Title: Comparison of the new LTDR AVHRR reflectance with the GIMMS AVHRR NDVI dataset

Article Type: Short Communication

Keywords: AVHRR NDVI
Long term data records
GIMMS

Corresponding Author: Dr. Molly Elizabeth Brown, PhD

Corresponding Author's Institution: NASA-Goddard Space Flight Center

First Author: Molly Elizabeth Brown, PhD

Order of Authors: Molly Elizabeth Brown, PhD; Michael T Marshall, M.S.
Comparison of the new LTDR AVHRR reflectance with the GIMMS AVHRR NDVI dataset

Short Communication for RSE

Authors: M.E. Brown, M.T. Marshall

Abstract

The GIMMS Advanced Very High Resolution Radiometer (AVHRR) NDVI dataset has nearly three decades of observations and is the most widely used vegetation index dataset. In this short communication, we compare AVHRR normalized difference vegetation index (NDVI) data from the GIMMS group with NDVI created from the Long Term Data Record (LTDR) daily reflectance. The two datasets have substantially different processing characteristics, with the LTDR offering daily channel data, higher resolution and more complete atmospheric correction than the reflectance from which the GIMMS data were derived, but the data ends in 1999. Understanding and characterizing the impact of the differences between the GIMMS and LTDR datasets is a first step in being able to exploit the temporal and spatial resolutions offered by the new LTDR AVHRR dataset. We found that although there are significant differences between the datasets, they are broadly comparable during the 1981-1999 period. The LTDR AVHRR dataset will provide a solid foundation for future AVHRR datasets once it has been extended through to the present.
1.0 Introduction

The AVHRR NDVI is one of the few consistent sources of observations that provide direct, daily and continuous measurements of the ground that can be processed into datasets of sufficient quality for global change analysis [Brown, et al., 2006]. These measurements have increased in importance for a variety of communities as the record has grown in length. The data are now being used in a wide variety of scientific [Neigh, et al., 2008], modeling [Yang, et al., 2006; Zhou, et al., 2003] and operational [Brown and De Beurs, 2008; Tarnavsky, et al., 2008; van Leeuwen, et al., 2006] settings to estimate how ecosystems have changed in the past three decades.

The AVHRR record requires substantial processing as the observations are contaminated by a wide variety of atmospheric, aerosol and radiometric effects. The dataset most frequently used in the scientific literature is the AVHRR NDVI dataset produced by the Global Inventory Monitoring and Mapping (GIMMS) group, described in Tucker et al. (2003). The dataset consists of 15-day composites that begin in July 1981 and go through the present, incorporating data from AVHRR/2 sensors on satellites NOAA-07, 09, 11, 14 and AVHRR/3 sensors on satellites NOAA-16, 17 and 18.

The version 2 Long Term Data Record (LTDR) data, described in [Nagol, et al., 2008; Pedelty, et al., 2007] is available online at http://ltdr.nascom.nasa.gov/ltdr/ltdr.html. This new dataset has a few advantages over the GIMMS data, including a comprehensive atmospheric correction that is applied to daily data and a higher spatial resolution. The
LTDR dataset does have a data gap for the six months in 1994 where the GIMMS data uses
NOAA-9. In addition, the LTDR dataset only runs until 1999, as it does not include data
from NOAA-17 or NOAA-18. This reduces its utility over GIMMS long-term studies until it
can be extended. Here we compare the GIMMS NDVI with NDVI calculated from LTDR
reflectance for the overlapping period to determine how similar the two datasets are, as
well as their compatibility. Although a diversity of data sources are useful for the
community, understanding how these two datasets compare is important for future
investments.

2.0 Data and Methods

Two AVHRR NDVI datasets are used in this study: the LTDR and the GIMMS data. Here each
dataset will be described along with the processing and analysis approach. The version 2
LTDR data used here are a five channel, daily reflectance dataset provided at a global CMG
geographic 0.05 degree resolution. Using new methods developed from their experience
with MODIS data [Nagol, et al., 2008], improved methods were developed and implemented
to correct the dataset for Rayleigh scattering, ozone and aerosol contamination [Nagol, et
al., 2008].

The LTDR data was aggregated into 15-day maximum value composites (day 1-15 and day
16-end of the month) for comparison purposes with the GIMMS dataset. A channel 5
temperature (< 0 °C) mask was applied to reduce interference from clouds. Because the
LTDR data are mapped daily, all available data is provided, including data at the extremes
of view angle (Figure 1). To composite data to match the GIMMS data, only data with a
view angle less than 42 degrees off nadir are used in the composite.

A weighted least squares temporal moving filter described in Swets et al. (1999) was
applied to the LTDR dataset and compared to the GIMMS dataset. The smoother calculates
the linear least squares regression between points bounded by a default window. Weights
are assigned based on magnitude—peaks are weighted more heavily than troughs.
Preferential weighting is chosen, because noise generally acts to lower NDVI. The values
determined from the lines at a given point are then combined using a simple average.
Values that exceed a maximum threshold and fail the chi squared test with two degrees of
freedom are considered outliers and subsequently omitted from the process.

The GIMMS group NDVI data are at 8km resolution and in a global composite of 0.07
degrees. Only the NDVI data are corrected, with the reflectance data uncorrected and
available only during certain periods and in certain areas. The GIMMS data have had the
effects of solar zenith angle removed for the tropics and the AVHRR/2 and AVHRR/3 data
have been integrated into one dataset through the use of SPOT Vegetation and the
empirical mode decomposition and reconstruction technique [Pinzon, 2002].

We selected the same sites as were used in the Brown et al (2006) paper to compare the
two datasets globally. A time series for each dataset was extracted at 25 km² resolution
(Table 1). We present SeaWiFS land [SeaWiFS-Land, 2004], and MODIS climate model grid
monthly MOD13 [Huete, et al., 2002] time series for each site for further comparison with
the LTDR dataset. The mean and standard deviation over all months for each dataset was used to characterize a site. The global maximum NDVI from 1981-1999 was calculated to compare the two datasets and identify differences spatially.

3.0 Results

Time series plots show a broad similarity in the LTDR and GIMMS AVHRR datasets (Figure 2). Overall, the GIMMS NDVI data has higher values than the LTDR, as can be seen in Table 1, and is therefore much closer to the NDVI values of SeaWifs and MODIS. The correlations between the AVHRR GIMMS and LTDR anomaly time series were higher than those between the GIMMS dataset and the other sensors. In some cases, there was essentially no relationship between the GIMMS anomaly and the other sensor anomaly, but a fairly high correlation between the GIMMS and LTDR datasets (the Ji Parana site in Brazil, for example).

Globally, Figure 3 shows the maximum NDVI from 1981-1999 for the GIMMS (top) and the LTDR (bottom) datasets. Figure 4 provides a histogram of the two images, showing that the highly vegetated pixels have an NDVI value of around 0.9 for the GIMMS and 0.7 for the LTDR. At the lower end, the GIMMS data is slightly lower than the LTDR for sparsely vegetated regions.

4.0 Discussion
The GIMMS group NDVI has become one of the primary long term data records [Eklundh and Olsson, 2003] which requires direct observations for scientific studies from multiple disciplines [Anyamba, et al., 2001; Slayback, et al., 2003; Stow, et al., 2003; Suzuki, et al., 2003] and for modeling and mapping efforts. This has been because it is one of the few sources of global measurements that are available from 1981 to a year or so behind the present day which are consistent, comparable and statistically robust across the whole AVHRR record.

The AVHRR/2 to AVHRR/3 design change introduced in 2000 with the flying of NOAA-16 consisted of several changes that make the incorporation of the data from before 1999 with that after non-trivial. These include

- embedded timing and global positioning coordinates in each scan, greatly improving the geolocation of the mapped data
- on-board satellite stabilization which removes the problem of a precessing orbit which afflicts the data from the AVHRR/2 series;
- changes in the spectral reflectance functions which would, without correction, increase the NDVI over what was seen in the AVHRR/2 data;
- A dual gain is also introduced in Channels 1 and 2 data that affect the response of the lower reflectance regions in comparison with the upper ranges.

The GIMMS data uses the empirical mode decomposition (EMD) and reconstruction technique [Huang, et al., 1998] to incorporate the AVHRR/2 with the AVHRR/3 data [Pinzon, 2002]. Here we compare the GIMMS data from the AVHRR/2 series with that from
139 the LTDR. The LTDR reflectance dataset does not attempt to combine the information from
140 the two series of instruments, but does have a comprehensive atmospheric correction,
141 daily time step and a higher resolution than the GIMMS data. Here we find that although
142 there are significant differences between the two datasets, which can be attributed both to
143 differences in processing as well as the atmospheric correction applied to data, they still
144 remain to be smaller than the differences among NDVI datasets from different sensors.
145
146 The effect of the EMD processing technique is to bring the two AVHRR series of
147 instruments together in the same time series, and to match the resulting AVHRR dataset to
148 the levels seen in other widely-used vegetation datasets such as from MODIS. Here we
149 show that the GIMMS AVHRR NDVI is comparable to the LTDR dataset, despite its lack of a
150 comprehensive atmospheric correction of the reflectance. The post-processing empirical
151 approach used by the GIMMS group results in the removal of spurious noise and not signal,
152 resulting in a similar end product as the more traditional atmospheric correction process
153 used in the LTDR. Although the results show that the contamination of the AVHRR signal
154 by aerosols and clouds is still much higher than the other two sensors, overall the two
155 AVHRR datasets are comparable to other global datasets at similar resolutions.
156
157
158 5.0 Conclusions
159
160 The new LTDR dataset poses new opportunities for the remote sensing community, as it
161 provides a daily global reflectance product with a far more comprehensive correction for
aerosols and atmospheric effects than what was available with the AVHRR Pathfinder Land (PAL) product [James and Kalluri, 1994]. Although the dataset is currently not available past 1999, it has the potential of being a strong resource for land observations. To understand its impact, we have here compared it to the widely used AVHRR GIMMS 15 day dataset, and to NDVI datasets from MODIS and SeaWiFS.

We conclude that the LTDR data has a lower range and requires significant post-processing to be used as a vegetation index dataset, specifically creating the NDVI, compositing the daily data and smoothing of the composit ed files. The data also is missing six months of data in 1994 and does not extend past 1999. It has potential, however, of being a good intermediate step to a new higher resolution, shorter compositing period AVHRR NDVI dataset in the coming years.

References


SeaWIFS-Land (2004), SeaWIFS land Data organized by Data Product groups, edited, GES Distributed Active Archive Center.


Table 1. Data comparison from 21 sites around the world, showing the relationship between the monthly NDVI anomalies of the GIMMS AVHRR (G) compared to the LTDR AVHRR NDVI (L), SeaWiFS (SW) and MODIS (MO) anomaly time series at a monthly timestep. Also reported are the mean NDVI for each site and the standard deviation of the anomaly for the period.

Figure 1. Image of July 1, 1997 reflectance of Channel 2 from the LTDR data web site. Notice the nearly complete coverage at 0.05 degree by using data at the edge of the AVHRR swath and a continuous mapping approach. Regions with very off-nadir data show as lighter diagonal stripes.

Figure 2. Time series of X site showing the variability of GIMMS, LTDR, SeaWiFS, 10-day SPOT Vegetation and monthly MODIS CMG data. The Landsat NDVI averages are plotted (circles) and were created from atmospherically corrected NDVI subsets around the study sites. All data are the average of a 25 km² window. The top panel shows the entire 1981-2008 record and the bottom shows the overlapping period of 1997-2004.

Figure 3. Global AVHRR NDVI from LTDR and GIMMS showing a comparison of the maximum NDVI for each pixel throughout the record.

Figure 4. Histogram of the data from the maximum NDVI images shown in Figure 3.
Table 1.

<table>
<thead>
<tr>
<th>Site Name</th>
<th>lat</th>
<th>long</th>
<th>Land Cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia NW</td>
<td>-31.1</td>
<td>-150.7</td>
<td>grasses</td>
</tr>
<tr>
<td>Barrow</td>
<td>71.3</td>
<td>-156.7</td>
<td>closed brushlands/shrubland</td>
</tr>
<tr>
<td>Bondville</td>
<td>40.1</td>
<td>-88.4</td>
<td>croplands</td>
</tr>
<tr>
<td>Cascades-1</td>
<td>44.2</td>
<td>-122.3</td>
<td>evergreen needleleaf forest</td>
</tr>
<tr>
<td>Crystal_City</td>
<td>28.7</td>
<td>-99.5</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>E.Longreach.AU</td>
<td>23.0</td>
<td>145.5</td>
<td>closed brushlands/shrubland</td>
</tr>
<tr>
<td>Gujarat</td>
<td>22.0</td>
<td>70.0</td>
<td>open shrublands</td>
</tr>
<tr>
<td>Harvard</td>
<td>42.5</td>
<td>-72.2</td>
<td>mixed forests</td>
</tr>
<tr>
<td>Ji-Parana</td>
<td>-10.8</td>
<td>-82.4</td>
<td>woodlands</td>
</tr>
<tr>
<td>Kozza</td>
<td>30.1</td>
<td>-96.6</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Louga.Senegal</td>
<td>18.0</td>
<td>-16.0</td>
<td>bare</td>
</tr>
<tr>
<td>Lyon.France</td>
<td>46.0</td>
<td>5.0</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Mongu</td>
<td>-15.3</td>
<td>23.2</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Parkfall_Wisc-1</td>
<td>45.9</td>
<td>-90.3</td>
<td>mixed forests</td>
</tr>
<tr>
<td>Rajasthan.india</td>
<td>26.0</td>
<td>74.5</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>50.0</td>
<td>-110.0</td>
<td>grasses</td>
</tr>
<tr>
<td>Sevilleta</td>
<td>34.4</td>
<td>-106.9</td>
<td>open shrublands</td>
</tr>
<tr>
<td>Skukuza</td>
<td>-25.0</td>
<td>31.6</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Tapajos</td>
<td>-2.4</td>
<td>-54.6</td>
<td>woodlands</td>
</tr>
<tr>
<td>USDAARS-BAI</td>
<td>39.0</td>
<td>-76.9</td>
<td>wooded grasslands/shrubs</td>
</tr>
<tr>
<td>Walker</td>
<td>36.0</td>
<td>-84.3</td>
<td>mixed forests</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Anomaly Correlations</th>
<th>G-L</th>
<th>G-SW</th>
<th>G-MO</th>
<th>Mean NDVI * 1000</th>
<th>Standard Deviation of Anomaly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=222)</td>
<td>(n=36)</td>
<td>(n=56)</td>
<td>G</td>
<td>L</td>
</tr>
<tr>
<td>0.62</td>
<td>0.63</td>
<td>0.70</td>
<td>534</td>
<td>473</td>
<td>496</td>
</tr>
<tr>
<td>0.03</td>
<td>0.02</td>
<td></td>
<td>184</td>
<td>293</td>
<td>111</td>
</tr>
<tr>
<td>0.13</td>
<td>0.01</td>
<td></td>
<td>432</td>
<td>396</td>
<td>399</td>
</tr>
<tr>
<td>0.10</td>
<td>0.31</td>
<td></td>
<td>679</td>
<td>662</td>
<td>766</td>
</tr>
<tr>
<td>0.48</td>
<td>0.34</td>
<td></td>
<td>404</td>
<td>368</td>
<td>387</td>
</tr>
<tr>
<td>0.47</td>
<td>0.39</td>
<td></td>
<td>346</td>
<td>292</td>
<td>275</td>
</tr>
<tr>
<td>0.26</td>
<td>0.60</td>
<td></td>
<td>265</td>
<td>242</td>
<td>237</td>
</tr>
<tr>
<td>0.04</td>
<td>0.10</td>
<td></td>
<td>652</td>
<td>579</td>
<td>696</td>
</tr>
<tr>
<td>0.02</td>
<td>0.02</td>
<td></td>
<td>647</td>
<td>590</td>
<td>642</td>
</tr>
<tr>
<td>0.15</td>
<td>0.57</td>
<td></td>
<td>466</td>
<td>438</td>
<td>457</td>
</tr>
<tr>
<td>0.63</td>
<td>0.60</td>
<td></td>
<td>218</td>
<td>196</td>
<td>204</td>
</tr>
<tr>
<td>0.08</td>
<td>0.08</td>
<td></td>
<td>553</td>
<td>493</td>
<td>543</td>
</tr>
<tr>
<td>0.31</td>
<td>0.31</td>
<td></td>
<td>497</td>
<td>437</td>
<td>441</td>
</tr>
<tr>
<td>0.09</td>
<td>0.09</td>
<td></td>
<td>522</td>
<td>465</td>
<td>540</td>
</tr>
<tr>
<td>0.08</td>
<td>0.58</td>
<td></td>
<td>260</td>
<td>202</td>
<td>248</td>
</tr>
<tr>
<td>0.43</td>
<td>0.68</td>
<td></td>
<td>253</td>
<td>272</td>
<td>249</td>
</tr>
<tr>
<td>0.29</td>
<td>0.44</td>
<td></td>
<td>193</td>
<td>185</td>
<td>186</td>
</tr>
<tr>
<td>0.23</td>
<td>0.53</td>
<td></td>
<td>495</td>
<td>432</td>
<td>452</td>
</tr>
<tr>
<td>0.29</td>
<td>0.25</td>
<td></td>
<td>389</td>
<td>423</td>
<td>552</td>
</tr>
<tr>
<td>0.00</td>
<td>0.25</td>
<td></td>
<td>533</td>
<td>499</td>
<td>551</td>
</tr>
<tr>
<td>0.07</td>
<td>0.10</td>
<td></td>
<td>623</td>
<td>574</td>
<td>640</td>
</tr>
</tbody>
</table>
Figure 1.

Figure 2.
Figure 3.
Figure 4.