISCUSSION

The de-noising algorithms have been initially applied to the synthetically generated signals (sine, square, sawtooth/triangular) in order to evaluate the performance of the algorithms. Gaussian noise at variance levels of 0.01, 0.02 and 0.03 were introduced into the raw signals and were further reconstructed using the de-noising algorithms. The efficiency of these algorithms in removing noise from the synthetic signals was evaluated using MSE and PSNR as shown in Table 2. From table 2, it can be seen that MA filter is performing better with higher PSNR values for all the signals.

The daily MODIS NDVI data was subjected to the de-noising algorithms Fast Fourier Transform (FFT), Savitzky-Golay (SG) and Moving-Average (MA) for the purpose of noise-reduction from the time-series datasets. These algorithms were tested for five land-cover types namely forest, built-up, agriculture, water-body and wasteland in study area and we proceed to analyze the results. Fig 2 – 6 shows the plots of temporal profiles for each land cover classes for four different locations for the duration of 2016-2019 which are reconstructed using the three de-noising algorithms.

From Fig 2, we identify the existence of a periodic trend in the overall NDVI values in each of the year. The values generally remain low during the winter months and gradually increase to peak value of 0.6 during March and then begin to decrease and reach a minimum of around 0.2 during the months of July and August. The values further increase and reach a peak during the months of September and October and further fall back during the peak winters. The increase in vegetation cover during the post-monsoon period has effectively been recorded in our analysis. The reduction in vegetation due to shedding of leaves during the winter months have also been captured. The rise in vegetation cover post winter during spring and the subsequent reduction during peak dry summers the region experiences have also been effectively portrayed. However, NDVI values of location 4 only increases in the onset of monsoons and then linearly decreases until the monsoon of subsequent year. From Figure 2 we also see the excessive noise present in the original signal which is physically unrealistic, since drastic fluctuations in vegetation cover would be required to mimic the trends seen. The various filters have effectively reduced these variations and have resulted in realistic smoothed trends.

From Fig 3, it is seen that all the built-up locations except location 1 has significant amounts of vegetation present, which is clearly visible from the higher NDVI values. All the locations show a periodic trend where NDVI values decreases periodically in the pre-monsoon season and sometimes in the end of the year i.e. the onset of winter season. The drastic fluctuations in the observed time series values have been clearly reduced when the de-noising algorithms are applied.

Fig 4 examines the performance of the de-noising algorithms on the Agriculture land-cover type. We observe higher NDVI values of 0.6 during the winter and monsoon seasonal periods. Location 1, 2 and 4 show a sudden drop in the NDVI values during the end of the year 2017.

Fig 5 examines the performance of the de-noising algorithms on water-bodies. Location 1 of water body has least amount of vegetation except during the end of 2018 and start of 2019, when NDVI values are varying from low to high. Constant NDVI values are maintained in location 4 throughout the time-series.

NDVI time series of wasteland is presented in Fig 6. We clearly see that location 1 has least the amount of vegetation as visible from the presence of lower NDVI values. Location 4 has sudden increase in NDVI values in winter of 2017, 2018 and 2019, which decreased to 0 during pre-monsoon. Location 1 has vegetation only in end of monsoon, which gradually decreases until onset of monsoon of the subsequent year. Location 3 has multiple sudden drops in NDVI values, whereas location 2 maintained constant NDVI values throughout the time series.

The de-noising algorithms were successful in removing sudden peaks and drops while maintaining the original overall trends of all the NDVI time-series. Table 3 justifies the figures by showing the overall improvement in the noise

reduction by giving lower RMSE values and higher PSNR and Pearson values. For forest, SG has performed better by giving lower RMSE value 0.198 and higher PSNR value 61.639 while higher Pearson value is obtained from MA. In case of built-up, MA has given lower RMSE value 0.108 and higher PSNR and Pearson value 67.911 and 0.563 respectively. For agriculture land-cover type, FFT has given lower RMSE value whereas higher PSNR 64.487 and higher Pearson 0.704 is given by MA. MA has also given lower RMSE 0.094 and higher PSNR 69.475 and Pearson 0.68 for water-body. For wasteland, FFT has given better RMSE value which is closely related to that of SG 0.1992 and again higher PSNR 0.685 and Pearson 67.35 is given by MA. Hence, it can be observed that for each land-cover class, MA has given the highest PSNR and Pearson values. The consistently lower values of RMSE and higher values of PSNR and Pearson coefficient throughout the evaluation process for all the averaged locations of land cover types clearly shows that MA has reduced the noise significantly better than the other algorithms. Hence, for temporal analysis, MA is the best de-noising algorithm for giving better quality NDVI time-series than other two techniques.

The performance of the de-noising algorithms was further evaluated on the averaged locations of the temporal profiles was evaluated by averaging the mean and standard deviation of each land-cover type for all the four locations as shown in Table 4. The significantly lower values of the standard deviation for each land-cover class, it can be said that the de-noising algorithms were successful in removing noise from the time-series to some extent. It is seen that for forest, FFT filter gives the lowest standard deviation 0.1380 with the highest mean value 0.385, whereas for built-up, SG filter is giving the lowest standard deviation 0.0643 and highest mean 0.1943. For agriculture, water-body and wasteland, again, the FFT filter gives the lowest standard deviation 0.1304, 0.0525 and 0.0933 however, highest mean value is given by SG filter 0.3111, 0.0755 and 0.2383 for agriculture, water-body and wasteland. Therefore, it was observed that FFT filter has performed better as compared to other two techniques.

All the NDVI datasets had considerable noise which was reduced using the above de-noising algorithms. All the processing is implemented and run on open-source tools (Python 3.8.2) and coordinates for pixels corresponding to each class was extracted from thematic maps freely available at Bhuvan (Indian web-based utility from ISRO). Figure 2 – 6 illustrates the smoothing performance on different land cover classes namely forest, agriculture, water-body, wasteland and built-up. This clearly shows that the de-noising algorithms are able to preserve the general form of the curves and simultaneously removed the sudden drops caused by atmospheric contamination from the time series. Taking into account the temporal evaluation for each land-cover classes (Table 3), performance of all the three de-noising algorithms clearly affirms that the time-series after noise removal has better quality. However, among all the techniques, MA has given best results temporally with consistently lower RMSE values and higher PSNR and Pearson coefficient values. The performance of algorithms was further validated by doing statistical analysis on the land-cover classes for averaged locations.

CONCLUSION

The time-series satellite data generally suffers from noise which greatly affects the monitoring of land cover and other remote sensing applications. Hence, it is necessary to remove the undesirable noise from the datasets before further analysis and interpretation. The current research focused on obtaining better temporal analysis by de-noising the time-series MODIS datasets using the de-noising algorithms Fast Fourier transform (FFT), Savitzky-Golay filter (SG) and Moving Average filter (MA). The time-series for each land-cover classes was reconstructed and analyzed statistically. All the three de-noising algorithms have performed well in removing considerable amount of noise from the datasets while preserving the temporal and spatial pattern for better analysis. However, MA technique proved to be significantly better than the other two techniques with its lower RMSE values and higher PSNR and Pearson coefficient values. Thus, reconstruction of NDVI data is useful for detecting long term LULC changes, monitoring climate change and vegetation dynamics on a global and regional level.

Table-1 Locations Used from the Study Area

Land-cover class	Label/Location	Coordinates (Lat/Long)
Forest	1	28°31'53.4" N 79°57'56.3" E
	2	28°26'57.8" N 80°00'02.6" E
	3	28°29'51.5" N 80°15'00.6" E
	4	28°31'03.5" N 80°14'33.4" E
Built-up	1	28°37'05.7" N 79°48'27.1" E
	2	28°22'26.0" N 79°26'35.4" E
	3	27°53'05.4" N 79°54'33.5" E
	4	28°30'51.0" N 80°08'48.5" E
Agriculture	1	28°36'27.4" N 79°47'19.8" E
	2	28°06'15.8" N 79°27'17.1" E
	3	27°56'51.2" N 79°54'15.4" E
	4	28°31'37.4" N 80°11'19.9" E
Water-body	1	28°36'55.1" N 79°47'13.0" E

	2	27°05'33.8" N 81°28'58.6" E
	3	26°47'19.3" N 82°00'35.4" E
	4	28°29'59.6" N 80°21'24.7" E
Wasteland	1	28°37'28.1" N 79°46'41.7" E
	2	28°21'39.0" N 79°28'52.5" E
	3	27°59'44.3" N 79°53'38.6" E
	4	28°30'46.4" N 80°21'00.3" E

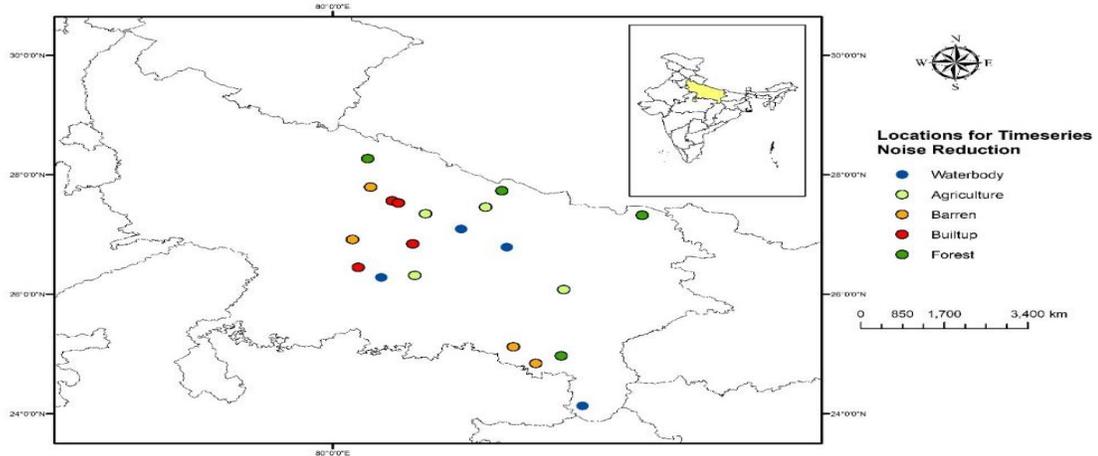
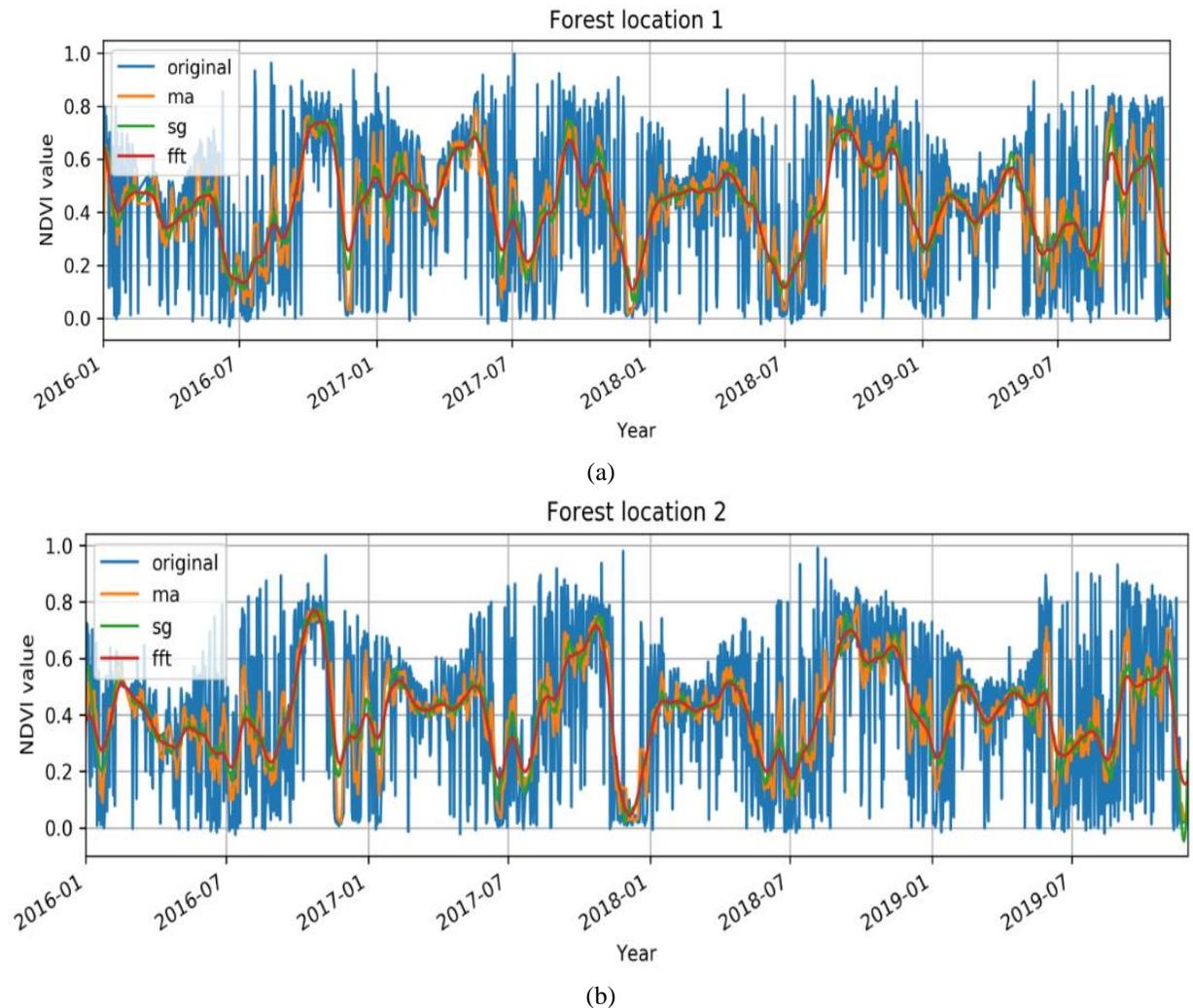
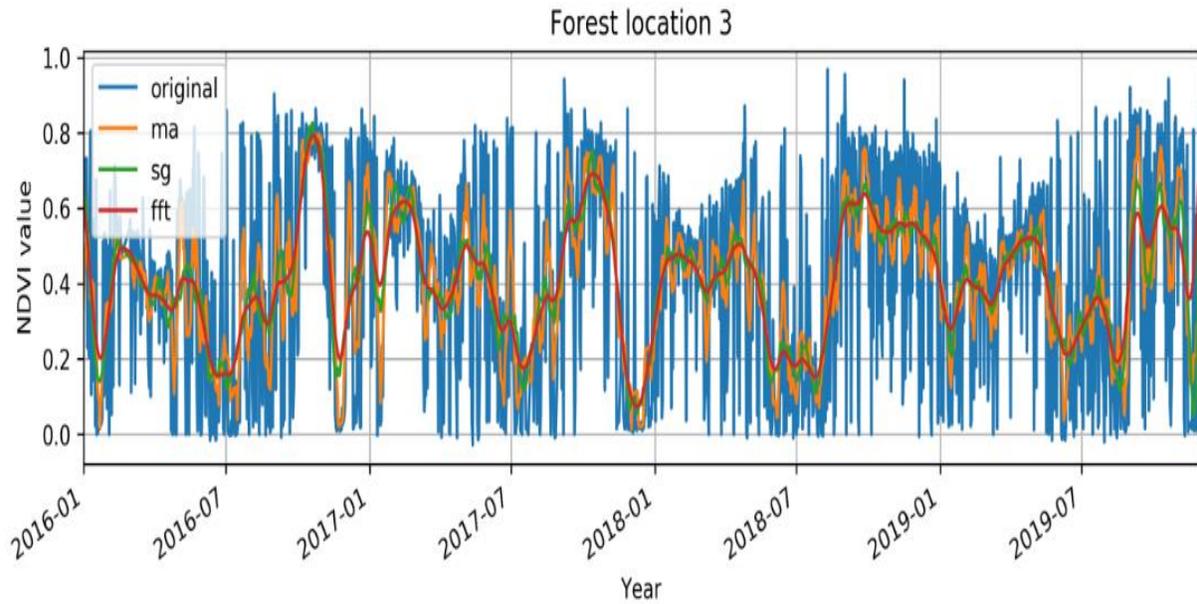
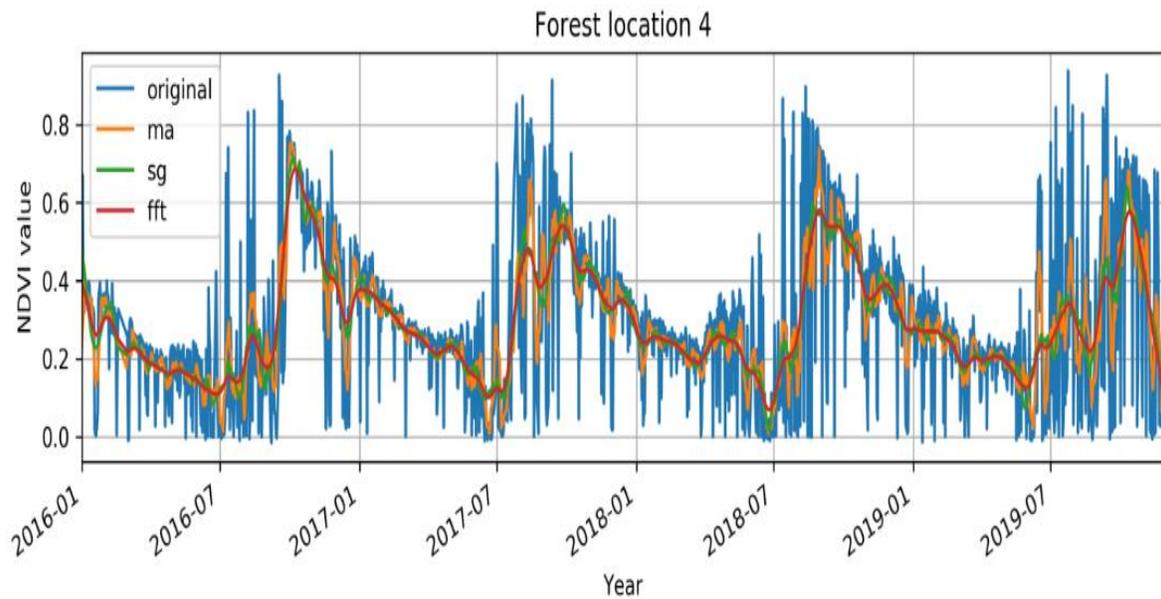


Fig. 1 Map of Uttar Pradesh, Study Area Used For Spatial Analysis and Location of Points Used For Noise Reduction From Time-Series



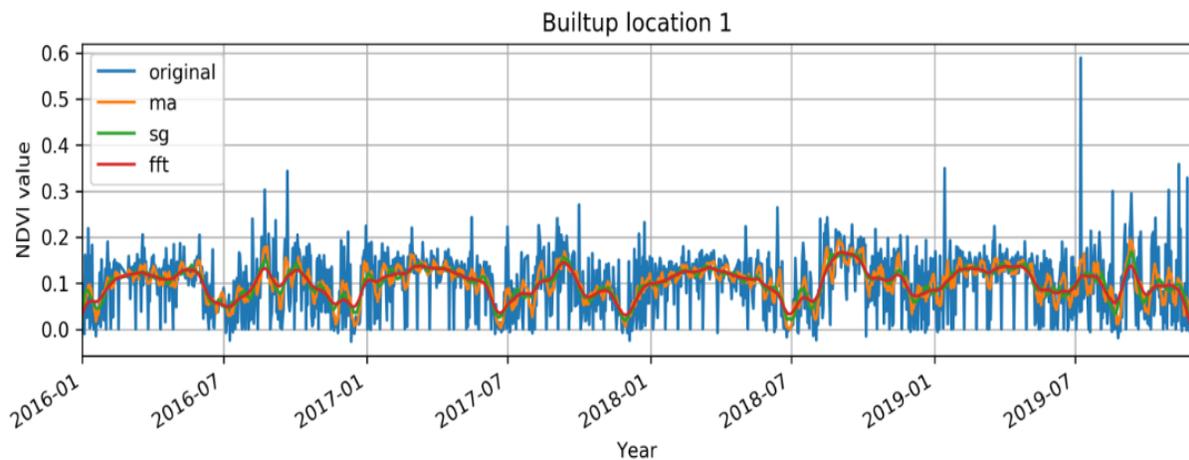


(c)

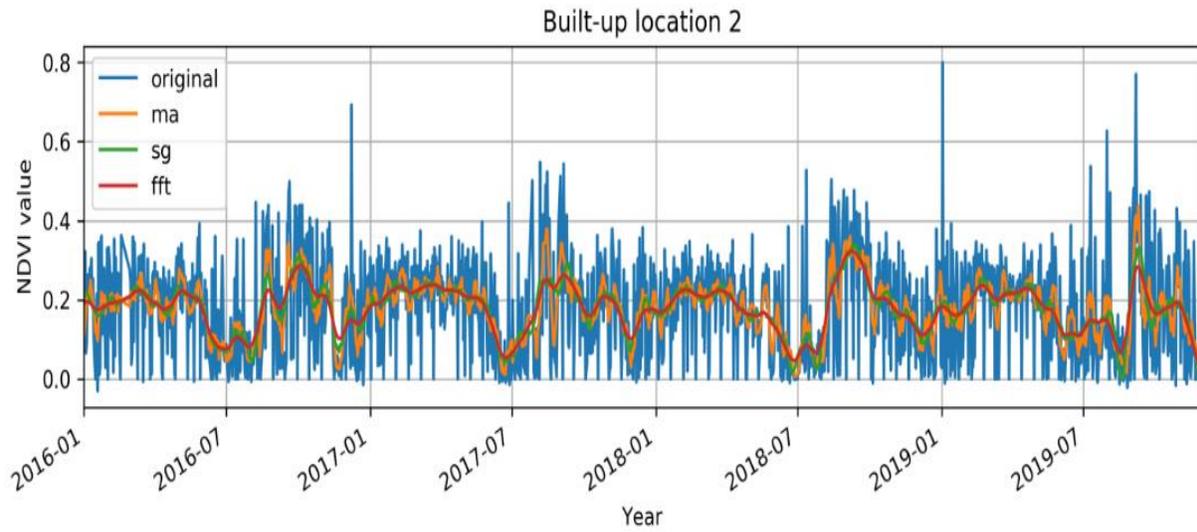


(d)

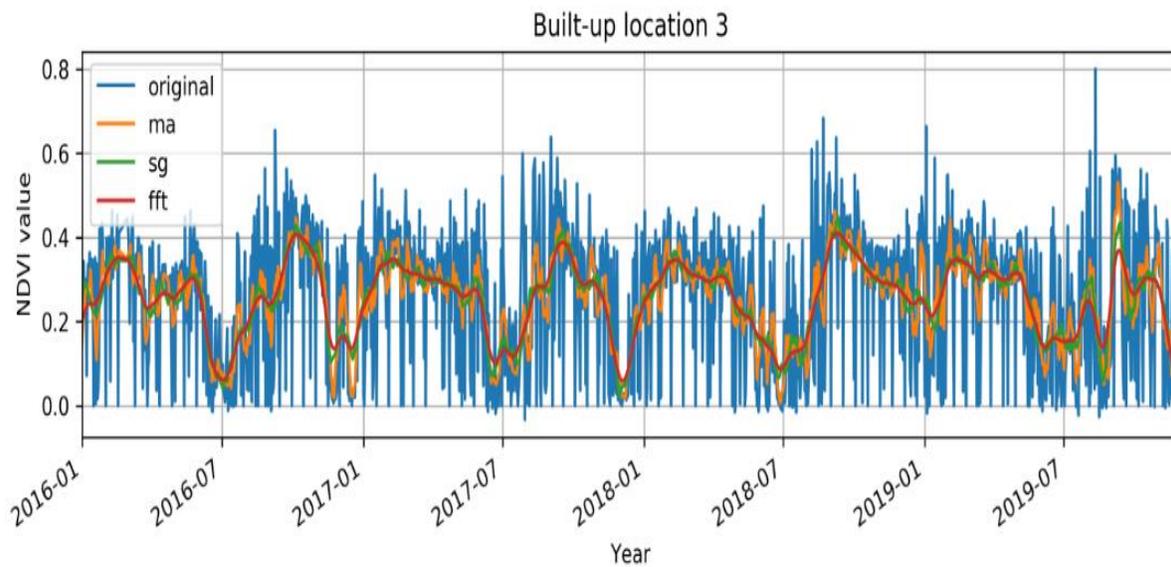
Fig. 2 Temporal NDVI profiles of raw and reconstructed datasets using FFT, MA and SG de-noising algorithms for the period of 2016-19 with subplots a) for location 1, (b) for location 2, (c) for location 3 and (d) for location 4



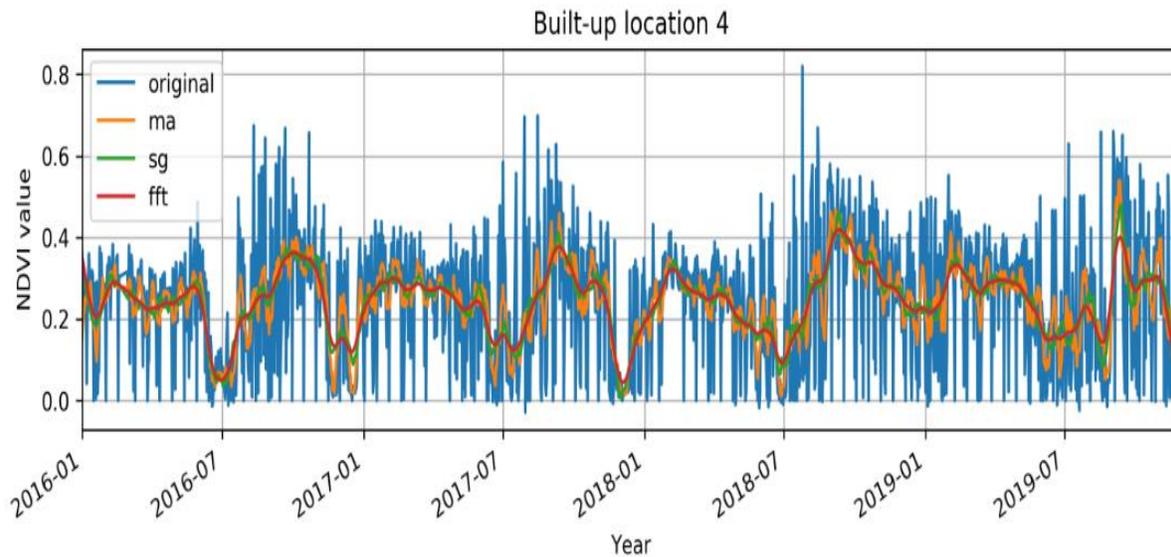
(a)



(b)

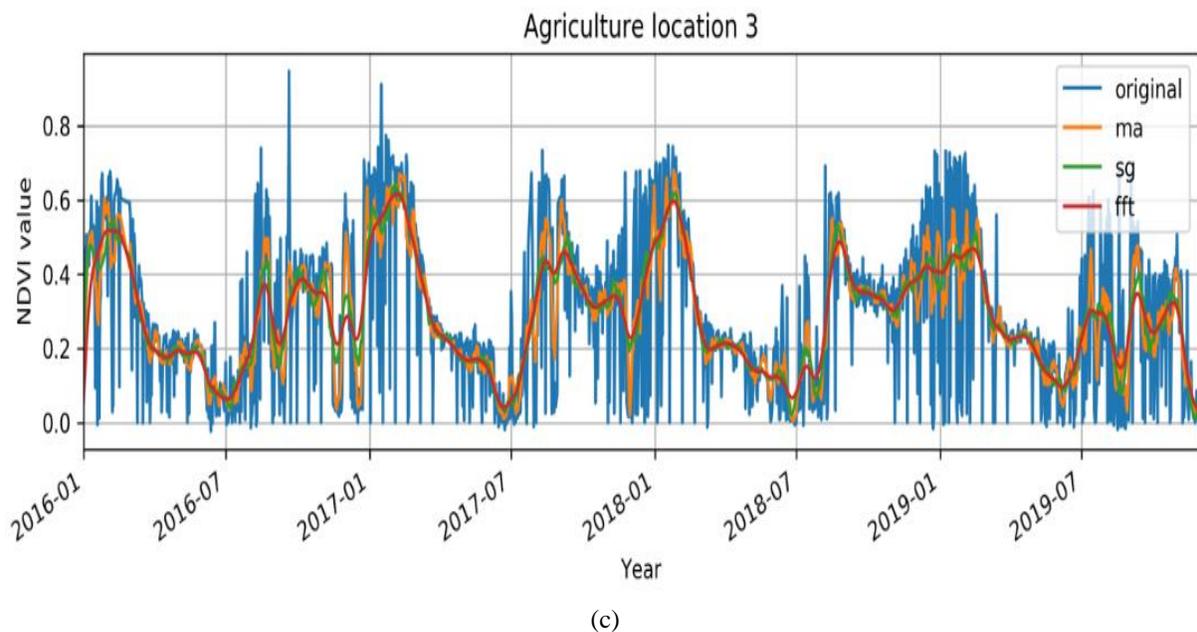
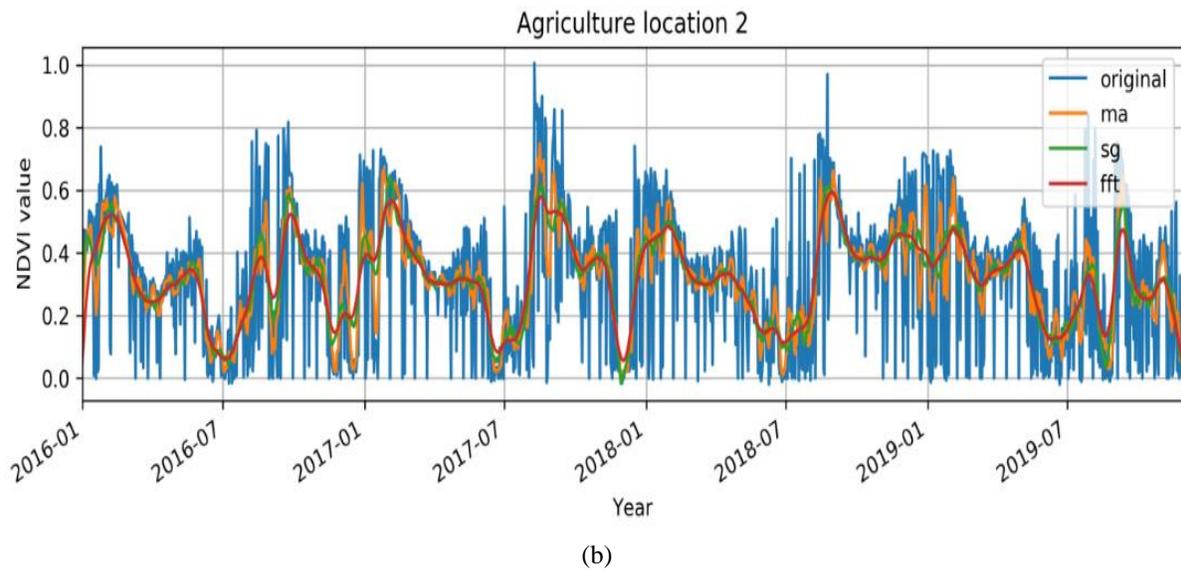
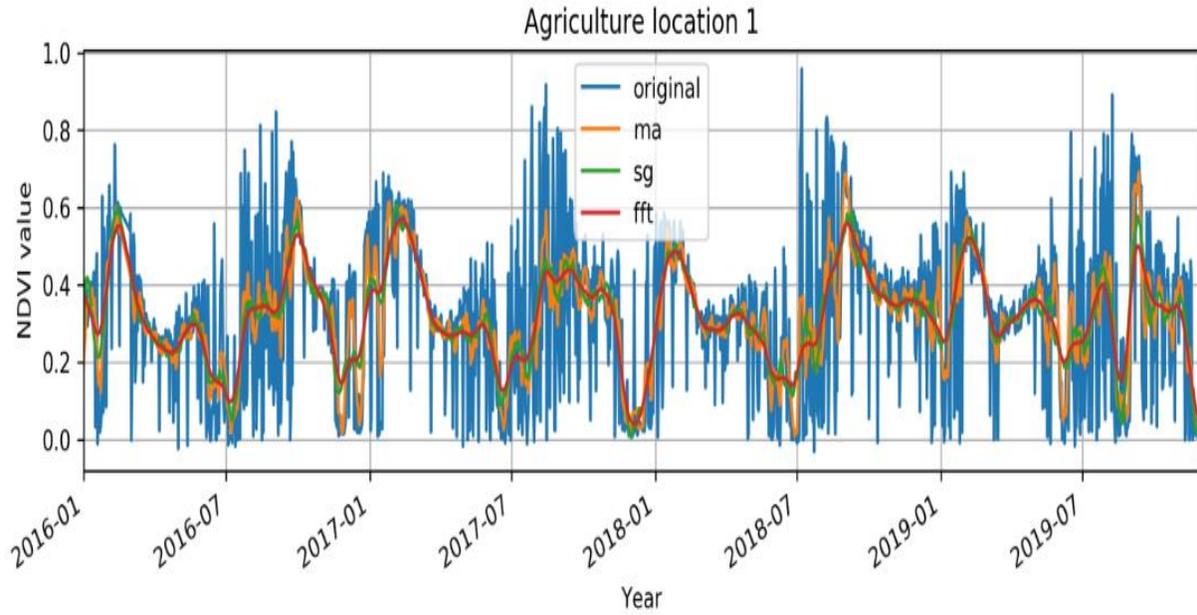


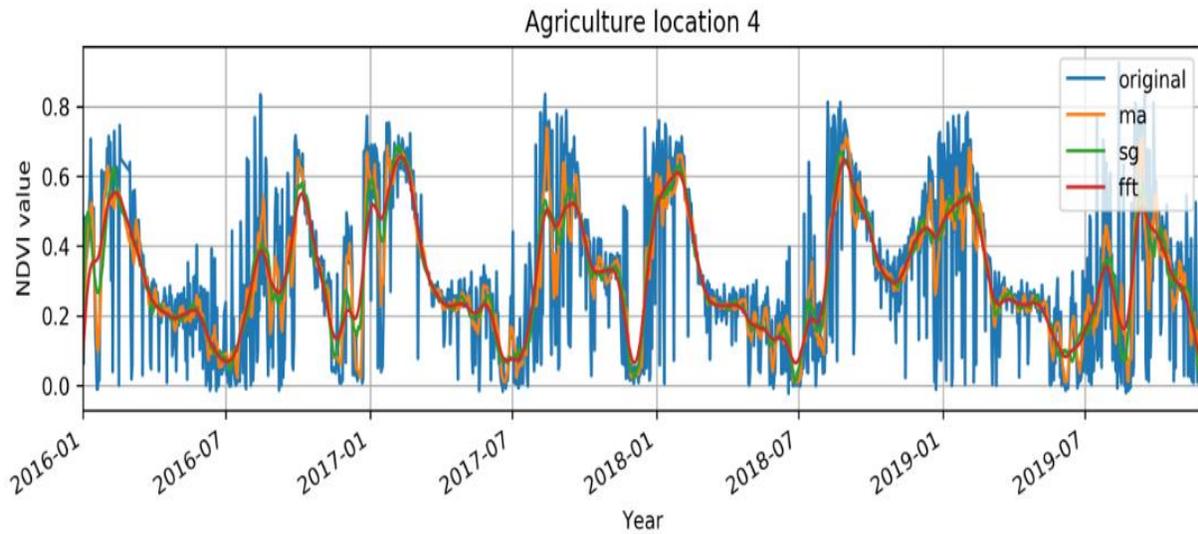
(c)



(d)

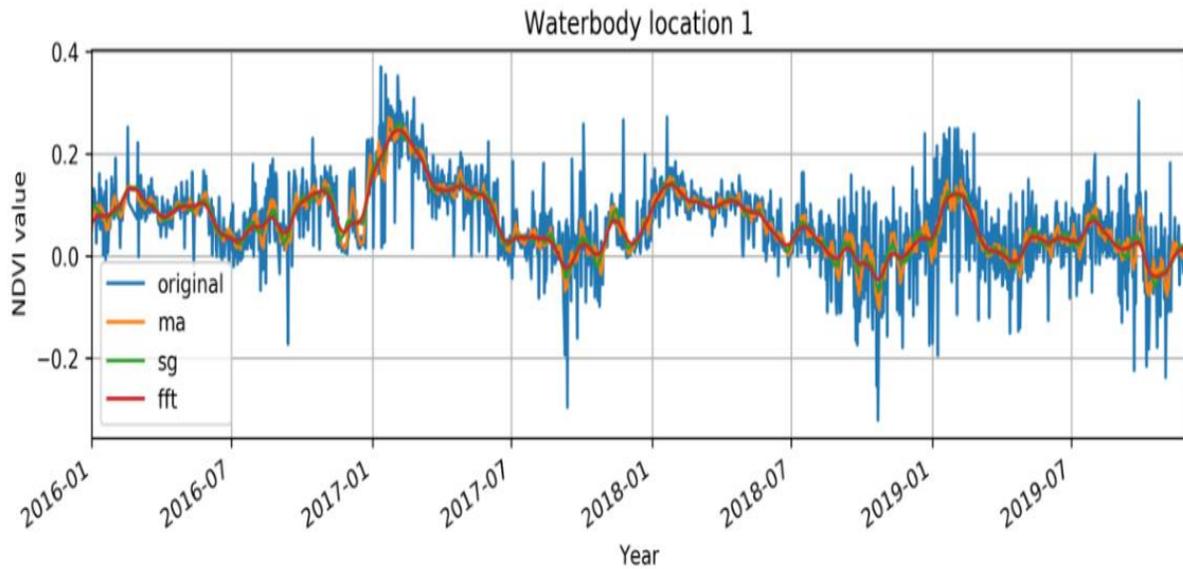
Fig.3 Temporal NDVI profiles of raw and reconstructed datasets using FFT, MA and SG de-noising algorithms for the period of 2016-19 with subplots a) for location 1, (b) for location 2, (c) for location 3 and (d) for location 4



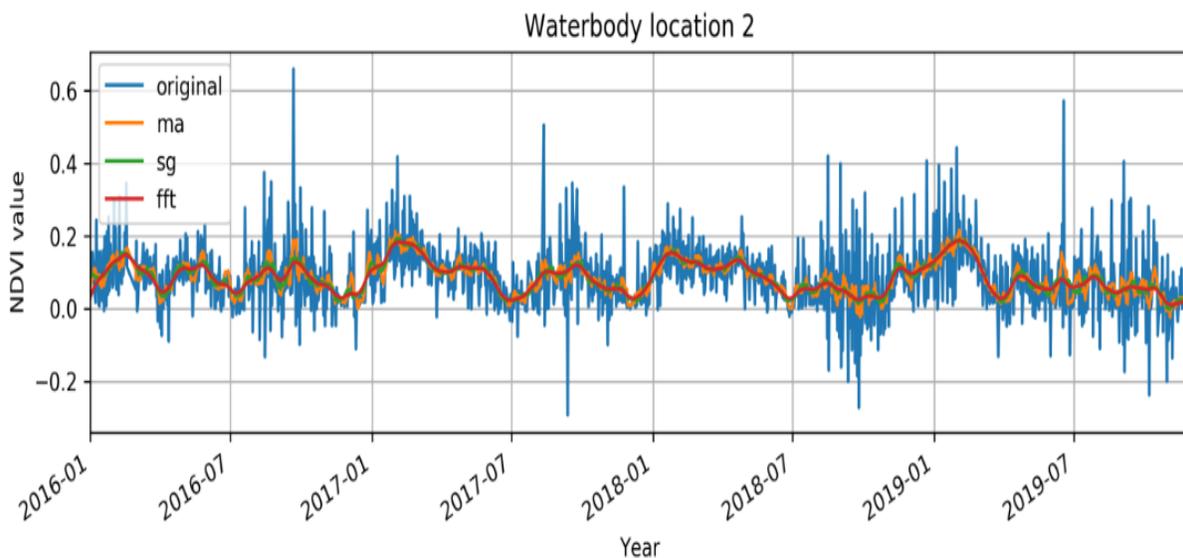


(d)

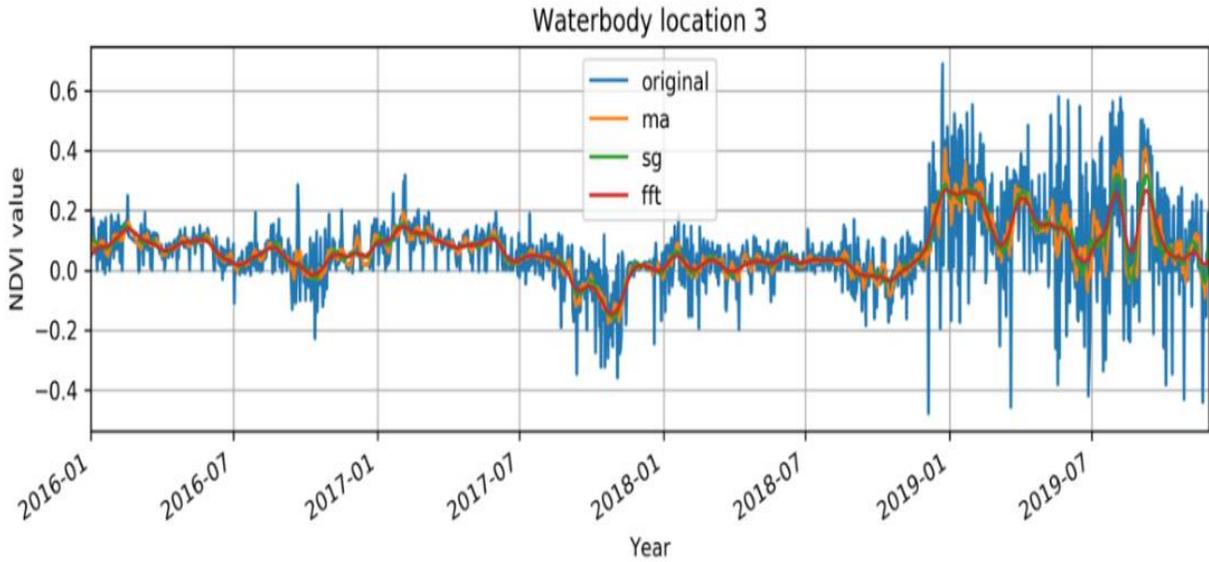
Fig. 4 Temporal NDVI profiles of raw and reconstructed datasets using FFT, MA and SG de-noising algorithms for the period of 2016-19 with subplots a) for location 1, (b) for location 2, (c) for location 3 and (d) for location [33]



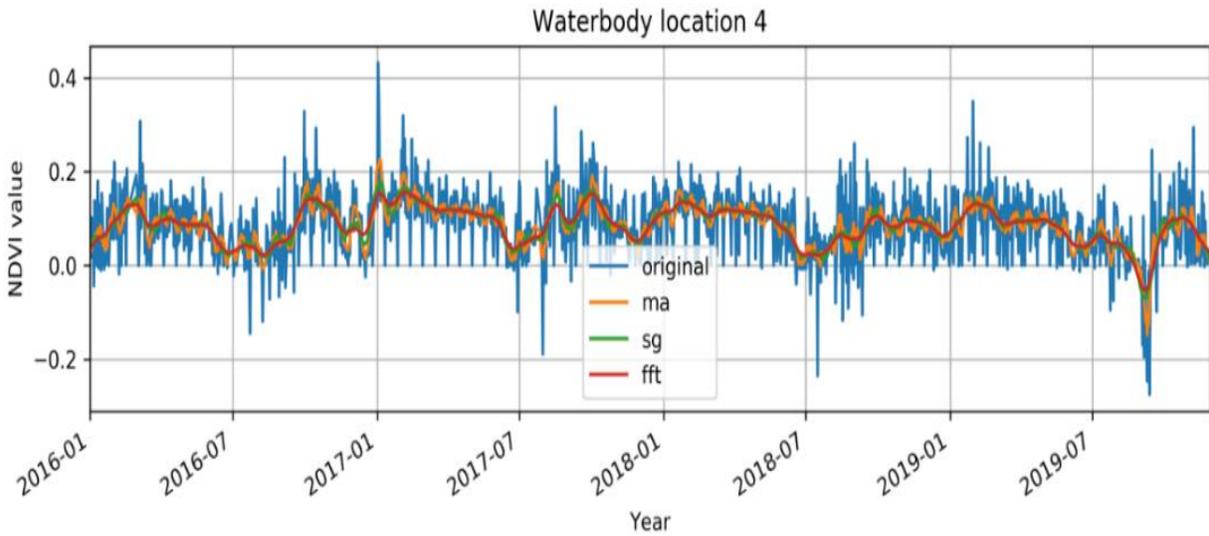
(a)



(b)

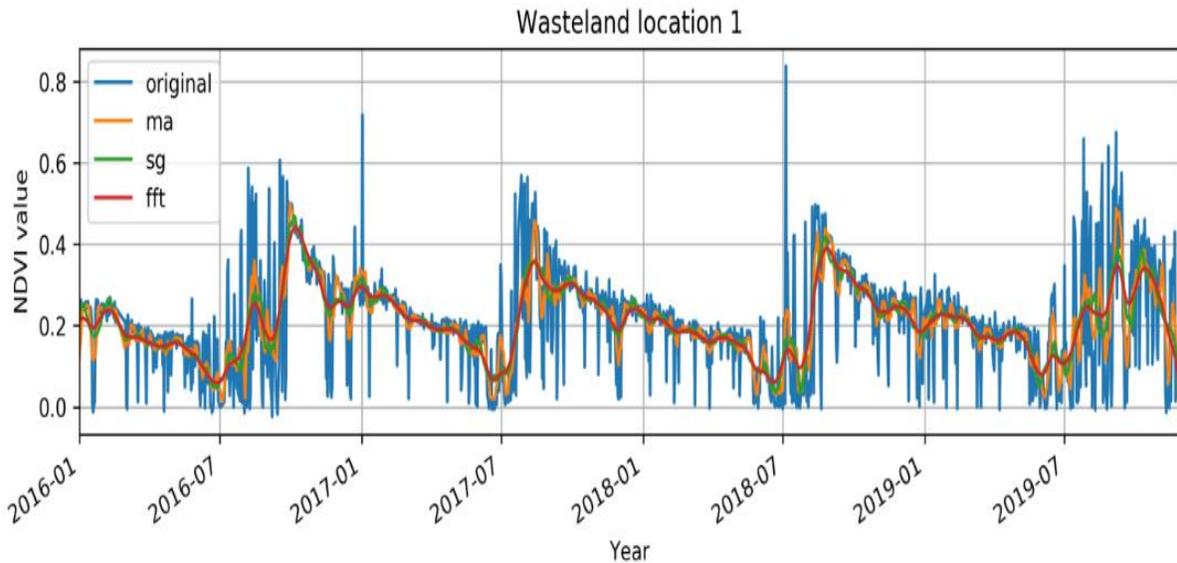


(c)



(d)

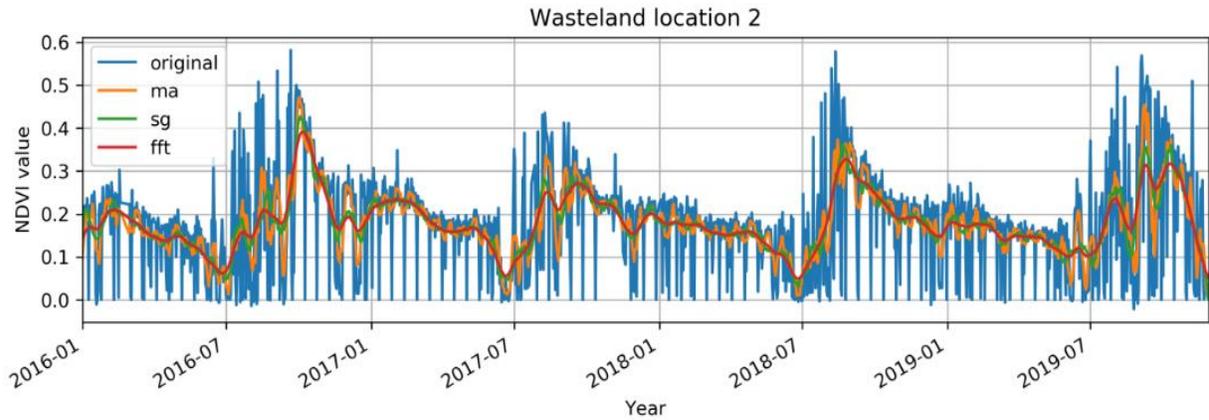
Fig. 5 Temporal NDVI profiles of raw and reconstructed datasets using FFT, MA and SG de-noising algorithms for the period of 2016-19 with subplots a) for location 1, (b) for location 2, (c) for location 3 and (d) for location 4



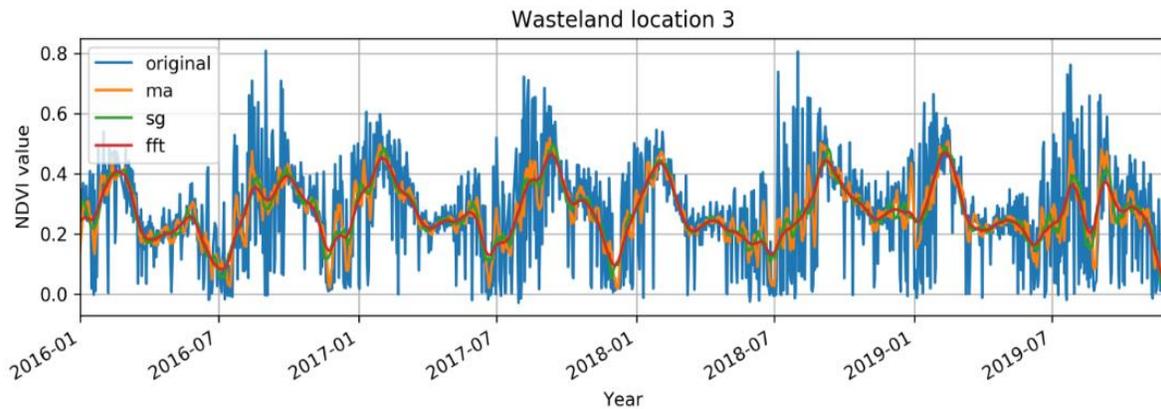
(a)

Table-2 Performance Evaluation of De-Noising Algorithms on Synthetic Signals Using Statistical Evaluation Techniques At Multiple Variance Levels

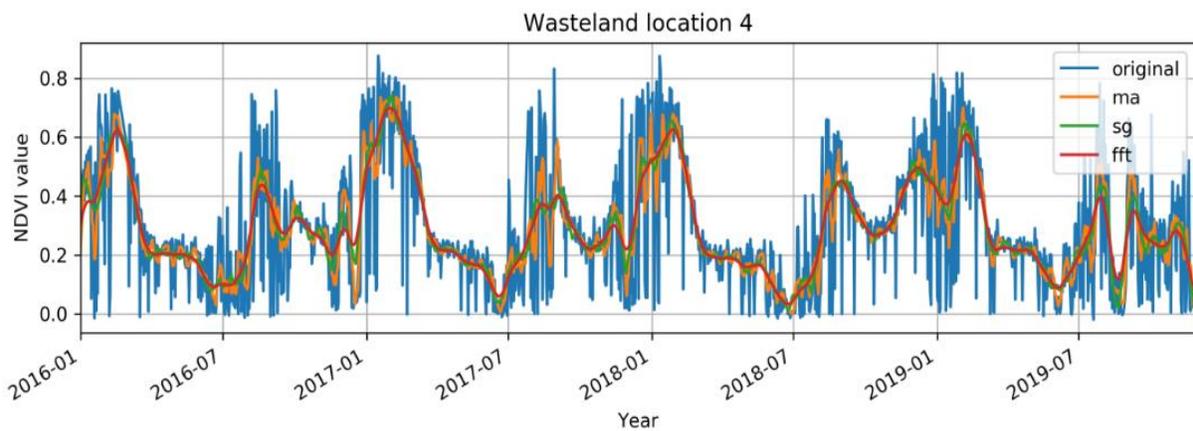
Noise type	Noise variance	Signal	SG		MA		FFT	
			MSE	PSNR	MSE	PSNR	MSE	PSNR
Gaussian	0.01	Sine	0.0092	68.464	0.0100	68.101	0.0098	68.205
		Square	0.1218	57.272	0.0672	59.852	0.0782	59.196
		Triangular	0.0594	60.387	0.0371	62.428	0.0426	61.835
	0.02	Sine	0.0392	62.192	0.0368	62.464	0.0352	62.662
		Square	0.1482	56.421	0.0875	58.706	0.1022	58.034
		Triangular	0.0882	58.671	0.0618	60.215	0.0749	59.384
	0.03	Sine	0.0780	59.220	0.0648	60.014	0.0853	58.818
		Square	0.1895	55.353	0.1302	56.983	0.1451	56.511
		Triangular	0.1305	56.972	0.1085	57.775	0.1237	52.204



(b)



(c)



(d)

Fig. 6 Temporal NDVI profiles of raw and reconstructed datasets using FFT, MA and SG de-noising algorithms for the period of 2016-19 with subplots a) for location 1, (b) for location 2, (c) for location 3 and (d) for location 4

Table-3 Performance Evaluation of De-Noising Algorithms For Different Land Cover Classes For Averaged Locations

Land-cover Class	MA			SG			FFT		
	RMSE	PSNR	Pearson	RMSE	PSNR	Pearson	RMSE	PSNR	Pearson
Forest	0.199	62.215	0.643	0.198	61.639	0.574	0.213	61.628	0.574
Built-up	0.108	67.911	0.563	0.114	67.445	0.489	0.114	67.456	0.493
Agriculture	0.151	64.487	0.704	0.162	63.873	0.647	0.164	63.8	0.641
Water-body	0.094	69.475	0.68	0.101	68.673	0.622	0.101	68.891	0.62
Wasteland	0.111	67.35	0.685	0.119	66.737	0.623	0.119	66.731	0.624

Table-4 Statistical Assessment of De-Noising Algorithms For Averaged Location

Land-cover Class	Parameter	Raw	MA	SG	FFT
Forest	Mean	0.3841	0.3839	0.3844	0.385
	STD	0.2606	0.1628	0.1477	0.138
Built-up	Mean	0.1941	0.194	0.1943	0.1942
	STD	0.1321	0.0732	0.0643	0.1252
Agriculture	Mean	0.3106	0.3108	0.3111	0.3104
	STD	0.2156	0.1488	0.1377	0.1304
Water-body	Mean	0.0755	0.0753	0.0755	0.0754
	STD	0.0949	0.0584	0.0549	0.0525
Wasteland	Mean	0.2382	0.2379	0.2383	0.2381
	STD	0.1556	0.1056	0.098	0.0933

REFERENCES

- [1] Atkinson, P.M.; Jeganathan, C.; Dash, J and C. Atzberger, "Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology," Remote Sens. Environ, vol. 123, p. 400–417, 2012.
- [2] Beck PS, Atzberger C, Høgda KA, Johansen B, Skidmore AK, "Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODISNDVI," Remote Sens Environ, vol. 100, p. 321–334, 2006.
- [3] Brown, J.C.; Kastens, J.H.; Coutinho, A.C.; Victoria, D.D.; Bishop, C.R, "Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data," Remote Sens. Environ. , vol. 130, p. 39–50., 2013.
- [4] Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.H.; Matsushita, B.; Eklundh, L., "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter," Remote Sens. Environ., vol. 91, pp. 332-344, 2004.
- [5] Daren Harmel, R.; Smith, P.K., "Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling," J. Hydrol, vol. 337, pp. 326-336, 2007.
- [6] Dass AK, Shial RK, Gouda BS, "Improvising MSN and PSNR for finger-print image noised by GAUSSIAN and SALT & PEPPER," Int J Multimedia Appl (IJMA), vol. Aug, p. 4:59, 2012.
- [7] Eilers, P.H.C, "A perfect smoother," Anal. Chem., vol. 75, p. 3631–3636. , 2003.
- [8] Fensholt, R.; Proud, S.R, "Evaluation of earth observation based global long term vegetation trends—Comparing GIMMS and MODIS global NDVI time series," Remote Sens. Environ. , vol. 119, pp. 131-147, 2012.
- [9] Filipova-Racheva, D., Hall-Beyer, M, "Smoothing of NDVI time series curves for monitoring of vegetation changes in time," Ecological Monitoring and Assessment Network National Science Meeting 2000, Toronto, Ontario, Canada, no. January, p. 17–22, 2000.
- [10] Gu J, Li X, Huang C, Okin GS., "A simplified data assimilation method for reconstructing time-series MODIS NDVI data," Adv Space Res. , vol. 44, p. 501–509, 2009.
- [11] Guiñón, J. & Ortega, Emma & García-Antón, José & Pérez-Herranz, Valentín, "Moving average and Savitzki-Golay smoothing filters using Mathcad.," 2007.
- [12] Hird, J.N.; McDermid, G.J., "Noise reduction of NDVI time series: An empirical comparison of selected techniques.," Remote Sens. Environ., vol. 113, pp. 248-258, 2009.
- [13] Holben BN., "Characteristic of maximum value composite images for temporal AVHRR data.," Int J Remote Sens., no. 7, p. 1417–1434., 1986.
- [14] Jêdrzej Bojanowski, Wanda Kowalik, Zbigniew Bochenek., "Noise Reduction of NDVI Time-Series: A Robust method based on Savitzky-Golay Filter," 2009.
- [15] Jiang, N.; Zhu, W.Q.; Mou, M.J.; Wang, L.L.; Zhang, J.Z, "A Phenology-Preserving Filtering Method to Reduce Noise in NDVI Time Series.," In Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, no. July 22-27, p. pp. 2384–2387, 2012 .

- [16] Jönsson, P.; Eklundh, L, "Seasonality extraction by function fitting to time-series of satellite sensor data.," *IEEE Trans Geosci. Remote Sens.*, no. 40, p. 1824–1832., 2002.
- [17] Julien, Y.; Sobrino, J.A, "Comparison of cloud-reconstruction methods for time series of composite NDVI data.," *Remote Sens. Environ.*, no. 114, pp. 618-625, 2010.
- [18] Kerr, J.T.; Ostrovsky, M, "From space to species: Ecological applications for remote sensing.," *Trends Ecol. Evol.*, no. 18, pp. 299-305, 2003.
- [19] Kobayashi, H.; Dye, D.G., "Atmospheric conditions for monitoring the long-term vegetation dynamics in the Amazon using normalized difference vegetation index.," *Remote Sens. Environ.*, no. 97, p. 519–525., 2005.
- [20] Forkel, M.; Carvalhais, N.; Verbesselt, J.; Mahecha, M.D.; Neigh, C.S.R.; Reichstein, M, "Trend change detection in NDVI time series: Effects of inter-annual variability and methodology," *Remote Sens.*, vol. 5, p. 2113–2144., 2013.
- [21] Li, J.L.; Zhao, H.B.; Yang, X.P, "Study on the seasonal dynamics of zonal vegetation of NDVI/EVI of costal zonal vegetation based on MODIS data: A case study of *Spartina alterniflora* salt marsh on Jiangsu Coast, China," *Afr. J. Agr. Res.*, no. 6, p. 4019–4024, 2011.
- [22] Lu X, Liu R, Liang S., "Removal of noise by wavelet method to generate high quality temporal data of terrestrial MODIS products," *Photogramm. Eng. Remote Sens.*, vol. 73, p. 1129–1139., 2007.
- [23] Lunetta RS, Joseph FK, Ediriwickrema J, Lyon JG, Worthy LD, " Land cover change detection using multi-temporal MODIS NDVI data.," *Remote Sens Environ.* , vol. 105, p. 142–154, 2006.
- [24] Ma M, Veroustraete F, "Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China.," *Adv Space Res.*, vol. 37, p. 835–840, 2006.
- [25] Moody EG, King MD, Platnick S, Schaaf CB, Gao F., " Spatially complete global spectral surface albedos: Valueadded datasets derived from Terra MODIS land products.," *IEEE Trans Geosci Remote Sens.* , vol. 43, p. 144–158, 2005.
- [26] Niraj Priyadarshi, V. M. Chowdary, Y. K. Srivastava, Iswar Chandra Das & Chandra Shekhar Jha, "Reconstruction of time series MODIS EVI data using de-noising algorithms," *Geocarto International*, vol. 33:10, no. DOI: 10.1080/10106049.2017.1333535 , pp. 1095-1113, , 2018.
- [27] Olsson L, Eklundh L, "Fourier Series for analysis of temporal sequences of satellite sensor imagery," *Int J Remote Sens.*, vol. 15, no. 18, p. 3735–3741., 1994.
- [28] Park, S., "Cloud and cloud shadow effects on the MODIS vegetation index composites of the Korean Peninsula.," *Int. J. Remote Sens.*, vol. 34, p. 1234–1247., 2013.
- [29] Roerink G, Menenti M, Verhoef W., " Reconstructing Cloud Free NDVI Composites Using Fourier Analysis of Time Series.," *Int J Remote Sens.*, vol. 21, no. doi:10.1080/014311600209814., p. 1911–1917., 2000.
- [30] Saribulut L, Teke A, Tumay M., "Fundamentals and literature review of Fourier transform in power quality issues.," *J Electr Electron Eng Res.*, no. doi: 10.5897/JEEER2013.0436.May;5, p. 9–22., 2013.
- [31] Savitzky A, Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal Chem.*, vol. 36, p. 1627–1639, 1964.
- [32] Sellers, P.J., Dickinson, R.E., Randall, D.A., Betts, A.K., Hall, F.G., Berry, J.A.,, "Modeling the exchanges of energy, water, and carbon between continents and the atmosphere.," *Science*, vol. 275, p. 502–509., 1997.
- [33] Setiawan Y, Yoshino K., " Temporal pattern analysis of wavelet-filtered MODIS EVI to detect land use change in JAVA island, Indonesia.," *Int Arch Photogramm Remote Sens Spatial Inf Sci.* XXXVIII, Part 8, Kyoto Japan., 2010.
- [34] Soni A, V., "De-noising of Gaussian noise using discrete wavelet transform.," *Int J Eng Trends Technol (IJETT)*. , vol. Feb;8, p. 309–312., 2014.
- [35] Viovy N, Arino O, Belward AS, "The best index slope extraction (BISE).A method for reducing noise in NDVI time-series.," *Int J Remote Sens.*, vol. 13, no. 8, p. 1585–1590., 1992.
- [36] X. Zhi, S. Yuexin, M. Jin, Z. Lujie and D. Zijian, "Research on the Pearson correlation coefficient evaluation method of analog signal in the process of unit peak load regulation," *13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, Yangzhou, 2017, vol. doi: 10.1109/ICEMI.2017.8265997, pp. 522-527., 2017.
- [37] Lovell, J.L.; Graetz, R.D., "Filtering pathfinder AVHRR land NDVI data for Australia," *Int. J. Remote Sens.*, no. 22, p. 2649–2654, 2001.