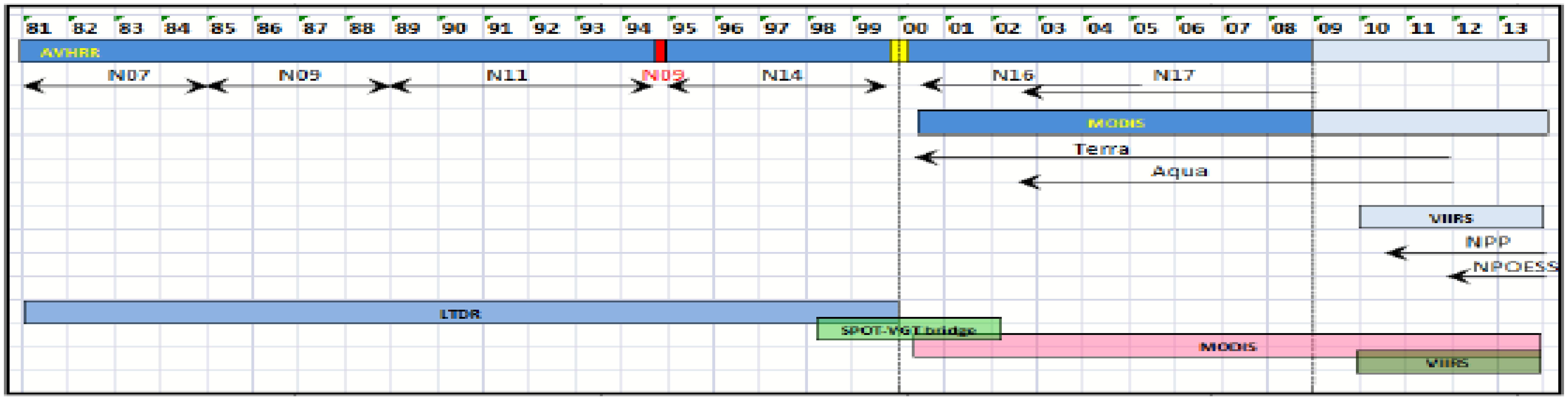
Data Processing Flow Chart

Go Back

Input

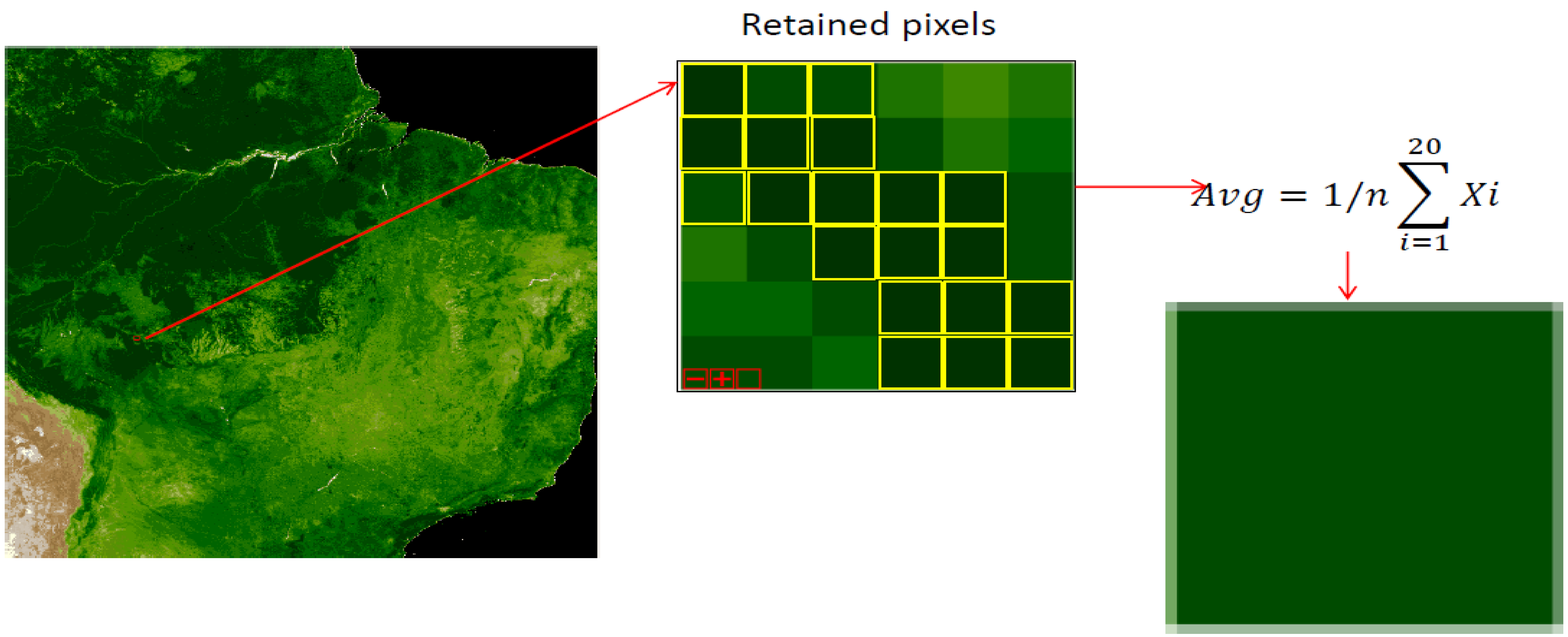
- Data Download
  A 30+ years global CMG daily dataset is downloaded, composed of the following sensors: AVHRR (1981-1999), SPOT (1998-2002) and MODIS (2000-2010). The daily global data from MODIS and LTDR both have 3600x7200 pixels.

- 

  - AVHRR (Missing days)
  - SPOT (Missing days)
  - MODIS (Missing days)

Go Back

SPOT Resampling



- Spatial resolution for SPOT is 1.0 km and for MODIS is 5.6 km, thus in order to combine the data, they must have the same resolution. First of all we have to inspect 6x6 pixels on SPOT image, then filter the data and finally determine the average of the retained pixels (see the figure above). This procedure will achieve a 6 km pixel which is good enough to combine with 5.6km pixel from MODIS.

Back

VIS Estimation

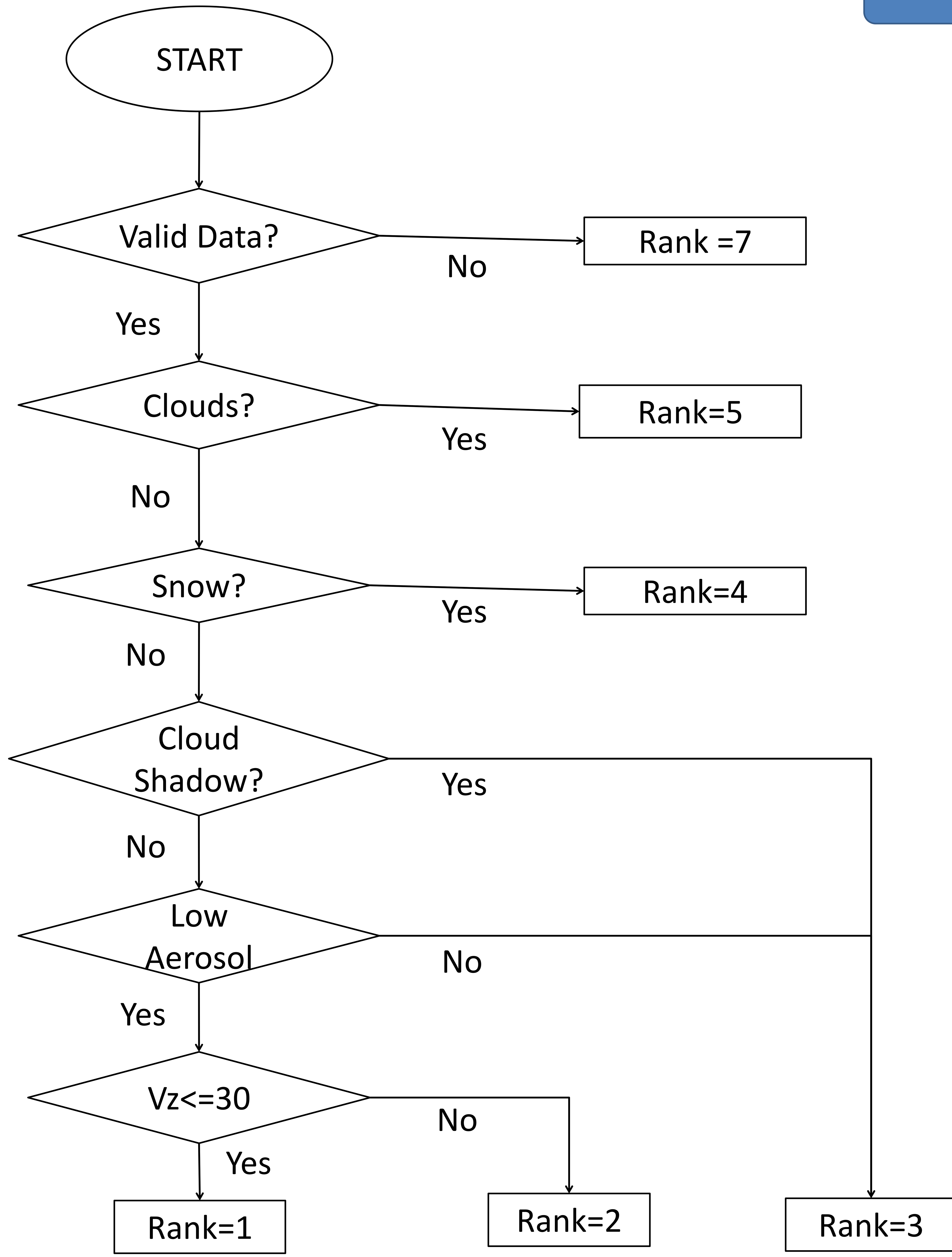
- **Vegetation indices** (VI) are empirical measures that quantities vegetation biomass of the vegetation at the land surface. They often are function of the red and near infrared spectral functions.
- **VIS Estimation:** NDVI and EVI2 sds's are estimated and added to the downloaded data. In addition a Rank layer, describing the quality of the data, based on QA information is added to each file.
- **NDVI & EVI2:** As a ratio, the NDVI has the advantage of minimizing certain types of band-correlated noise (positively-correlated) and influences attributed to variations in direct/diffuse irradiance, clouds and cloud shadows, sun and view angles, topography, and atmospheric attenuation. On the other hand, EVI (Enhance Vegetation Index) was developed to minimize the atmospheric effect by using the difference in blue and red reflectances as an estimator of the atmosphere influence level.

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

$$EVI2 = 2.5 * \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + 2.4 * \rho_{red} + 1}$$

Go Back

Data Filtering:



Note:
The rank 6 was used later on in the process to identify the data generated using the gap filled technique.

Go Back

Rank 7

The first aspect evaluated was the validity of the data. The data was considered not valid when at least one of the following factors occurred:

- surface reflectance value is out of the range,
- the area is not coverage by the sensor swath,
- instrumentation failure and/or high view zenith angles ($>85^{0}$).

Go Back

Rank 5 and 4

- The second aspect was the presence of clouds on the data. If there is clouds, then the pixel is ranked as 5.
- The presence of snow on pixels was ranked 4.

Go Back

Rank 1, 2 and 3

The pixels which passed the above filtering (clouds and snow) were taken to the next step where they were analyzed for cloud shadows and for aerosols which are normally the cause of poor quality when there are no clouds. Then, if the aerosols were low the data was evaluated to determine the influence of the view zenith and if this was larger than a pre-defined value (i.e.30°) this data was considered negatively affected by this aspect.

- 1 being ideal data,
- 2 good to marginal data and requires additional post-processing,
- 3 marginal to questionable data

Go Back

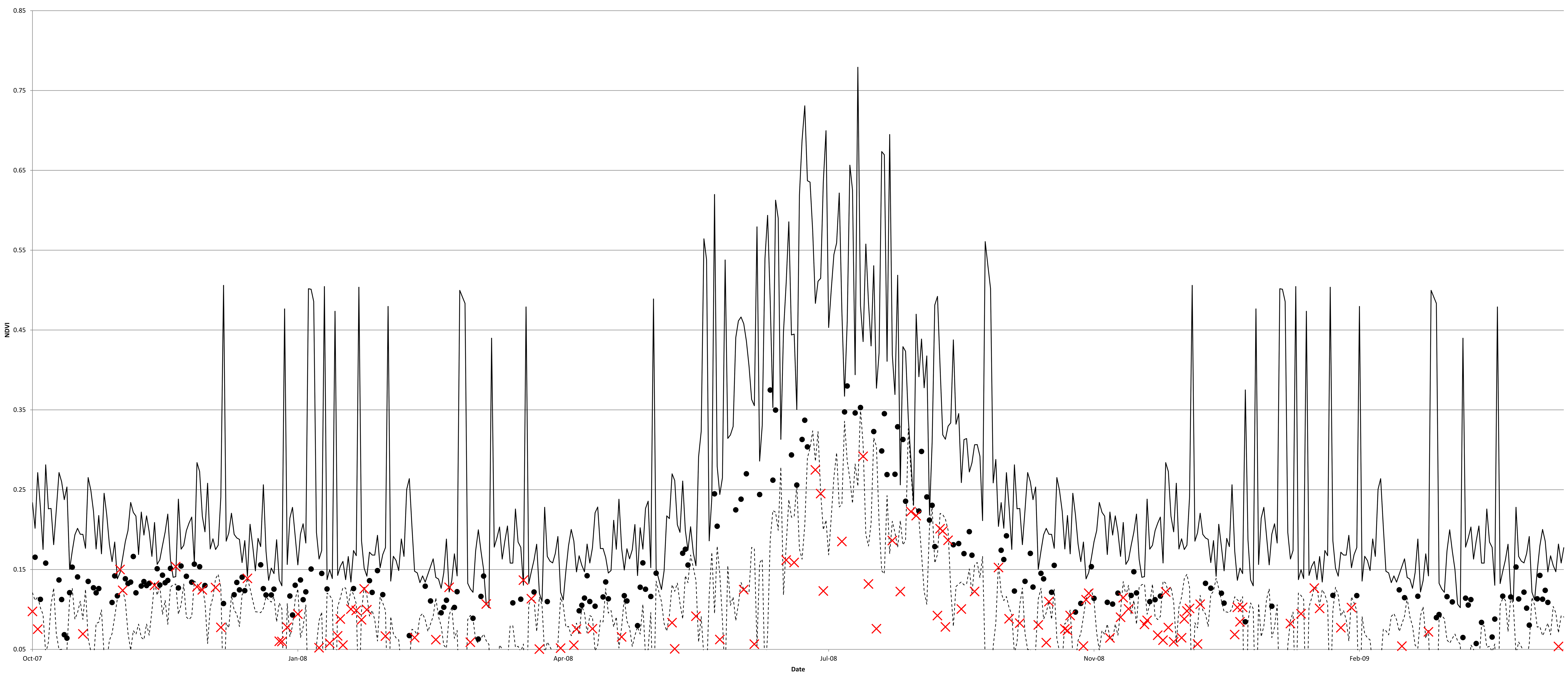
Long Term Average Estimation:

A second filter, using a long term data record, was considered to ensure the quality of the data. A long term average (LTAvg) profile was determined using both MODIS and AVHRR datasets and a confidence interval based on the standard deviation was established. A moving window of five years was used to determine the long term average profile for most pixels. For pixels where five years did not provided enough data, longer periods were used as necessary. The long term averages periods used in this project were 5, 10, 20 and 30 years period (Figure below). Example

Go Back

Data Filtering using Long Term Average Data:

Vegetation Index profile for one year constrained by the long term average using daily information (see the black dots, •). The continuous line is the long term average plus one and a half standard deviations and the dashed line is the long term average minus one standard deviation. In this case only the data point denoted by the X's are rejected.



Go Back

Continuity

Continuity Data: A seamless continuous dataset is produced by applying the continuity equations derived from MODIS, SPOT and AVHRR data records from the overlap period. Two different methods are used:

1) Top-Down
2) Bottom-up (this approach was implemented just in version 1)

Top-down, Direct Image Comparison (for LTDR v.3)

Go Back

Spectral Transformation Equations to MODIS-equivalents (TOC, CMG)

NDVI (x variable)	Equation	Uncertainty (95% PI)
N-7 AVHRR, ROW, GAC	$y = -0.0646111 + 1.2409713x - 0.0304219x^2$	±0.0138
N-9 AVHRR, ROW, GAC	$y = -0.0621082 + 1.2487272x - 0.0307315x^2$	±0.0138
N-11 AVHRR, ROW, GAC	$y = -0.0606805 + 1.2456808x - 0.0335204x^2$	±0.0138
N-14 AVHRR, ROW, GAC	$y = -0.0571829 + 1.2372178x$	±0.0138
S-4 VEGETATION, TOC, CMGV	$y = 0.0156834 + 1.0610148x$	±0.061

EVI2 (x variable)	Equation	Uncertainty (95% PI)
N-7 AVHRR, ROW, GAC	$y = -0.0403338 + 1.2400319x$	±0.088
N-9 AVHRR, ROW, GAC	$y = -0.0403338 + 1.2400319x$	±0.088
N-11 AVHRR, ROW, GAC	$y = -0.0403338 + 1.2400319x$	±0.088
N-14 AVHRR, ROW, GAC	$y = -0.0403338 + 1.2400319x$	±0.088
S-4 VEGETATION, TOC, CMGV	$y = 0.0085842 + 1.1557716x$	±0.037

Bottom-up, Hyperspectral Analysis

Go Back

Spectral Transformation Equations to MODIS-equivalents (TOC, CMG)

NDVI (*x* variable)	Equation	Uncertainty (95% PI)
N-7 AVHRR, ROW, GAC	y = 0.0105080 + 1.1144501x	±0.033
N-9 AVHRR, ROW, GAC	y = 0.0127476 + 1.1215841x	±0.032
N-11 AVHRR, ROW, GAC	y = 0.0143102 + 1.1167148x	±0.032
N-14 AVHRR, ROW, GAC	y = 0.0143951 + 1.1336442x	±0.030
S-4 VEGETATION, TOC, CMGV	y = 0.0381324 + 1.0064999x	±0.013

EVI2 (*x* variable)	Equation	Uncertainty (95% PI)
N-7 AVHRR, ROW, GAC	y = -0.000084 + 1.2339542x	±0.023
N-9 AVHRR, ROW, GAC	y = 0.0023720 + 1.2298151x	±0.022
N-11 AVHRR, ROW, GAC	y = 0.0033594 + 1.2256970x	±0.022
N-14 AVHRR, ROW, GAC	y = 0.0044528 + 1.2244740x	±0.022
S-4 VEGETATION, TOC, CMGV	y = 0.0232545 + 1.0324644x	±0.006

Go Back

GAP Filling

Gaps are filled using

1. Linear Interpolation
2. Inverse Distance Weighting.

$$VI_j = \frac{\sum_i \frac{VI_i}{d_{ij}^n}}{\sum_i \frac{1}{d_{ij}^n}}$$

VI_i is the vegetation index value of the known points
d_{ij} is the distance to the known point
VI_j is the vegetation index value of the unknown point
n is a power parameter , user selects the exponent (often 1, 2 or 3)

3. Values are constrained by the long term average moving window of 5, 10, 20 or 30 years. One standard deviation is used to restrict the boundaries of the values. Values outside of boundaries are replace with a long term average value and labeled within the Rank sds.

Go Back

Compositing

- Compositing is a procedures to improve the quality of land products. It combines multiple daily images to generate a single cloud and problem free image over a predefined temporal intervals. This method reduces the noise due to the clouds and atmospheric constituents [Jonsson et. al. 2004]. The compositing can be the first filter to get a better and more accurate time series data. One type of composting is the maximum value composite (MVC). MVC compares all the images taken by a satellite, such as MODIS, during a pre-defined period of time and selects the pixels with the highest vegetation index value since it is assume that contamination reduces the VI values [Viovy et. al. 1992].
- Daily data is used to generate composed images. A 15-days and Monthly datasets are generated. Each one based on the following approaches

    a) CV-MVC (Constrain View-Maximum Value Compositing): it minimizes the off-nadir tendencies of MVC.

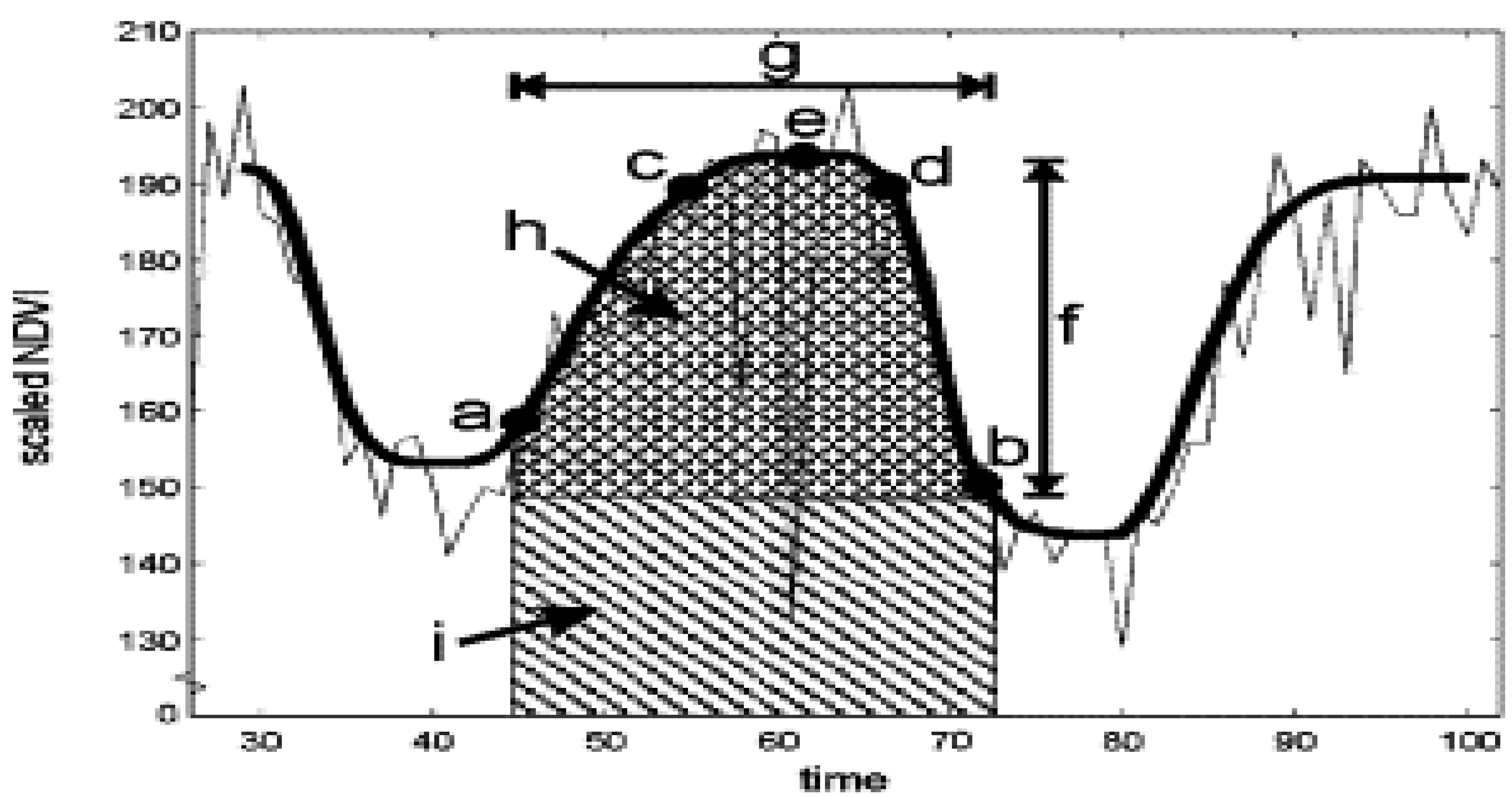
    b) Average of All values

    c) Average of N max values

Go Back

Phenology

- Vegetation phenology can be defined as the plants study of the biological cycle events throughout the year and the seasonal and interannual response by climate variations. Phenology products, produced daily or on any compositing period, provided different parameters which describe the seasonal behavior of the vegetation.
- In general, the phenology is represented graphically it has a bell shape. The graphic below exhibits the following parameters: start of season (a), end of the season (b), length of the season (g), day of pick (e, time), rate of greening (, between a and c), rate of senescencing (, between d and b), cumulative green (h), pick green (e, NDVI), and average green. All of these parameters are shown below [Jonsson].



Go Back

AVHRR missing days

Year	Missing Days
1981	177, 178, 182-201
1982	22, 88, 104-107, 114, 119-121, 187, 202, 237, 268, 269
1983	218
1984	14, 15, 51, 53, 62, 82, 101, 107, 205, 341, 342, 366*
1985	1, 2, 18, 19, 39, 40, 41, 42, 70, 310
1986	38, 73, 74, 247-365
1987	
1988	4, 72, 73, 81, 90, 135, 136, 170, 197-199, 206-208, 235, 262, 281, 313-315, 335
1989	80, 81, 96
1990	1, 3, 59, 201-205, 210-213, 307, 321
1991	1-4, 10-14, 41-43, 262-365
1992	50, 213-365
1993	90, 213-365
1994	11, 257-365
1995	
1996	121-128
1997	285-365
1998	
1999	1, 287, 288

Go Back

SPOT missing days

Year	Missing Days
1998	0-90
1999	70, 199-365
2000	
2001	1, 2, 303, 304
2002	29, 80, 133, 250, 332, 337-365

Go Back

MODIS missing days

Terra	
Year	Missing days
2000	0-54, 117-118, 219-230, 342-366
2001	167-182, 238, 239, 267-294
2002	79-86, 105, 253-265, 292
2003	351-357
2004	
2005	
2006	301
2007	33, 316, 317
2008	356, 357
2009	219, 329
2010	52, 118-151, 182, 249, 359-365

Aqua	
Year	Missing days
2002	1-184, 211-219, 256
2003	
2004	
2005	
2006	186
2007	18, 158, 336
2008	
2009	
2010	360-365