

DEVELOPMENT OF AN AUTOMATED SYSTEM FOR MONITORING THE PERFORMANCE OF REFORESTATION ACTIVITIES USING GOOGLE EARTH ENGINE

Rusty A. Lopez (1), Aeron Adrian C. Maralit (1), Enrico C. Paringit (1,2), Czar Jakiri S. Sarmiento (1,2), Regine Anne G. Faelga(1), Fe Andrea M. Tandoc (1), Celeste Z. Vidad (1), Carla Mae M. Arellano (1), Fatima Joy O. Pamittan (1), Ayin M. Tandong (1,2)

¹ Project DIME: Monitoring and Assessment of Planting Activities and other Applications, University of the Philippines, Diliman, Quezon City 1101, Philippines

² Department of Geodetic Engineering, University of the Philippines, Diliman Quezon City 1101, Philippines
Email: rusty.lpz@gmail.com; aeronadrianmaralit@gmail.com; paringit@gmail.com; cssarmiento@up.edu.ph; reginefaelga@gmail.com; tandoc.fe@gmail.com; celeste.z.vidad@gmail.com; carlamae.arellano@gmail.com; fopamittan@up.edu.ph; amtamondong@up.edu.ph

KEY WORDS: Reforestation, NDVI, Time Series Data, Algorithm

ABSTRACT: Reforestation activities are vital to preserve the remaining forest resources of the Philippines. The National Greening Program was established in 2011 to facilitate these activities. It is logically and physically challenging for forest managers to monitor the vegetation growth of all planting sites by conducting field observations. Remote sensing contributes to a more efficient forest growth assessment. The purpose of the study is to test the capability of Google Earth Engine (GEE) to promptly evaluate the growth performance of all reforestation sites using available satellite imageries. Through the GEE platform, satellite data were collected and corresponding vegetation indices were computed for each planting site. Normalized Difference Vegetation Index time series covering the period from 2005 to present were generated to evaluate the vegetation growth of 50,000 reforestation sites. Trend line comparisons of the vegetation index before and after planting were compared to detect vegetation change. On a nationwide level, results prove that GEE is capable of producing fully-automated vegetation growth assessments based on decades-worth of time series data for the Philippines.

1. INTRODUCTION

The National Greening Program (NGP) is a forest rehabilitation program of the Philippine government that started in 2011 and targeted an area of roughly 1.5 million hectares across the whole country. In 2015, the program increased its scope to 7.1 million hectares (DENR, 2019). Given the scale, it is impossible to acquire complete ground data from field observations. Even partial sampling would be logistically impractical for any statistically significant percentage of the sites. Coupled with the fact that all of the analyses must be done within at least a year for annual reporting, balancing fieldwork logistics to the capability of the monitoring system is a major hurdle for a project of this size.

Remote sensing (RS) can be used to address the aforementioned need for large-scale data acquisition methods. Satellite imaging operate through various principles around recording the incoming electromagnetic spectrum. Reflected (optical), emitted (thermal or passive microwave), or scattered (active radar) electromagnetic signals carry information about the properties of a land surface (Joshi, et al., 2016). These data are acquired periodically since the satellites have a fixed revisit time in sun-synchronous orbits (thus it also follows that the images are acquired in similar illumination conditions). The wealth of spatiotemporal data available allows researchers to use time-series analysis in detecting land cover changes.

To quantify these changes, a metric is needed that can reflect the health of vegetation in an area. A simple indicator is the unitless Normalized Difference Vegetation Index (NDVI), defined as:

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (1)$$

Vegetation indices capture the contrast between the 0.65-0.85 μ m wavelength interval because of the distinct absorption spectra of chlorophyll (Myneni, et al., 1995). The absorption peaks at about 0.69 μ m and suddenly dips at around the near-infrared (NIR) region. The larger the contrast between the red and NIR band values generally means that a pixel is dominated by healthy vegetation. The denominator in Eq. 1 normalizes this index to values between -1 and 1. In general, healthy vegetation has the highest positive values, while bare soil, water, snow, ice, or clouds have NDVI values of zero or that are slightly negative (Turvey & McLaurin, 2012). Vegetation under

stress or with a small leaf area has lower positive NDVI values. The simplicity of NDVI gives it wider applicability at the expense of specificity. It can describe the general presence and health of vegetation in an area, but cannot take into account other detailed factors such as species differences, environment moisture content, soil type, etc.

Although RS techniques provide an exponential step in terms of data gathering compared to field observations, it is still cumbersome to acquire and process satellite imagery at scale. This is the advantage of using Google Earth Engine (GEE). Developed by Google, GEE is a web portal providing global time-series satellite imagery and vector data, cloud-based computing, and access to software and algorithms for processing such data. Apart from providing an extensive archive of most publicly available RS data, it also offers free cloud computing thus enabling researchers anywhere in the globe to perform large-scale analysis by harnessing the cutting-edge storage and computation infrastructure of Google (Gorelick, et al., 2017). The availability of the full archive of Landsat Surface Reflectance (SR) data, as opposed to Top-of-Atmosphere (TOA) reflectance, allows for analyses of spectral values across multiple scenes, multiple times, or both without worrying for atmospheric effects to affect the fidelity of the results (Young, et al., 2017). Before GEE, acquiring a single scene would require submitting a request in the USGS website and waiting for a day or more before it can be downloaded (USGS, 2018).

The aim of this study is to assess the performance and feasibility of a system developed around GEE for monitoring nationwide-scale reforestation activities using NDVI time-series data. The bulk work of archiving, downloading, and preprocessing satellite data was offloaded to GEE. Direct NDVI statistical values of a given site and date were directly acquired to generate a time-series database for all sites that was then locally pre-processed, analyzed and maintained. For this study, the NDVI is an adequate index since the nationwide-scale precludes any kind of covariance adjustment and metrics with minimal assumptions are favored, a site's trend is simply compared before and after the implementation of the NGP. The objective is to get a rough idea of the performance of the whole project and not generate a sophisticated model or forecast the growth of the all sites. The scheme presented here only acts as a placeholder part to simulate the performance of site assessment in the context of the overall processing stream. Hopefully, the system could be used as a part of a prototype measurement, reporting and verification (MRV) operationalization scheme for implementing REDD+ projects (Mitchell, et al., 2017) in the Philippines.

2. METHODOLOGY

2.1 Software Used / Environment

The system developed in this study was developed entirely in Python 3.7.3 and uses GEE's Python API (v.1.182) to provide a programmatic and flexible interface to Earth Engine, allowing for automation of cloud-based tasks and extensive scripting capabilities. The system is housed in a virtual machine (VM) and is activated when the necessary input files are detected. The user uploads these files to the virtual machine through a web-synchronized folder, in this case, Dropbox. This current scheme using VMs allows for greater control on the system's access, especially when sensitive datasets are involved.

2.2 Hardware Specifications

The local machine used has 4 cores @ 2.4GHz and 24GB RAM. The settings used by the programs were made to adjust automatically by assessing the operating system's capabilities, so no further tweaking is expected in the eventuality of implementing the system on a different machine. However, it is advised to allocate enough RAM that can hold the expected input size of the local dataset.

2.3 Local Dataset

A file with digitized georeferenced boundaries (*shapefile*) of a subset of all project sites of the NGP from 2011 to 2018 was obtained from the Department of Budget and Management. The attribute table of the shapefile has 50000 rows, one for each site, and contains various columns with pertinent information for each NGP site such as location, year planted, species planted, area size, unique feature ID etc. The total area coverage of the included sites in this file is roughly 1.08 million hectares.

2.4 Google Earth Engine Dataset

For this study, 2005 was chosen as the earliest year to acquire images for the time-series NDVI data. This provides 6 years' worth of historical data for the earliest NGP reforested sites of 2011. Landsat 5 and 7 has data for this timeframe and were both chosen to maximize the number of potential data points, while Landsat 7 & 8 should sufficiently cover the more recent dates. With this in mind, Landsat 5 (LANDSAT/LT05/C01/T1_SR), 7 (LANDSAT/LE07/C01/T1_SR), and 8 (LANDSAT/LC08/C01/T1_SR) SR Tier 1 data was used to ensure consistent

quality through time and across instruments (USGS, 2019). Details about the source satellites are shown in Table 1. Tier 1 includes Level-1 Precision and Terrain (L1TP) corrected data that have well-characterized radiometry and are inter-calibrated across the different Landsat instruments. These data are also radiometrically calibrated and orthorectified using ground control points (GCPs) and digital elevation model (DEM) data to correct for relief displacement.

Table 1. Summary of Landsat Imagery in GEE

Image Collection	Description	Data Availability (Time)	Resolution (Meters)	Revisit Interval (Days)	Provider
Landsat 5	Tier 1 and 2 (Raw, TOA reflectance, surface reflectance); 8 day, 32 day and annual composites	1 January 1984–5 May 2012	30	16	USGS
Landsat 7		1 January 1999–present			
Landsat 8		11 April 2013–present			

Shuttle Radar Topography Mission (SRTM) digital elevation and slope data was also acquired as auxiliary data that could be useful when deliberating performance or planning future field work activities by taking into account the general terrain conditions of a site. This SRTM V3 product (SRTM Plus) is provided by NASA JPL at a resolution of 1 arc-second (approximately 30m).

2.5 System Outline

The system is structured to archive each run such that the database accumulates over time. For example, if the first run acquired GEE NDVI time-series data for n sites from 2005 to 2010, and the next run requested to process m sites from 2005 to 2012, the system also updates the first n sites with new 2010-2012 data, resulting to an updated database with $n+m$ sites with 2005-2012 NDVI data and a static SRTM derived elevation and slope data.

These are the required input files:

1. Shapefile ensemble (.cpg, .shp, .dbf, .prj, .shx), containing the geographical bounds and attribute table of the sites to be processed
2. Settings text file, that contains the end-date for image acquisition, since 2005 is the default starting time.

Initially, a script runs every minute that checks if a new shapefile and settings file are present in the specified Dropbox folder. Upon finding the required files, the checking script is paused to avoid the possibility of simultaneous runs that could clash with each other and affect the integrity of the database. Then, the main script for the system is executed and a run process is started.

2.5.1 Initial Input Checks and Shapefile Uploading (Steps 1-4)

On activation, the user inputs are first checked for validity. The shapefile is tested for overlaps to the existing database to prevent possible site duplication or user error while the user settings are tested that a valid date is given. If there are errors, a text file is created in the Dropbox folder to notify the user to retry and the faulty input files are deleted. If the inputs are good, the run is now formally started. New directories are generated to store the intermediate and resulting files of this run, and a timing benchmark text file is generated. Raw files from previous runs are copied as well, for use in updating and merging to this new run's resulting database. The shapefile is then prepared for upload by storing the attribute table contents to a new .csv file, the metadata file. A new smaller file with only geographical and projection information stored in it is automatically uploaded to GEE as an asset.

2.5.2 Post-processing and Analytics Generation (Step 6-18)

The downloaded metadata and the raw GEE data are then merged with their previous run's counterparts. This updated dataset is then converted from .db files to more easily viewed .csv files. The slope data is stored to a newly created results file. At step 7 we have 3 main files that are needed for post-processing.

1. Metadata – containing additional info about the sites
2. Data – time-series NDVI data
3. Results – new file to contain newly calculated data. The static slope data is also stored here.

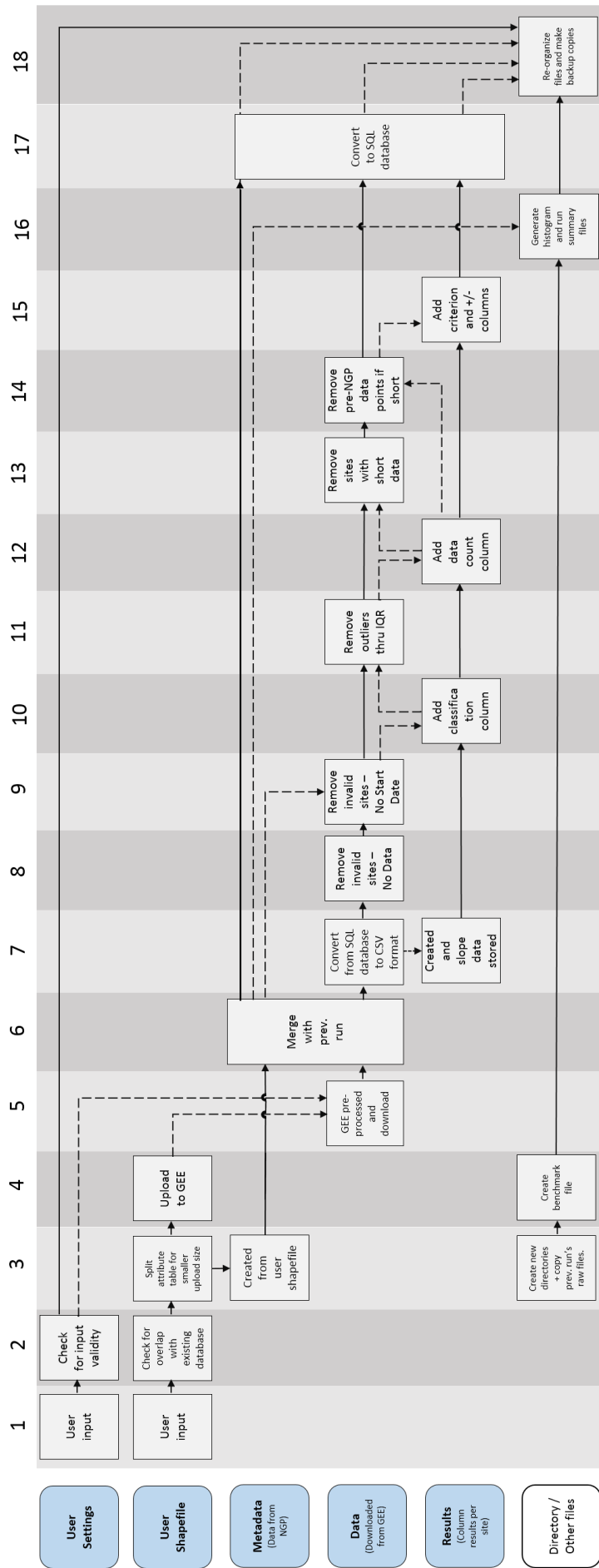


Figure 1. System flowchart that highlights major operations done to the files involved. Solid arrows indicate a major or direct manipulation of data, while dashed lines indicate that the information from the file was used but the file itself is not altered. Step numbers are shown above as a guide in discussion.

Table 2. Post-processing summary table

Step	Action done	File modified	Information from file:	Notes
8	Remove sites with no data gathered	Data	-	Usually no data gathered due to failing the GEE filters.
9	Remove sites with no start date	Data	Metadata	Input error from the user. This is necessary since the trend analysis rely on the declared start date in the metadata.
10	Add classification column	Results	Data	This is a simple script that attempts to determine if a site is urban or mangrove in behavior, due to an unnatural abundance of erratically low NDVI values. Those flagged as such will then be excluded from the outlier removal step or risk removing actual data points.
11	Remove outliers thru IQR	Data	Results	Inter-quartile range (IQR) is used to improve the performance against outliers and non-normally distributed data points (i.e. consistently well vegetated areas). See Fig. 2 for a comparison against Z-score.
12	Add data count column	Results	Data	Used when removing sites based on data point count
13	Remove sites with short data	Data	Results	For this study, a minimum threshold of 10 points was used.
14	Remove pre-NGP data points if short	Data	Results	Since we are actually comparing two trends, not only should the overall data point count be greater than the selected threshold value, but each part as well. For example, a site could have 20 data points, but only 5 of those are from the pre-NGP dates, so in essence only the post-NGP data points are compliant to the threshold rule. Thus, pre-NGP data points are removed. See Fig. 2.
15	Add criterion columns	Results	Data	The criterion results are calculated based on a simple flowchart. See Fig. 3. A simple “early-data” behavior column is also added. It is done by checking if the mean NDVI of the first $n\%$ of the post-NGP data points are already above 0.6, or densely-vegetated (D, et al., 2016). This could be a useful metric to see if an area is already densely vegetated during the earlier dates post-NGP.
16	Generate histograms	Histograms	Metadata	Area size, species planted histogram are generated for additional possible useful data for analysis.

Table 2 shows other details from the various post-processing steps in Fig. 1. The post-processing step cleans the dataset by removing invalid sites based on different factors, as well as outlier data points. At step 15, a rule-based flowchart based on linear regression statistics comparison is used to act as a simple metric for assessing a site’s performance. Fig. 3 shows the flowchart. This was chosen because a simple linear fit comparison between two temporal trends has minimal assumptions to the data, and thus can be generally applied on different data sets. This is especially appropriate for the NGP because most sites have multiple species of tree planted in it. In addition, sites are scattered across the Philippines, making any modelling with environment variables challenging. For sites without pre-NGP data, the slope simply needs to be positive to get a positive rating, interpreted as that the site’s vegetation density/health has grown post-NGP activities, but unfortunately no pre-NGP data can be used for comparison. For sites with pre-NGP data points however, both the slope and mean values must be better post-NGP. This ensures that sites with increased rate of growth and average vegetation density gets flagged with a positive rating. There are 8 possible combinations in this scheme, with 5 and 3 cases for negative and positive performance

ratings, respectively.



Figure 2. Top: IQR and Z-Score outlier detection comparison. Bottom: Example of a site with <10 pre-NGP and >10 pre-NGP data points. The pre-NGP data points will be removed on post-processing step 14. The red dashed lines are linear fits, the black vertical dashed line is the year of NGP planting, orange data points are tagged as outliers, and the blue rectangles indicate the wet season.

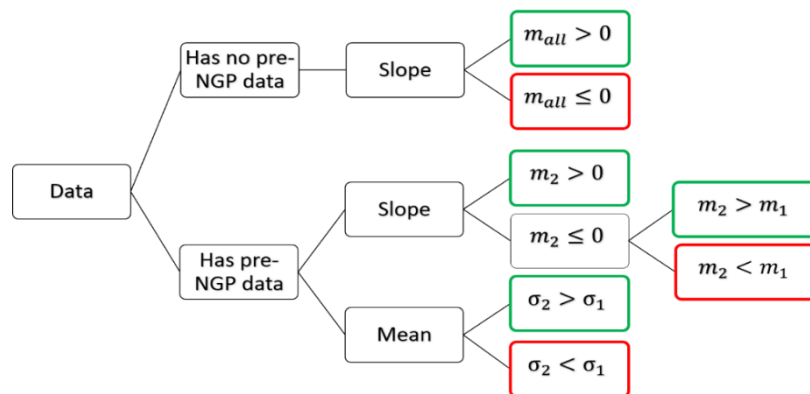


Figure 3. Assessment flowchart used in Step 15. Green and red outlines denote a positive and negative performance for the site, respectively. The subscripts 1 and 2 correspond to pre and post-NGP timeframes.

Note that this is entirely based on values comparison, and not by using specified threshold values. This is meant to be a temporary system that can be modified or completely replaced as per the request of an implementing agency for the use of their own internal metrics. The next step is to generate histograms for the area and species planted in sites as ancillary data. Along with the site elevation value, these were included to help in contextualizing the analysis of time-series data. Finally, the run's directories are organized and temporary files deleted in preparation of the next run to allow seamless integration.

3. RESULTS AND DISCUSSION

The system completed its run after 40h 19m 18.5s. As expected, downloading the GEE data took the longest, as seen in Table 3. Although each individual thread processed each site at around 30-50s (sometimes even reaching 200-300s, probably depending on GEE's traffic), the parallel computing done by multithreading became equivalent to downloading a single site's data for 2.8s in a single thread implementation. Including the post-processing duration of an hour would still yield 2.9s per site since local processing only constituted 2.8% of the total duration.

Table 3. Various result statistics

Result Statistic	Value
Total run time	40h 19m 19s 145159
GEE data download time	39h 09m 36s (97.1%)
Post-processing time	01h 07m 08s (2.7%)
Shapefile upload, etc.	00h 02m 35s (0.1%)
Data timeframe	2005-01-01 to 2019-01-01
Total number of sites	50,000
Total eligible for criterion evaluation	44,147 (88.3%)
Total skipped	5,853 (11.7%)
No start date in metadata file	120
No data (either GEE Error or no collected data due to masking thresholds)	469
Below data point count threshold	5,259 (5 overlaps with no start date)
Total number of NDVI time series data points downloaded	4,549,652
Median number of images acquired per site	41

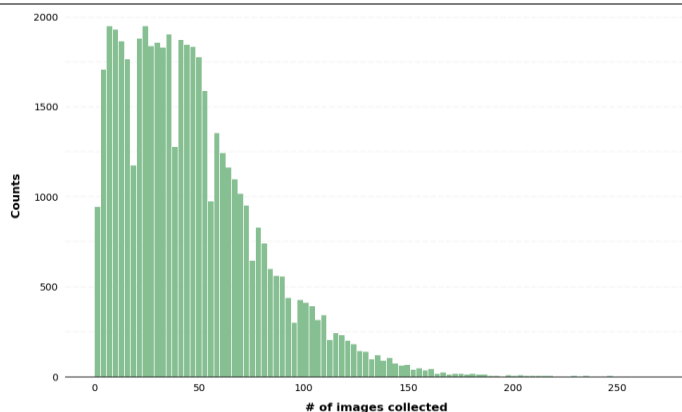


Figure 4. Histogram of the number of data points collected per site. The median value is 41.

For the post-processing step, majority of the skipped sites are due to having less than the threshold value set at 10

data points. Looking at Fig. 4, the histogram of the data point population per site is skewed to the lower values, with a median value of 41. Given the 14-year timeframe, this is roughly equivalent to getting only 3 valid data points per year. This shows the effect of the stringent GEE mask thresholds used to maintain cloud and error-free images. It is worth considering the limitations of acquiring optical images due to clouds cover. If an analyses required a greater density of images in a given time period (e.g. growth modelling of fast growing trees), using optical imagery might be insufficient. There were also only 469 (<1%) sites with no data collected, and this includes sites with totally

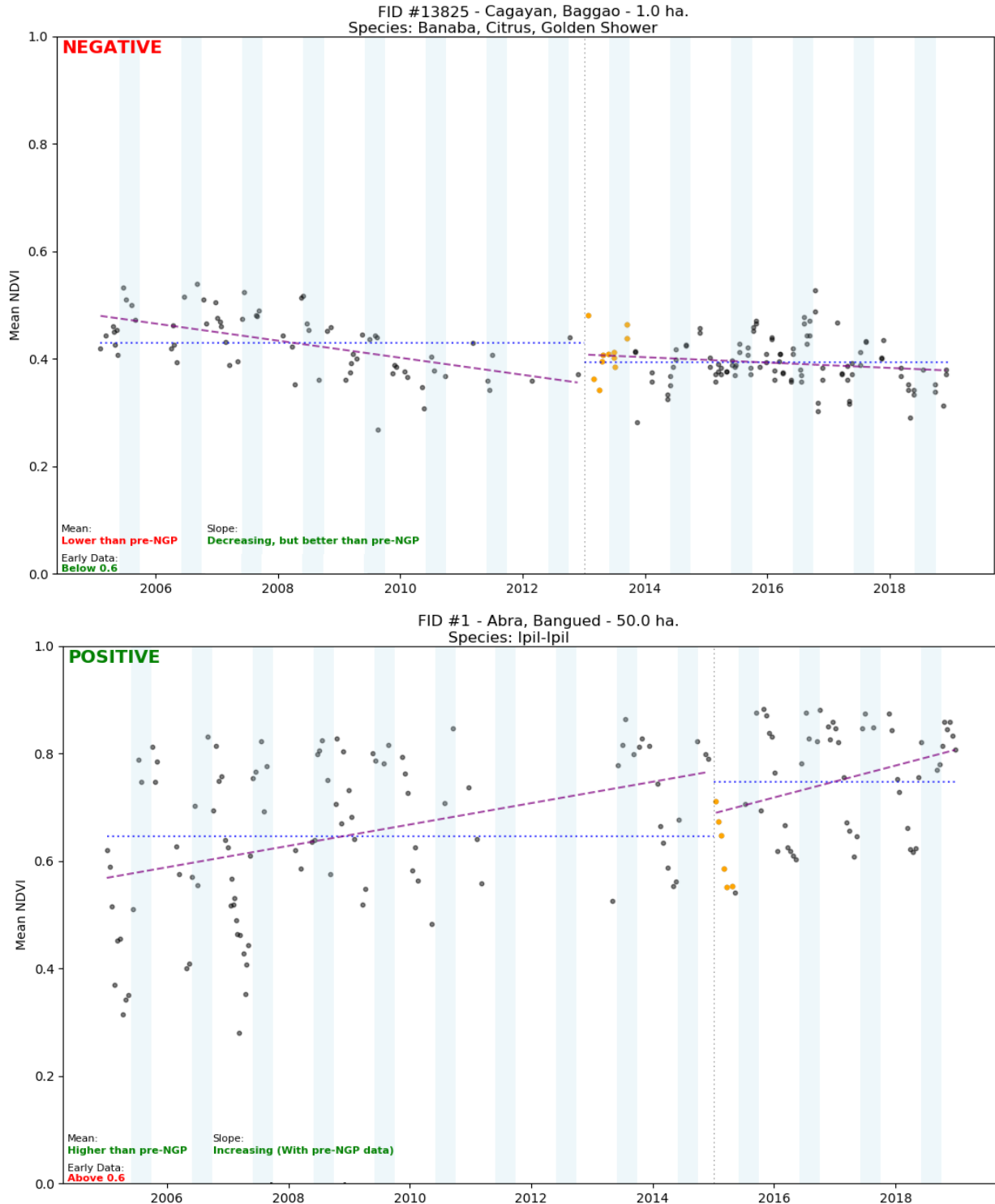


Figure 5. Examples of the assessment results. Purple dashed lines are linear fits, blue dotted horizontal lines are the mean value, the black vertical dashed line is the year of NGP planting, orange data points are the first 10% post-NGP data points, and the blue rectangles indicate the wet season. The location, size, and species planted in the site are on top. The text at the top-left is the assessment result.

filtered images due to the used GEE thresholds. This is a measure of GEE’s reliability as a cloud database for Landsat

SR images. This value is expected to further improve upon finishing multiple runs since they system would queue these sites that previously timed-out their computation in the GEE server.

The assessment results in Fig. 5 provide a graphical understanding of the rule-based assessment flowchart of Fig. 4. The top image shows a site with a negative assessment rating. From the pre-NGP NDVI, the vegetation density is decreasing. After the NGP activities, the slope is still negative, but better than before. Due to the relatively high values of NDVI circa 2005-2008, the post-NGP mean is lower ultimately giving the site a negative rating. The bottom image shows a site with NDVI data points that decreases during the dry season but still generally increases over time. The gap in data at around 2012 further highlights the effect of the GEE mask thresholds. The post-NGP values are around 0.7, signifying the presence of a consistently dense vegetation on the site.

The current rule-based assessment system is capable of a rough estimation of a site's performance by using simple trend comparison. It is best applied when looking at the entire dataset's statistics due to the generality of the methodology used. Even outliers in the individual site data assessment is still useful for exploring the field conditions that could have caused those. As an example from the top image of Fig. 5, it could be argued from NDVI behavior that there could be issues with the site's environment (soil quality, climate, anthropogenic factors, species compatibility, etc.) because regardless of human intervention, the vegetation in the area seems to struggle to grow properly.

4. CONCLUSION

Using only a single machine and without any user intervention after the initial data input, the GEE-based system proved capable in downloading and processing 14 years' worth of satellite imagery for 50000 sites to produce analysis ready datasets in less than 2 days. From the results, it was found that it is feasible to use more computationally expensive post-processing methods in the future because it only constitutes <3% of the total run time duration. With the modular design of the system, using a different GEE dataset and an altered post-processing chain can be implemented without any major refactoring. Error handling and data update is also built-in, so that even consecutive runs can be done without having to manually re-organize the database.

5. ACKNOWLEDGEMENTS

The authors would like to thank the Department of Budget and Management for the NGP data used in this study.

6. REFERENCES

DENR, 2019. ENHANCED NATIONAL GREENING PROGRAM, Retrieved September 1, 2019, from: <https://www.denr.gov.ph/index.php/priority-programs/national-greening-program>

D, J., Reddy, S. N. & Manikiam, B., 2016. Land Cover Classification based on NDVI using LANDSAT8 Time Series: A Case Study Tirupati Region. International Conference on Communication and Signal Processing, pp. 1332-1335.

Gorelick, N. et al., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, Volume 202, pp. 18-27.

Joshi, N. et al., 2016. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. Remote Sensing, 8(70).

Mitchell, A. L., Rosenqvist, A. & Mora, B., 2017. Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. Carbon Balance and Management, 12(9).

Myneni, R. B., Hall, F. G., Sellers, P. J. & Marshak, A. L., 1995. The Interpretation of Spectral Vegetation Indexes. IEEE Transactions on Geoscience and Remote Sensing, 33(2), pp. 481-486.

Turvey, C. G. & McLaurin, M. K., 2012. Applicability of the Normalized Difference Vegetation Index (NDVI) in Index-Based Crop Insurance Design. *American Meteorological Society*, Volume 4, pp. 271-284.

USGS, 2018. Landsat Data Access, Retrieved September 1, 2019, from: <https://www.usgs.gov/land-resources/nli/landsat/landsat-data-access>

USGS, 2019. Landsat Collection 1 Level 1 Product Definition, Retrieved September 1, 2019, from: <https://www.usgs.gov/media/files/landsat-collection-1-level-1-product-definition>

Young, N. E. et al., 2017. A survival guide to Landsat preprocessing. *Ecology*, 98(4), pp. 920-932.