Time Series Properties

# Input Data

#### Normalized Difference Vegetation Index

The normalized difference vegetation index (NDVI) is commonly used to monitor the status of crops, forests, and ecosystems. NDVI is sensitive to the amount of chlorophyll in any location and used to observe approximate levels of plant productivity.

Given the relatively small scale of agriculture in Belize and the lack of high frequency observations from other satellites, we derive the NDVI using the 250m vegetation products from the MODIS satellites, specifically the 16-day MODIS product MOD13Q1 from the Terra satellite (Didan and Huete 2006). Although relatively low-resolution MODIS satellites aquire images of each part of the globe daily. As such, MODIS’s low spatial resolution is compensated by its high temporal resolution thereby providing rich information for time series statistics that is not plagued by cloud cover. Moreover, a longer time series is available, in this case we summarize the properties of the five years leading up to the study date, Jan 1 2002 – Dec 31 2017. See Table 1 for low-spatial resolution summary statistics.

We supplement this with additional observations from Sentinel 2 data which provides a much higher resolution view of the earth. Sentinel 2 data resolution is handled at 10m resolution. This higher resolution comes at the cost of having fewer images over time, which in cloudy areas like Belize can severely limit the amount of useful information available throughout the year. As such, a more limited set of annual statistics are provided. Due to the limited available of Sentinel data generally, and cloud-free observations generally, summary properties from Sentinel are limited to the study year 2017. **See table XX for high-resolution summary statistics.**

We find that the combination of these two products provides a stable and informative time series. Each set of data provides a unique set of properties based either on high temporal of spatial resolution.

#### Rainfall

We also examine the time series properties of rainfall as measured by the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). CHIRPS is a 30+ year quasi-global rainfall dataset. CHIRPS incorporates 0.05° resolution satellite imagery with in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring[[1]](#footnote-1). In this case we resample the rain data to 75m spatial resolution to ensure that each enumeration area has a observation associated with it. We collect precipitation by dekad (Funk et al. 2015). There are three dekads in a month, the first two being 10 days long, and the third being the remaining days in the month. Because CHIRPS data has a similar high frequency and availability as MODIS data above, we provide the denser set of summary statistics outlined in **Table X for low-spatial resolution data.**

*Radar*

# Specialized Algorithms

#### LandTrendr – Greatest magnitude disturbance mapping

LandTrendr (LT) is a broadly used algorithm that looks to detect sudden shifts in an index, in this case NDVI, on a pixel-by-pixel basis. Effectively LT fits local regressions, the uses a series of metrics to detect sudden shifts in slope, or intercept on a year-by-year basis. As such, LR is effective at identifying land cover change, for instance conversion of forest to agriculture, or agriculture to urban settlement. Because LT only needs one clear observation per year it is effective for use with high resolution data like Sentinel or Landsat another higher resolution satellite. LT has a series of underlying assumptions and parameters, which we will not cover here for the sake of clarity. Detailed information on the algorithm can be found here https://www.mdpi.com/2072-4292/10/5/691.

#### Ts-Raster – Dense temporal feature extraction

Ts-raster (TS) is a python package for analyzing time-series characteristics from raster data. It allows feature extraction, dimension reduction and applications of machine learning techniques for geospatial data https://github.com/mmann1123/ts-raster. TS’s primary significance is the ability to provide an extensive set of time-series properties, including simple metrics like minimums or maximums, but also more complex ones like the number of peaks observed within a year, or the number of observations above or below the mean. Although new, and under development by this coauthor, we believe that TS will be able to meaningfully characterize the time series of higher frequency data products like those from MODIS or CHIRPS.

Table 1: Summary statistics for high temporal resolution data

|  |  |  |
| --- | --- | --- |
| Feature Name | Description | Uses |
| agg\_linear\_trend\_\_f\_agg\_"max"\_\_  chunk\_len\_6\_\_attr\_"slope" | Maximum observed trend during any 6 periods | Sudden positive shocks (flood) |
| agg\_linear\_trend\_\_f\_agg\_"min"\_\_  chunk\_len\_6\_\_attr\_"slope" | Minimum observed trend during any 6 periods | Sudden negative shocks (drought, land use change) |
| count\_above\_mean | Number of observations above the global mean | Persistent shifts up (increase in rainfall) |
| count\_below\_mean | Number of observations below the global mean | Persistent shifts down (decrease in rainfall) |
| last\_location\_of\_maximum | Location of the periods maximum value | Time since maximum observed value (declining productivity) |
| last\_location\_of\_minimum | Location of the periods minimum value | Time since minimum observed value (increasing productivity) |
| longest\_strike\_above\_mean | Longest period of values observed above the global mean | Duration of persistent shifts up (flooding) |
| longest\_strike\_below\_mean | Longest period of values observed below the global mean | Duration of persistent shifts down (drought) |
| maximum | Global maximum value | Highest observed greenness / rainfall |
| mean | Global mean value | Average observed greenness / rainfall |
| mean\_change | Average change between any two periods in series | Instability in time series (irregular rain) |
| median | Global median value | Average observed greenness / rainfall |
| minimum | Global minimum value | Minimum observed greenness / rainfall |
| number\_cwt\_peaks\_\_n\_12 | The highest number of peaks that occur in 12 periods | Number of crop rotations, unstable rainfall |
| number\_cwt\_peaks\_\_n\_6 | The highest number of peaks that occur in 6 periods | Number of crop rotations, unstable rainfall |
| quantile\_\_q\_0.05 | Value of the 5th percentile | Minima correcting for outliers |
| quantile\_\_q\_0.15 | Value of the 15th percentile | Minima correcting for outliers |
| quantile\_\_q\_0.85 | Value of the 85th percentile | Maxima correcting for outliers |
| quantile\_\_q\_0.95 | Value of the 95th percentile | Maxima correcting for outliers |
| ratio\_beyond\_r\_sigma\_\_r\_2 | Ratio of values that are more than 2\*std(x) away from the mean | Frequency of extreme values, flooding, shocks |
| ratio\_beyond\_r\_sigma\_\_r\_3 | Ratio of values that are more than 3\*std(x) away from the mean | Frequency of very extreme values, flooding, shocks |
| skewness | Sample skewness of x (calculated with the adjusted G1 coefficient) | Changes in distribution over time |
| sum\_values | Sum of all values across the period | Total rainfall, total productivity |

Table 2 Summary statistics for low temporal resolution data

|  |  |  |
| --- | --- | --- |
| Feature Name | Description | Uses |
|  | Maximum observed trend during any 6 periods | Sudden positive shocks (flood) |
|  | Minimum observed trend during any 6 periods | Sudden negative shocks (drought, land use change) |

1. https://developers.google.com/earth-engine/datasets/catalog/UCSB-CHG\_CHIRPS\_PENTAD [↑](#footnote-ref-1)