Discussion on the protocol for inter-comparison of VI datasets

CEOS WGCV, Subgroup on Land Products Validation

Focus Area: Vegetation Indices

ESA Living Planet Symposium 13-17 May 2019, Milan

Establishing a CEOS WGCV LPV protocol for the inter-comparison of Vegetation Index datasets

Else Swinnen[1], Toté Carolien[1], Tomoaki Miura[2], Miguel Román[3], Fernando Camacho[4], Jamie Nickeson[3]

[1]VITO Remote Sensing, Belgium
[2]University of Hawaii/NOAA
[3]NASA-GSFC
[4]EOLAB, Spain

Vegetation indices are one of the most widely-used satellite products for studies involving vegetation dynamics and vegetation anomaly monitoring. The Vegetation Index (VI) focus area was established as a new focus area within the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration/Validation (WGCV) Land Product Validation (LPVC) Subgroup in September 2016. The goal of this VI focus area is to establish a set of community-wide, peer-reviewed protocol recommendations for validating and inter-comparing satellite VI products.

The first workshop of the VI focus area was held in conjunction with the CEOS LPV Land Surface Phenology focus area's workshop in Fort Collins, Colorado, United States on 9-10 November 2016. At the workshop, the following components were identified as those that define VI uncertainty and, thus, that satisfy the user needs: (1) uncertainty of VIs in their units; (2) characterization of VI value changes with respect to changes in actual vegetation condition (biophysical and/or physiological); (3) long-

Scientific questions to be answered

- What is the overall **similarity**?
- What is the **magnitude** of the difference?
- What is the **spatial pattern** of the difference?
- What is the **temporal pattern** of the difference?
- What are the (possible) **causes** of the difference?
- What is the **impact on VI anomalies**?
- …?

Inter-comparison approach

- VI quality is related to quality of surface reflectance bands from which they are derived
- Factors contributing to differences
  - Absolute calibration
  - Sensor characteristics: relative spectral response, scanning systems, IFOV
  - Processing choices: atmospheric correction, bad pixel screening, angular normalization,...
- Focus of inter-comparison
  - Understand the differences
  - Evaluate the spatial and temporal stability of these differences

Inter-comparison criteria

- **Product completeness**
  spatial distribution and temporal evolution, length of gaps,…
- **Spatial consistency**
  absence of artefacts, spatial distribution of inter-comparison metrics,…
- **Statistical consistency**
  evaluation of bias, geometric mean regression,…
- **Temporal consistency**
  temporal profiles over point locations, temporal variation of inter-comparison metrics, temporal noise, inter-annual stability, VI anomaly analysis,…

Sample selection and stratification (1/2)

- SM / QA interpretation, uncertainties, …
- Temporal frequency and compositing methods
- Identical day of observation
- Observation and illumination geometry (in case of VI derived from directional surface reflectances): near-nadir, no mixed scatter,…
- Common grid, systematic spatial subsampling
- …

Sample selection and stratification (2/2)

- BELMANIP2 and DIRECT sites
- Other?

- Latitude bands
- Differentiation per biome
- Other?

- **Should we define a common set?**
- Global representativeness!

Metrics for inter-comparison (1/4)

- What are the best metrics?
  - Product completeness
  - Spatial consistency
  - Statistical consistency
  - Temporal consistency

- **Please share relevant papers / reports**

Metrics for inter-comparison (2/4)

- Geometric Mean regression / Orthogonal Distance Regression
- R^2, slope, intercept
- Root Mean Squared Difference (RMSD) or Uncertainty (U)

$$RMSD = U = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2}$$

  - Root of Mean Unsystematic (or random) Product Difference (RMPDu)
  - Root of Mean Systematic Product Difference (RMPDs)

$$RMPD_u = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(|X_i - \hat{X}_i|)(|Y_i - \hat{Y}_i|)} \qquad RMPD_s = \sqrt{MSD - MPD_u}$$

Metrics for inter-comparison (3/4)

- Mean Bias Error (MBE) or Accuracy (A) or average bias

$$MBE = A = \frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i) = \bar{X} - \bar{Y}$$

- Precision (P) or repeatability
  [Vermote & Kotchenova, 2008]

$$P = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - Y_i - A)^2}$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i|$$

Metrics for inter-comparison (4/4)

- Temporal smoothness
  [Vermote et al. 2009]

$$\delta(d_n) = \left| P(d_{n+1}) - P(d_n) - \frac{P(d_n) - P(d_{n+2})}{d_n - d_{n+2}}(d_n - d_{n+1}) \right) \right|$$

- Time series noise
  [Vermote et al. 2009]

$$Noise = \sqrt{\frac{\sum_{n=1}^{N-2} \delta(d_n)^2}{N-2}}$$

- Relative noise
  [Claverie et al, 2013]

$$RelativeNoise = 100 \times \frac{Noise}{\bar{X}}$$

- Inter-annual precision: anomalies of upper and lower percentiles from year to year
  [Fernandes et al. 2014]

Inter-comparison of VI anomalies

- Class agreement
  [Meroni et al. 2016]
- Heidke Skill Score (e.g. detection of drought events)
- ?

Additional points discussed

Product providers should make Quality Assessment Reports available, with basic statistics calculated (cfr. protocol)

NRT filtered products comparison / request