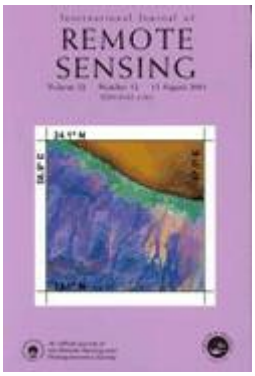

Neural Networks as a Tool for Constructing Continuous NDVI Time Series from AVHRR and MODIS

Journal:	*International Journal of Remote Sensing*
Manuscript ID:	TRES-PAP-2007-0740.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	14-May-2008
Complete List of Authors:	Brown, Molly; SSAI, NASA-Goddard Space Flight Center Lary, David; University of Maryland Baltimore College, GEST Vrieling, Anton; Joint Research Centre of the European Commission Stathakis, Demetris; Joint Research Centre of the European Commission Mussa, Hamse; University of Cambridge
Keywords:	AVHRR, MODIS, NDVI, NEURAL NETWORKS
Keywords (user defined):	AVHRR, MODIS, NDVI

Neural Networks as a Tool for Constructing Continuous NDVI Time Series from AVHRR and MODIS

Molly E. Brown[1]

David J. Lary[2]

Anton Vrieling[3]

Demetris Stathakis[3]

Hamse Mussa[4]

1 Science Systems and Applications, Inc., NASA Goddard Space Flight Center, MD, USA
Ph: 301-614-6616, Fax: 301-614-6015
Email: molly.brown@gsfc.nasa.gov

2 UMBC GEST, NASA Goddard Space Flight C, MD, USA

3 Joint Research Centre of the European Commission, Ispra (VA), Italy

4 Department of Chemistry, University of Cambridge, England

Revised, for International Journal of Remote Sensing

Abstract

The long term AVHRR-NDVI record provides a critical historical perspective on vegetation dynamics necessary for global change research. Despite the proliferation of new sources of global, moderate resolution vegetation datasets, the remote sensing community is still struggling to create datasets derived from multiple sensors that allow the simultaneous use of spectral vegetation for time series analysis. To overcome the non-stationary aspect of NDVI, we use an artificial neural network (ANN) to map the NDVI indices from AVHRR to those from MODIS using atmospheric, surface type and sensor-specific inputs to account for the differences between the sensors. The NDVI dynamics and range of MODIS NDVI data at one degree is matched and extended through the AVHRR record. Four years of overlap between the two sensors is used to train a neural network to remove atmospheric and sensor specific effects on the AVHRR NDVI. In this paper, we present the resulting continuous dataset, its relationship to MODIS data, and a validation of the product.

Keywords: Normalized difference vegetation index (NDVI), MODIS, AVHRR, Neural Networks

1.0 Introduction

Consistent, long term vegetation data records are critical for analysis of the impact of global change on terrestrial ecosystems. Continuous observations of terrestrial ecosystems through time are necessary to document changes in magnitude or variability in an ecosystem (Eklundh and Olsson, 2003; Slayback et al., 2003; Tucker et al., 2001). Satellite remote sensing has been the primary tool for scientists to measure global trends in vegetation, as the measurements are both global and temporally frequent. To extend measurements through time, multiple sensors with different design and resolution must be used together in the same time series. This presents significant problems as sensor band placement, spectral response, processing, and atmospheric correction of the observations can vary significantly and impact the comparability of the measurements (Brown et al., 2006). Even without differences in atmospheric correction, vegetation index values for the same target recorded under identical conditions will not be directly comparable because input reflectance values differ from sensor to sensor due to differences in sensor design and spectral response of the instrument (Miura et al., 2006; Teillet et al., 1997).

Several approaches have been taken to integrate data from multiple sensors. Steven et al. (2003), for example, simulated the spectral response from multiple instruments and with simple linear equations created conversion coefficients to transform NDVI data from one sensor to another. Their analysis is based on the observation that the vegetation index is critically dependent on the spectral response functions of the instrument used to calculate it. The conversion formulas the paper presents cannot be applied to maximum value NDVI datasets because the weighting coefficients are land cover and dataset dependent, reducing their efficacy in mixed pixel situations

(Steven et al., 2003). Trishchenko et al. (2002) created a series of quadratic functions to correct for differences in the reflectance and NDVI to NOAA-9 AVHRR-equivalents (Trishchenko et al., 2002). Both the Steven et al. (2003) and the Trishchenko et al. (2002) approaches are land cover and dataset dependent and thus cannot be used on global datasets where multiple land covers are represented by one pixel. Miura et al (2006) used hyper-spectral data to investigate the effect of different spectral response characteristics between MODIS and AVHRR instruments on both the reflectance and NDVI data, showing that the precise characteristics of the spectral response had a large effect on the resulting vegetation index. The complex patterns and dependencies on spectral band functions were both land cover dependent and strongly non-linear, thus we see that an exploration of a non-linear approach may be fruitful.

In this paper we experiment with powerful, non-linear neural networks to identify and remove differences in sensor design and variable atmospheric contamination from the AVHRR NDVI record in order to match the range and variance of MODIS NDVI without removing the desired signal representing the underlying vegetation dynamics. Neural networks are 'data transformers' (Atkinson and Tatnall, 1997), where the objective is to associate the elements of one set of data to the elements in another. Relationships between the two datasets can be complex and the two datasets may have different statistical distributions. In addition, neural networks incorporate a priori knowledge and realistic physical constraints into the analysis, enabling a transformation from one dataset into another through a set of weighting functions (Atkinson and Tatnall, 1997). This transformation incorporates additional input data that may account for differences between the two datasets.

Our objective in this paper is to demonstrate the viability of neural networks as a tool to produce a long term dataset based on AVHRR NDVI that has the data range and statistical distribution of MODIS NDVI. Previous work has shown that the relationship between AVHRR and MODIS NDVI is complex and nonlinear (Brown et al., 2006; Gallo et al., 2003; Miura et al., 2006), thus this problem is well suited to neural networks if appropriate inputs can be found. The impact of atmospheric contamination, such as clouds, smoke, pollution and other aerosols, variations in soil color and exposure through vegetation, and land cover type has a differential effect on AVHRR data as compared to MODIS data. Here we explore how neural networks can be used to account for these impacts and create an AVHRR NDVI dataset with similar characteristics as the MODIS dataset. Overlapping years of observations are used to train the network. Examination of the resulting MODIS-fitted AVHRR dataset both during the overlap period and in the historical dataset enabled an evaluation of the efficacy of the neural net approach compared to other approaches to merge multiple-sensor NDVI datasets.

2.0 Neural Networks

Neural networks are algorithms used for either classification or function approximation (Lippmann, 1987). A good introduction on neural networks is given by Lippmann (1987). Since their first introduction, they have been used for almost two decades in remote sensing (Benediktsson et al., 1990). The most commonly used type of neural network is the Multi-Layer Perceptron, of which Kalman filters are one type. Artificial neural networks (ANN) are made up of input layers, hidden layers and output layers.

The MLP neural network has an input layer where the data samples are fed, typically after being normalized. The data from the input layer is then fed into a number of hidden layers, typically either one or two. The choice of how many hidden layers and number of nodes per hidden layer that should be used is currently an open research question (Stathakis, 2008). Several heuristics exist to assist in selecting the number of nodes in the hidden layers, some of which developed explicitly in the domain of remote sensing such as the Kanellopoulos – Wilkinson (1997) rule (Stathakis and Vasilakos, 2006). Finally the hidden layers feed one or more input layers.

To summarize the ANN topology, a relation of *x*:*y*:*z* is frequently used. This implies a neural network with *x* input nodes, one hidden layer with *y* hidden nodes and *z* output nodes (for example, 7:20:1). The neural network is trained by adjusting the values of the connections, called weights, between nodes. The most commonly used training algorithm is back-propagation introduced by Rumelhart et al. (1986). Several modifications to the original algorithm have greatly boosted performance (Rumelhart et al., 1986). Neural networks can learn in an either supervised or unsupervised mode depending on whether target vectors are presented along with input vectors or not. In the supervised mode, several spectral bands (or in this study, time series) per data sample are typically presented to the network. At the same time the desired output is also used to modify the weights so that the deviation between actual and obtained output is minimized. Typically the samples available, i.e. input and output vectors, are split in order to train the network and independently validate the results. A three-set strategy has been proposed to offer a more objective validation by Bishop (1995). According to this strategy three subsets are created, one of training, one for validation and on for testing (Bishop, 1995).

One of the main advantages of neural networks is the fact that multiple sources, including non-spectral, data can be used as input (Benediktsson et al., 1990; Stathakis and Kanellopoulos, 2008). This is because neural networks make no assumptions, e.g. about statistical distributions, regarding the input data. One of their main drawbacks is that they require experience in selecting values for the numerous parameters that need to be set. Recent results show that global search methods can be used to make near-optimal choices (Stathakis, 2008). Additionally, neural networks are often accused of being black-box techniques because the knowledge learned can not be expressed in a meaningful way. Several efforts have been made towards building transparent neural networks. One way to do this is to deploy neuro-fuzzy methods (Stathakis and Vasilakos, 2006).

3.0 Data

This study uses global NDVI products derived from AVHRR and MODIS NDVI sensors at one degree resolution and for a monthly time window. Ancillary files are used in this study to determine the impact of clouds and other atmospheric effects on the vegetation measurement from different sensors through time. We have restricted the number of inputs to six besides the AVHRR NDVI to reduce redundancy and over-fitting of the neural network. These are three atmospheric products from TOMS, a soil type map, a digital elevation model (DEM), and a land cover map.

3.1 NDVI datasets at one degree

AVHRR and MODIS NDVI products were downsampled to one degree resolution to reduce processing time of the artificial neural network and to match the resolution of the atmospheric

TOMS inputs. To further reduce processing time, average monthly composites were made of the two products. The spatial and temporal downsampling was done by averaging all pixels falling in a one-degree cell for the two nearest periods in a month (MODIS products do not respect month limits).

The maximum value AVHRR NDVI composites have an 8-km resolution (Holben, 1986; Tucker, 1979) and were from the NASA Global Inventory Monitoring and Modeling Systems (GIMMS) group at the Laboratory for Terrestrial Physics (Brown et al., 2006; Tucker et al., 2005) from July 1981 to May 2004. A post-processing satellite drift correction has been applied to this dataset to further remove artifacts due to orbital drift and changes in the sun-target-sensor geometry (Pinzon et al., 2005). As a result of AVHRR's wide spectral bands, the AVHRR NDVI is more sensitive to water vapor in the atmosphere than MODIS. An increase in water vapor results in a lower NDVI signal, which can be interpreted as an actual change if no correction is applied (Pinheiro et al., 2004; Pinzon, 2002). The maximum value composite should lessen these artifacts (Holben, 1986). The GIMMS operational dataset incorporates AVHRR data from sensors aboard NOAA-7 through 14 with the data from the AVHRR on NOAA-16 and 17.

The Terra-MODIS 16 day L3 land surface NDVI product was selected. NDVI data for MODIS was computed from the (White-Sky) Filled Land Surface Albedo Map Product, which is a value-added product from the MODIS Atmospheres group. The global, one kilometer, 16 day MODIS NDVI composites from February 2000 to December 2004 were used to create averaged one degree monthly data for this analysis. The resulting one degree time series include only pixels with more

than 50% land and conforms to the ISCLSCP convention as described by Sellers et al. (1996).

3.2 Ancillary datasets

To account for the differences between the AVHRR and MODIS data, we use four ancillary data products in the neural network: TOMS Data which provides information on water vapor in the atmosphere, soil maps, land cover maps and elevation. Each of these accounts for an aspect of the sensor design differences and provide key information so that the neural network can work. Preliminary work (not described here) demonstrated that the most important factors controlling the relationship between the NDVI of MODIS and that of AVHRR are the surface reflectance, the land surface type, aerosols and total ozone column. Variations in atmospheric contamination have direct impact on the AVHRR NDVI used here because no atmospheric correction was implemented during its processing, only volcanic aerosols and maximum value compositing (Tucker et al., 2005). We know that ozone is a key atmospheric absorber of light in the visible region, and water, as measured by aerosols, in the infrared. The AVHRR NDVI, calculated using the wide bands of the instrument, will therefore be influenced by these elements.

The Nimbus-7 TOMS data is the only source of high resolution global information about the atmospheric composition (and hence depression of AVHRR NDVI) for much of the AVHRR record. As an instrument that measures the atmosphere back to 1981, TOMS has the advantage of being co-located for much of its record on the same platform as AVHRR, which is particularly important as the NOAA satellites from which the AVHRR NDVI are derived are subject to non-linear orbital drift through time (McPeters et al., 1998). The TOMS data is from Version 8, includes reflectance, aerosols and ozone measurements and is derived from three sensors: Nimbus 7, Meteor

3 and Earth Probe (Table 1). All three products are used in order to capture the impact of atmospheric variations on the uncorrected AVHRR NDVI data. During the missing period of 1994-96, we use a climatology created by taking the median value of the preceding 2, 4, and 6 years and the following 2, 4, and 6 years. This approach was used as ozone has a quasi-biennial oscillation (QBO). Although not optimal, this performed well and is required if we want to use these datasets for a correction of the entire series.

The NASA Goddard Institute for Space Studies (GISS) soil type map is used to account for the difference in sensitivity to underlying soil color from AVHRR and MODIS (Huete et al., 1994; Huete and Tucker, 1991). The soil type map is at one degree resolution and contains 26 soil units, and values for water and ice. The soil type data file was derived from the highest level of the FAO soil units and is based on the work of Zobler (1986).

A one degree 'surface type' land cover dataset was created from the SPOT Global Land Cover (GLC) 2000 dataset (Giri et al., 2004). Previous research has shown that variations in land cover affect the strength of the impact of atmospheric thickness (Pinzon, 2002). This dataset has 22 land cover classes based on the FAO land cover classification system. We aggregated the data to a one-degree resolution using a vote procedure. We used the GLC2000 data instead of MODIS or AVHRR-based land cover datasets as an independent surface classification for the ANN training. We use a single land cover map to represent the land cover for the 25 year record. Even though we acknowledge that land cover change may have occurred during this period, they are unlikely to span an entire one by one degree pixel. The neural network uses this parameter to identify regions

with very low signal due to small amounts of vegetation. These regions are approximately static through time globally.

A one degree DEM was used to ensure the identification and maintenance of mountainous regions that may otherwise be confused with clouds or other atmospheric effects. This DEM was derived from the USGS SRTM 90-m dataset, and has been aggregated to one degree using averaging.

3.3 Global Rainfall Data

We used Global Precipitation Climatology Centre (GPCC) rain gauge data from the Global Precipitation Climatology Project (GPCP). These data were used to evaluate the ability of the NDVI data products for capturing interannual vegetation dynamics related to rainfall. The GPCC data are area-averaged and time-integrated precipitation fields based on surface rain gauge measurements. The GPCC collects monthly precipitation totals received from the World Weather Watch GTS (Global Telecommunication System) of the World Meteorological Organization (WMO). The GPCC acquires monthly precipitation data from international/national meteorological and hydrological services/institutions. Surface rain-gauge based monthly precipitation data from 6700 meteorological stations are analyzed over land areas and gridded using a spatial objective analysis method (Rudolf et al., 1994).

4.0 Methods

4.1 Application of the ANN

When mapping AVHRR to MODIS NDVI using ANNs, factors that explain differences in the sensors and their processing must be accounted for by the input variables. Here we use historical data derived from the total ozone mapping spectrometer or TOMS, which is available with some interruption back to 1978 (McPeters et al., 1998). The AVHRR is also more sensitive to differences in background soil contamination than MODIS (Huete and Jackson, 1988), thus we use a soil type map (Zobler, 1986), a DEM, and a land cover map to account for these differences (see section 3 for a description of the datasets).

The neural network used here is a fully-connected feed-forward Multi-Layer Perceptron with 7:20:1 topology. Biases are connected to both hidden and output layers. The selection of the nodes in the hidden topology conforms well to the Kanellopoulos – Wilkinson rule commonly used in remote sensing. In this study we employed a feed-forward ANN with 20 nodes in a single hidden layer using a Kalman filter training algorithm. The Kalman filter algorithm provides rapid convergence for the weight estimation and is described by Lary and Mussa, (2004).

Besides the additional data sources, the neural net is trained with time-series data of AVHRR and MODIS from the overlapping period of 2000-2003. Subsequently, the resulting weighting functions were applied to the AVHRR data from 1982-2003, using the ancillary files. The functions enable the correction of the entire dataset, enabling the production of an AVHRR dataset with similar characteristics as the MODIS dataset. For simplicity, throughout this paper this new dataset will be referred to as NNndvi, or the neural net corrected AVHRR NDVI. The result is an experimental product, whose objective is to demonstrate how a seamless AVHRR to MODIS dataset may be created. We do not assume that the method used is the only possible or even the most optimal

method, but one that can produce a far closer integration between the datasets than has been demonstrated before using the actual processed data instead of modeled data. For this feasibility demonstration we operated on the one degree scale at a monthly resolution to reduce processing time of the neural net. The same training procedure could be conducted at a higher temporal and spatial resolution with more computing time and/or for smaller areas.

4.2 Evaluation Methods

The obtained NNndvi dataset is evaluated in two ways to determine if it is closer to the target MODIS NDVI than the original AVHRR dataset, and if it retains important interannual vegetation dynamics that have previously been identified in the AVHRR data (Bounoua et al., 2000; Zeng et al., 1999). First, time series for selected one degree boxes are presented to demonstrate the effect of the neural net procedure on particular locations. Second, the NNndvi is compared to the GPCC dataset to determine whether or not the correction has changed the relationship with observed rainfall.

5.0 Results

Figure 1 shows a schematic representation of the neural net mapping of the AVHRR NDVI to the MODIS NDVI during the years of overlap. Table 2 shows that the most important variable for linking the two datasets is the AVHRR NDVI (as would be expected) followed by the surface reflectance and total ozone column. In the TOMS data, the reflectance includes the degree of

cloudiness. Given the wide bands of the AVHRR sensor and the differences in processing, it is expected that the TOMS reflectance is important in the correction (Cihlar et al., 2001).

Figure 2 shows the NDVI difference between the MODIS and AVHRR, and the MODIS and the NNndvi by latitude band for a single image from December 2003. The biggest differences are in the tropics which have high concentrations of atmospheric aerosols and water vapor that interfere more with the AVHRR NDVI data than with the MODIS data (Huete et al., 2006). Another substantial difference between the datasets is seen in the northern latitudes. The histogram is from January, 2003, so the regions north of 40N have little active photosynthetic activity, the NDVI is largely measuring differences in ground cover and atmospheric thickness. The GIMMS AVHRR NDVI reports data over snow, ice, and during periods when there is no light, relying on the NDVI to correctly record the very low photosynthetic activity during these months. MODIS NDVI data incorporates much more sophisticated snow and ice detection, which results in large differences between the AVHRR and MODIS data. Because we have inputs into the neural net that can account for these differences (soil type, monthly changes in reflectivity), the differences between MODIS and AVHRR are considerably reduced by the neural network processing.

Figures 3a and 3b show the spatial average of all pixels in the same latitudinal band for the difference between the AVHRR and MODIS (3a) and NNndvi and MODIS (3b). The plots show the significant improvement in the correspondence between the datasets in the tropics and in the northern latitudes seen in Figure 2 is present in all years. Differences at the beginning and end of the growing season in the far north are clearly seen. These differences will be significant to scientists attempting to measure changes in phenology through time due to a warming climate. The

northern latitudes have experienced the largest degree of warming, thus these systematic differences are important to both recognize and remove if a consistent, sensor-independent dataset is to be developed.

The neural network process provides coefficients that were applied to the input data, to produce an NDVI fit to MODIS from AVHRR back to 1982. Figure 4 shows the zonal averages of the resulting dataset, displaying both seasonality and interannual variability as is expected. Table 3 shows the mean and standard deviation of the MODIS, AVHRR and NNndvi datasets. The mean NNndvi is closer to the MODIS data than to the original AVHRR data. The differences in the means can be seen in Figure 5, which shows the root mean square error (RMSE) in NDVI units between the AVHRR - MODIS (Figure 5A), and the NNndvi – MODIS (B). The NNndvi dataset is on average within 0.2 NDVI units of the MODIS data, removing the land-cover and regional differences that can be seen in the top panel. The scatter above 0.2 RSME are seen in the map of the RMSE in Figure 5B as being concentrated along the coastlines and where a sharp land-cover gradient is located, such as along the Himalayas and Andes mountain ranges. This is likely to be due to differences in the original land cover map between MODIS, AVHRR and TOMS and the other ancillary datasets, as well as averaging procedures to make the one degree datasets. This effect may be ameliorated by using a higher resolution, as at one degree much mixing of vegetated and non-vegetated features occurs, particularly along sharp land cover and topographic features which reduces the effectiveness of the neural network training.

Figure 6 shows the time series from MODIS, AVHRR, and the NNndvi from six selected one degree pixels (Brown et al., 2006). These locations were selected from the Earth Observing System

land validation core sites described in Brown et al (2006) and were meant to display a range of ecosystems and climates. The figure shows that the NNndvi is much closer to the MODIS series than the original GIMMS AVHRR, particularly in areas with high humidity such as in the Cascades of Washington state or Ji-Parana, Brazil. The NNndvi is higher than the GIMMS data, especially during the winter months. In some regions where the match between MODIS and AVHRR was fairly good originally, such as in the Harvard Forest, the fit between the datasets is extremely good.

Figure 7 shows the correlation coefficient, R, between the GPCC monthly gridded rainfall product at one degree and the GIMMS AVHRR, NNndvi, and MODIS from 2000-2003. The maps in the top two panels show that the NNndvi has a similar relationship with rainfall in semi-arid regions as has been documented with the GIMMS data (Brown et al., 2004). It demonstrates that at one degree, the correction maintains the datasets' basic integrity and relationship with rainfall in semi-arid zones. Panel D shows the histogram of the global correlation, showing a similar structure to the data for the three datasets.

The results of this procedure are fairly robust, but they are not sufficiently good to be used for scientific investigations. To determine if the data are usable immediately, we produced an anomaly for August 2003 from each dataset versus the four year August mean for MODIS. Figure 8 shows the histogram of the anomaly for August 2003 (when there was a major drought in Europe), which shows the improvement of the NNndvi over AVHRR, but the data is still quite a bit different than the MODIS data. Depending on the user requirements, this may be sufficiently similar. The bias in the AVHRR has been removed so that the NNndvi is far more normally distributed. The Rp statistic, a modified version of the Shapiro-Wilks test, measures the degree of normality of a dataset

by correlating the data with the standard normal distribution (Wilks, 1995). The Rp for the MODIS anomaly shown in Figure 8 is 0.17, whereas the NNndvi anomaly has a value of 0.45, and the AVHRR 0.47. So although the neural net correction has improved the data significantly, there are still differences that are systematic for every pixel. The quality of the corrected data is significantly better, however, as can be seen in Figure 9. The removal of cloud contamination in regions, such as the Gulf of Guinea, that have always had depressed NDVI signal in the AVHRR dataset, is a contribution that should not be underestimated.

6.0 Discussion

The lack of reliable climate observations throughout the AVHRR record is a major limitation in all attempts to correct the AVHRR data to match the quality of the MODIS record. In order to remove the systematic difference between the AVHRR and MODIS data due to atmospheric water vapor, we need accurate observations of the amount of water vapor in the atmosphere at the time of data acquisition. For AVHRR, the instrument that provides this data are derived from the Total Ozone Mapping Spectrometer (TOMS) data (McPeters et al., 1998). TOMS data has its own problems with data continuity and algorithms which may reduce the effectiveness of the neural network because the issues may interfere with the NDVI differences we are trying to remove.

One reason for the lack of strong results in this experiment is the use of aggregated data. The temporal mismatch between the 15 day AVHRR data, the 16 day MODIS data and the monthly TOMS datasets has consequences that are difficult to identify. Although an effort was made to minimize these problems through aggregation to the monthly time step, they may confound the neural net. Aggregated data is much cleaner than daily observations, requires far less

computational effort (a key factor in running neural networks), and are the most widely used products. In addition, daily data for the AVHRR NDVI and reflectances are currently not available, thus they are not used here.

An effort is being made in the context of a NASA funded collaborative project called the Long Term Data Record at the University of Maryland. In this project, daily AVHRR NDVI from NOAA 7 through 14 (1981 to 1999) will be combined directly with MODIS data from 2000 onward. The data from the year 2003 will be used to relate the two datasets. The research presented in this paper will illuminate the efforts of this project.

7.0 Conclusion

Remote sensing datasets are the result of a complex interaction between the design of a sensor, the spectral response function, stability in orbit, the processing of the raw data, compositing schemes, and post-processing corrections for various atmospheric effects including clouds and aerosols. The interaction between these various elements is often non-linear and non-additive, where some elements increase the vegetation signal to noise ratio (compositing, for example) and others reduce it (clouds and volcanic aerosols) (Los, 1998). Thus, although other authors have used simulated data to explore the relationship between AVHRR and MODIS (Trishchenko et al., 2002; van Leeuwen et al., 2006), these techniques are not directly useful in producing a sensor-independent vegetation dataset that can be used by data users in the near term.

There are substantial differences between the processed vegetation data from AVHRR and MODIS [3, 7]. In order to have long data record that utilizes all available data back to 1981, we must find

practical ways of incorporating the AVHRR data into a continuum of observations that include both MODIS and VIIRS. The results in this paper show that the TOMS data record on clouds, ozone and aerosols can be used to identify and remove sensor-specific atmospheric contaminants that differentially affect the AVHRR over MODIS. Other sensor-related effects, particularly those of changing BRDF, viewing angle, illumination, and other effects that are not accounted for here, remain important sources of additional variability. Although this analysis has not produced a dataset with identical properties to MODIS, it has demonstrated that a neural net approach can remove most of the atmospheric-related aspects of the differences between the sensors, and match the mean, standard deviation and range of the two sensors. A similar technique can be used for the VIIRS sensor once the data is released.

References

Atkinson, P.M. and Tatnall, A.R.L., 1997. Introduction: Neural networks in remote sensing. International Journal of Remote Sensing, 18(4): 699 - 709.

Benediktsson, J.A., Swain, P.H. and Ersoy, O.K., 1990. Neural network approaches versus statistical methods in classification of multisource remote sensing data. IEEE Transactions in Geoscience and Remote Sensing, 28: 540-552.

Bishop, C., 1995. Neural Networks for Pattern Recognition. Oxford University Press.

Bounoua, L., Collatz, G.J., Los, S.O., Sellers, P.J., Dazlich, D.A., Tucker, C.J. and Randall, D.A., 2000. Sensitivity of climate to changes in NDVI. Journal of Climate, 13(13): 2277-2292.

Brown, M.E., Pinzon, J. and Tucker, C., 2004. New Vegetation Index Dataset Available to Monitor Global Change. EOS Transactions, 85(52): 565-569.

Brown, M.E., Pinzon, J.E., Didan, K., Morisette, J.T. and Tucker, C.J., 2006. Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-Vegetation, SeaWiFS, MODIS and LandSAT ETM+. IEEE Transactions Geoscience and Remote Sensing, 44(7): 1787-1793.

Cihlar, J., Tcherednichenko, I., Latifovic, R., Li, Z. and Chen, J., 2001. Impact of Variable Atmospheric Water Vapor Content on AVHRR Data Corrections over Land. IEEE Transactions Geoscience and Remote Sensing, 39(1): 173-180.

Eklundh, L. and Olsson, L., 2003. Vegetation index trends for the African Sahel 1982-1999. Geophysical Research Letters, 30(8).

Gallo, K.P., Ji, L., Reed, B.C., Dwyer, J. and Eidenshink, J.C., 2003. Comparison of MODIS and AVHRR 16-day normalized difference vegetation index composite data. Geophysical Research Letters, 31: L07502-5.

Giri, C., Zhu, Z. and Reed, B.C., 2004. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. Remote Sensing of Environment, 94: 123-132.

Holben, B., 1986. Characteristics of Maximum-Value Composite Images from Temporal AVHRR Data. International Journal of Remote Sensing, 7(11): 1417-1434.

Huete, A., Justice, C. and Liu, H., 1994. Development of Vegetation and Soil Indices for MODIS-EOS. Remote Sensing of Environment, 49: 224-234.

Huete, A.R., Didan, K., Shimabukuro, Y.E., Ratana, P., Saleska, S.R., Hutyra, L.R., Yang, W., Nemani, R.R. and Myneni, R.B., 2006. Amazon rainforests green-up with sunlight in dry season. Geophysical Research Letters, 33: L060405.

Huete, A.R. and Jackson, R.D., 1988. Soil and Atmosphere Influences on the Spectra of Partial Canopies. Remote Sensing of Environment, 25: 89-105.

Huete, A.R. and Tucker, C.J., 1991. Investigation of soil influences in AVHRR red and near-infrared vegetation index imagery. International Journal of Remote Sensing, 12(6): 1223-1242.

Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE ASSP Magazine: 4-22.

Los, S.O., 1998. Estimation of the Ratio of Sensor Degradation Between NOAA AVHRR Channels 1 and 2 from Monthly NDVI Composites. IEEE Transactions on Geoscience and Remote Sensing, 36(1): 206-213.

McPeters, R., Bhartia, P.K., Krueger, A., Herman, J., Wellemeyer, C., Seftor, C., Jaross, G., Torres, O., Moy, L., Labow, G., Byerly, W., Taylor, S., Swissler, T. and Cebula, R., 1998. Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Product User's Guide, National Aeronautics and Space Administration, Greenbelt, MD.

Miura, T., Huete, A. and Yoshioka, H., 2006. An empirical investigation of cross-sensor relationships of NDVI and red/near-infrared reflectance using EO-1 Hyperion data. Remote Sensing of Environment, 100(2): 223-236.

Pinheiro, A.C., Privette, J.L., Mahoney, R. and Tucker, C.J., 2004. Directional Effects in a Daily AVHRR Land Surface Temperature Data Set over Africa. IEEE Transactions in Geoscience and Remote Sensing, 42(9): 1941-1954.

Pinzon, J., 2002. Using HHT to successfully uncouple seasonal and interannual components in remotely sensed data, SCI 2002 Conference Proceedings Jul 14-18. SCI International, Orlando, Florida.

Pinzon, J., Brown, M.E. and Tucker, C.J., 2005. Satellite time series correction of orbital drift artifacts using empirical mode decomposition. In: N. Huang and S.S.P. Shen (Editors), Hilbert-Huang Transform: Introduction and Applications. World Scientific, Hackensack NJ, pp. 167-186.

Rudolf, B., Hauschild, H., Rueth, W. and Schneider, U., 1994. Terrestrial Precipitation Analysis: Operational Method and Required Density of Point Measurements. In: M. Desbois and F. Desalmand (Editors), Global Precipitations and Climate Change. NATO ASI Series. Springer-Verlag, pp. 173-186.

Rumelhart, D.E., McClelland, J. and group, t.P.r., 1986. Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 1 - Foundations. MIT, 547 pp.

Sellers, P.J., Collatz, J., Hall, F.G., Meeson, B.W., Closs, J., Corprew, F., MeManus, J., Myers, D., Sun, K.-J., Dazlich, D., Kerr, Y., Koster, R., Los, S., Mitchell, K. and Try, P., 1996. The ISLSCP Initiative I Global Datasets: Surface Boundary Conditions and Atmospheric Forcings for Land-Atmosphere Studies. Bulletin of the American Meteorological Society, 77(9): 1987-2005.

Slayback, D.A., Pinzon, J.E., Los, S.O. and Tucker, C.J., 2003. Northern hemisphere photosynthetic trends 1982-99. Global Change Biology, 9(1): 1-15.

Stathakis, D., 2008. How many hidden layers and nodes? International Journal of Remote Sensing, submitted.

Stathakis, D. and Kanellopoulos, I., 2008. Global elevation ancillary data for land use classification using granular neural networks. Photogrammetric Engineering and Remote Sensing, 74(1).

Stathakis, D. and Vasilakos, A., 2006. Comparison of several computational intelligence based classification techniques for remotely sensed optical image classification. IEEE Transactions in Geoscience and Remote Sensing(44): 8.

Steven, M.d., Malthus, T.J., Baret, F., Xu, H. and Chopping, M.J., 2003. Intercalibration of vegetation indices from different sensor systems. Remote Sensing of Environment, 88: 412-422.

Teillet, M., Staenz, K. and Williams, D.J., 1997. Effects of Spectral, Spatial and Radiometric Characteristics on Remote Sensing Vegetation Indices of Forested Regions. Remote Sensing of Environment, 61: 139-149.

Trishchenko, A.P., Cihlar, J. and Li, Z., 2002. Effects of spectral response function on surface reflectance and NDVI measured with moderate resolution satellite sensors. Remote Sensing of Environment, 81: 1-18.

Tucker, C.J., 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. Remote Sensing of Environment, 8: 127-150.

Tucker, C.J., Pinzon, J.E., Brown, M.E., Slayback, D., Pak, E.W., Mahoney, R., Vermote, E. and El Saleous, N., 2005. An Extended AVHRR 8-km NDVI Data Set Compatible with MODIS and SPOT Vegetation NDVI Data. International Journal of Remote Sensing, 26(20): 4485-4498.

Tucker, C.J., Slayback, D.A., Pinzon, J.E., Los, S.O., Myneni, R.B. and Taylor, M.G., 2001. Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. International Journal of Biometeorology, 45(4): 184-190.

van Leeuwen, W., Orr, B.J., Marsh, S.E. and Herrmann, S.M., 2006. Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications. Remote Sensing of Environment, 100: 67-81.

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences, an Introduction. Academic Press, San Diego.

Zeng, N., Neelin, J.D. and Lau, W.K.-M., 1999. Enhancement of Interdecadal Climate Variability in the Sahel by Vegetation Interaction. Science, 286: 1537-1540.

Zobler, L., 1986. A world soil file for global climate modeling, NASA.

Captions

Table 1. Global datasets used in this paper.

Table 2. Statistics of the MODIS, AVHRR, and NNndvi datasets for 48 months of data (2000-2003).

Figure 1. Schematic representation of the neural network used in this paper.

Figure 2. Graph showing the latitudinal means of the difference between MODIS, AVHRR and NNndvi for January 2003. The figure highlights the zones where the neural net correction is the strongest.

Figure 3. Zonal mean (averaged per latitude) of the difference between MODIS and AVHRR (Panel A) and MODIS and NNndvi (Panel B) through time from 2000 to 2003.

Figure 4. Latitude-averaged mean of NNndvi from 1982 to 2003.

Figure 5. Root mean square error from MODIS-AVHRR (above) and the MODIS-NNndvi (below) from 2000 to 2003 in NDVI units.

Figure 6. Time series plots of six latitude-longitude locations: A. Louga, Senegal (16, -16), Tigray Ethiopia (14, 40), Bondville Illinois (10, -88), Cascades Washington (44,-122), Harvard Forest Massachusetts (43,-72), and Ji-Parana Brazil (-11,-62).

Figure 7. Correlation coefficient of AVHRR, (A), NNndvi (B), and MODIS (C) vs GPCC rainfall data. Panel D shows the histogram of the correlation coefficient of the NDVI vs gridded rainfall by percent.

Figure 8. The August 2003 anomaly, defined as the difference between the MODIS, AVHRR and NNndvi image for August 2003 and the mean of four August MODIS images (2000-2003).

Figure 9. Africa subset of one degree images for July 2002 for the AVHRR (A), NNndvi (B), and the difference between the two (C).

Table 1.

Sensor	AVHRR NDVI	MODIS NDVI	GPCC Rain	TOMS reflectivity, ozone and aerosol
Data Source	GIMMS NDVIg Operational Dataset	MODIS-Land and Atmospheres	Gridded Gauge data	NASA GSFC Ozone Processing Team
Native Spatial Resolution	8000 m	250 m	1 degree	26 km
Temporal Resolution	15 day	16-day	monthly	Daily
Period Available	July 1981 – present (NOAA 7, 9, 11,14,16 and 17)	Feb 2000 – present	April 1986 – present	11/1978-5/1993(Nimbus 7) 5/1993-11/1994 (Meteor 3) 7/1996-12/2005 (Earth Probe)
Equatorial Crossing	~9 AM - ~6 PM	10.30 AM	NA	~9 AM - ~6 PM
Field of View (FOV)	±55.4º	±55º	NA	±55.4º

Table 2.

Element	Accumulated weight
AVHRR NDVI	0.6
TOMS Reflectance	0.5
TOMS Column Ozone	0.3
Land Surface Type	0.3
TOMS Aerosol Index	0.2
Soil cover	0.2
Digital Elevation Model	0.2

Table 3.

Sensor	NNndvi	AVHRR	MODIS
Global Mean NDVI	0.4834	0.2982	0.4830
Global Std NDVI	0.2384	0.2460	0.2522

Figure 1.

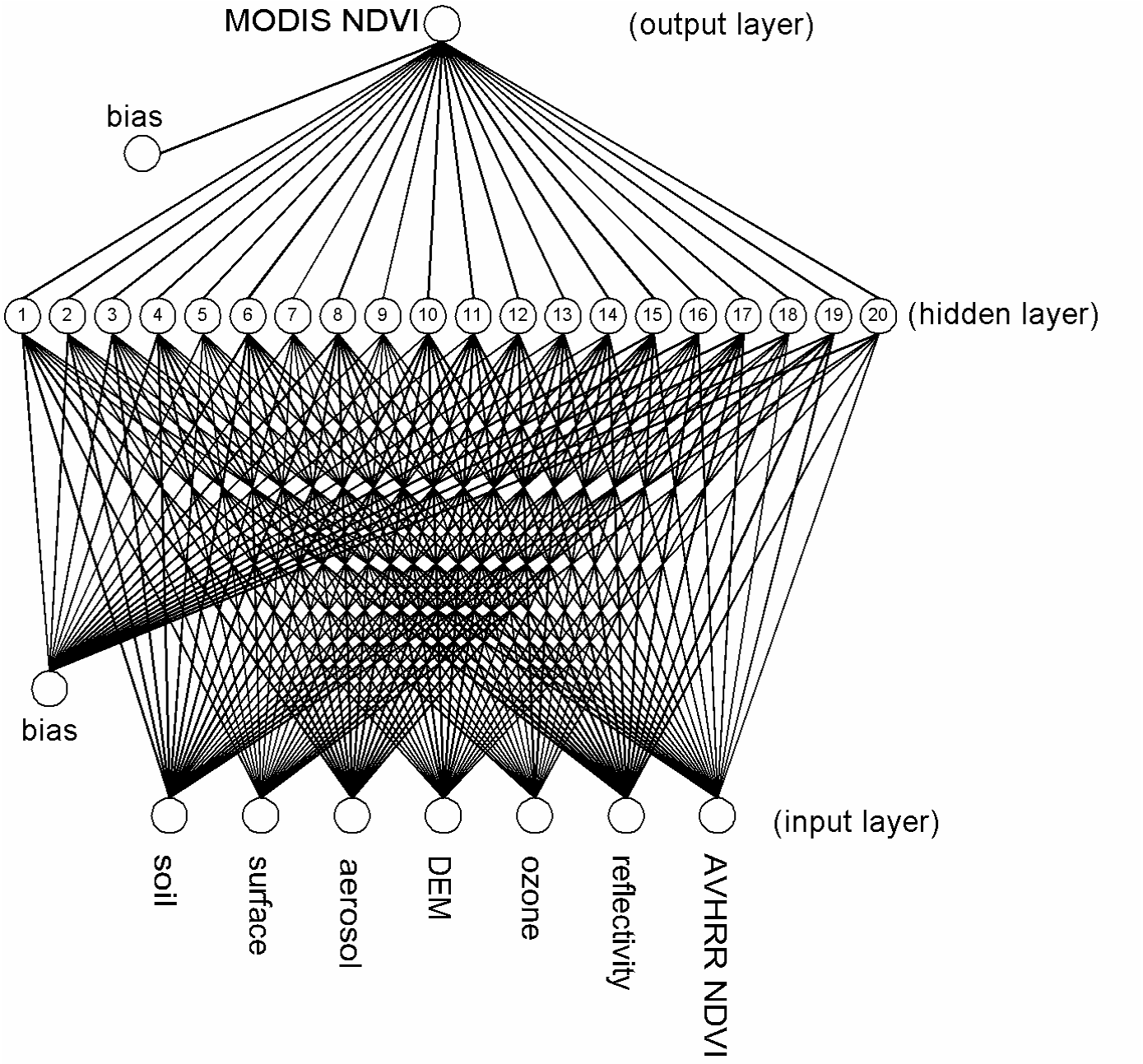
MODIS NDVI
(output layer)
bias
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
(hidden layer)
bias
(input layer)
soil
surface
aerosol
DEM
ozone
reflectivity
AVHRR NDVI

Figure 2.

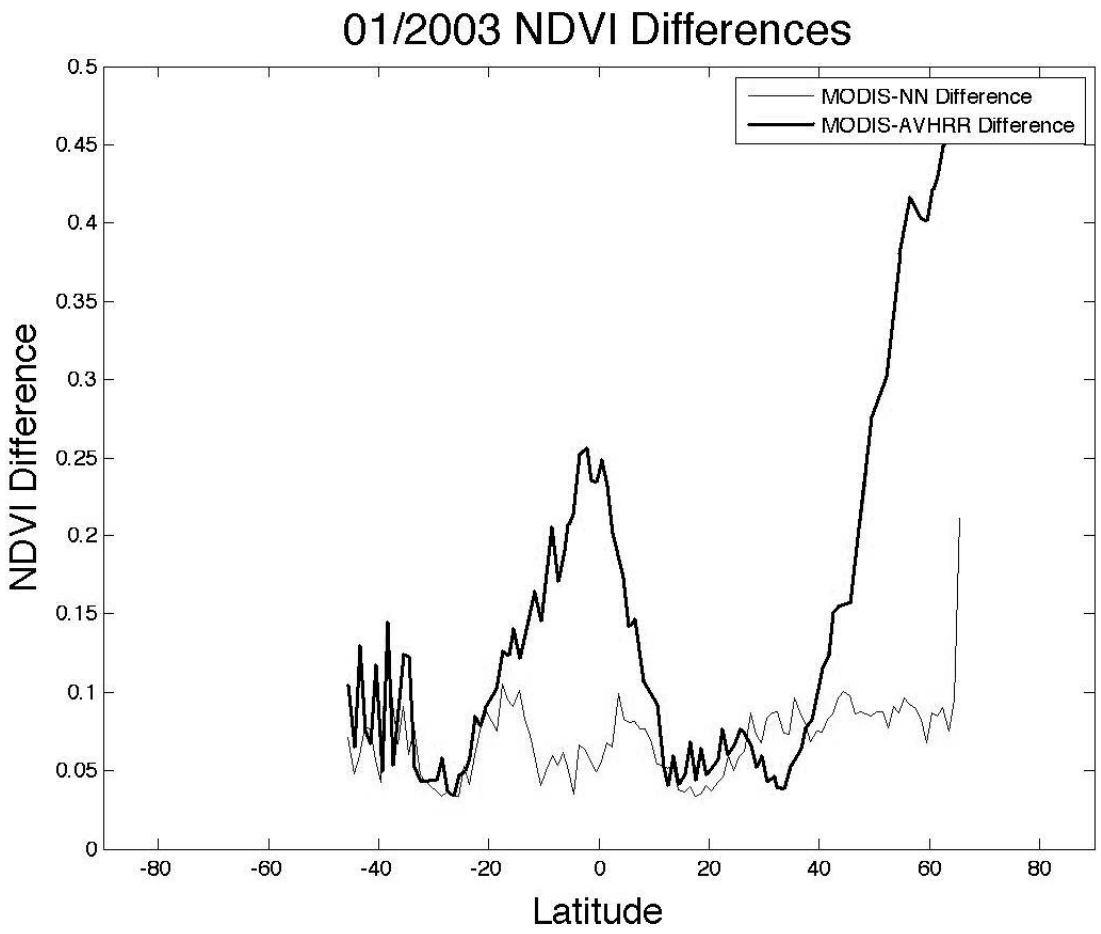

Figure 3.

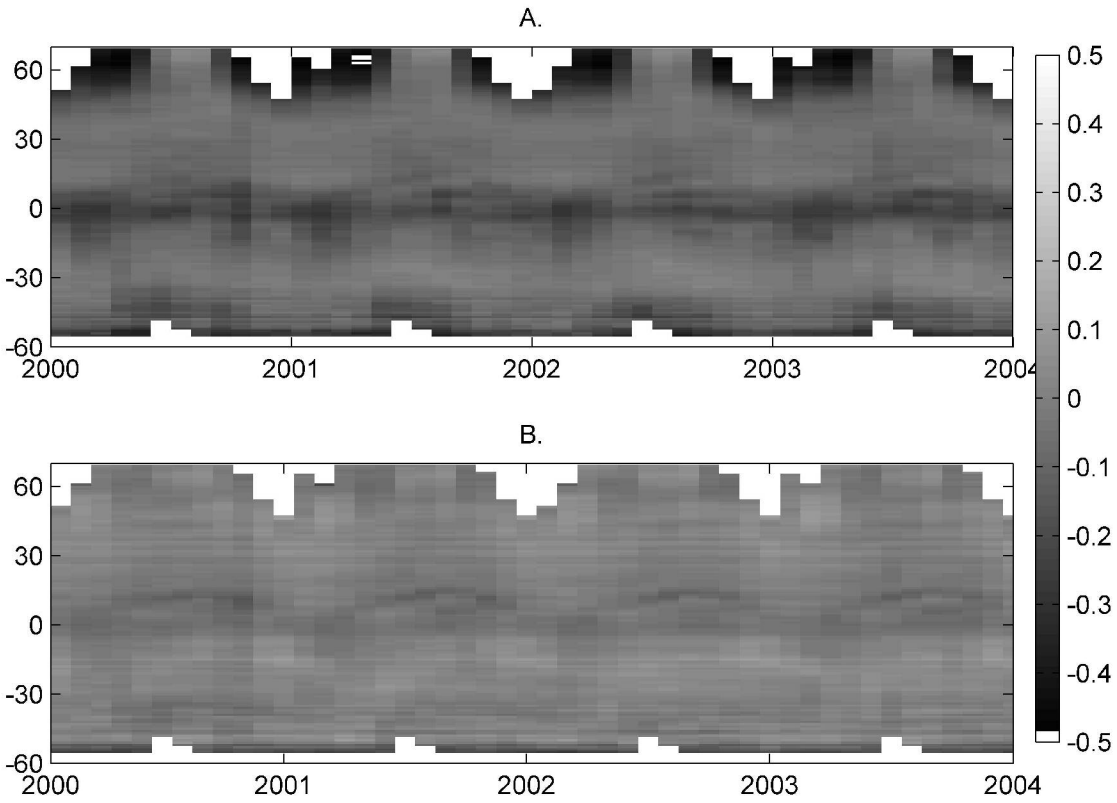

Figure 4.

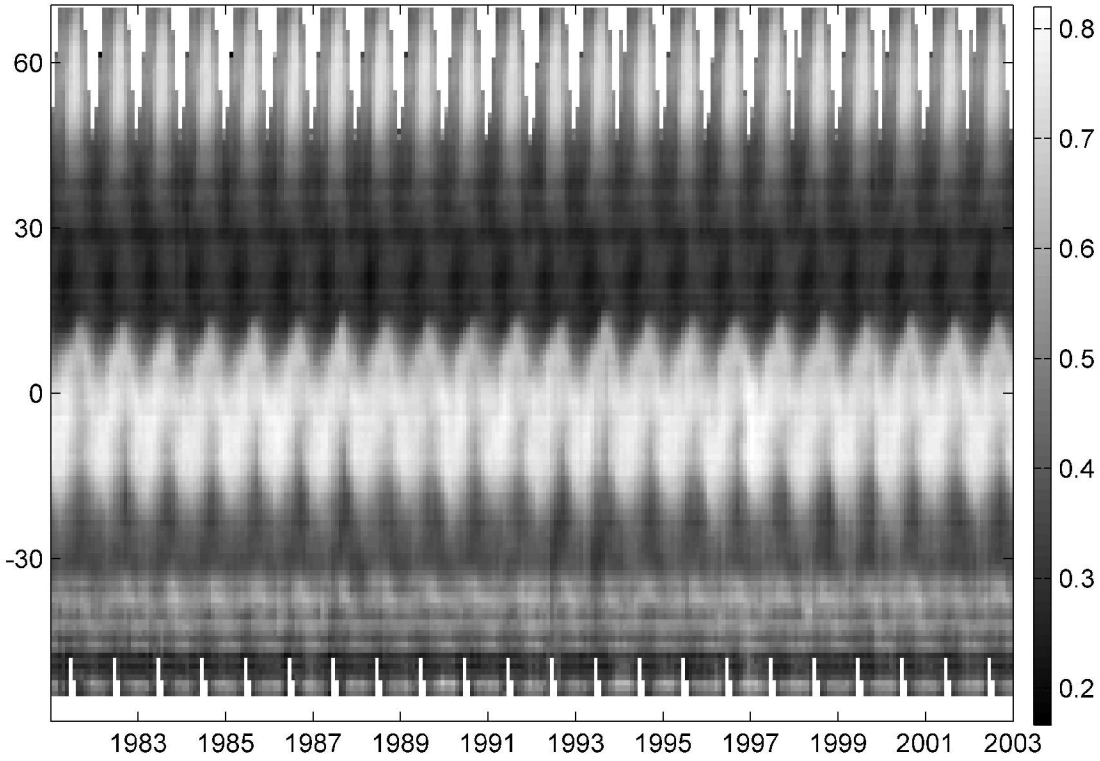

Figure 5.

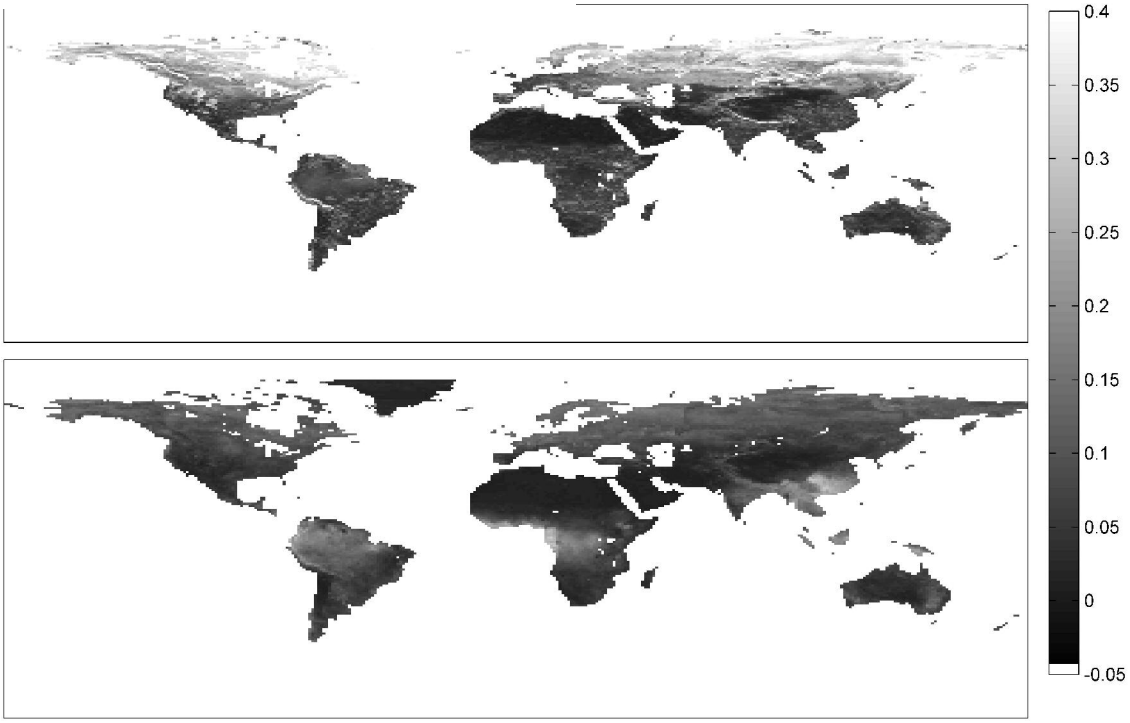

Figure 6.

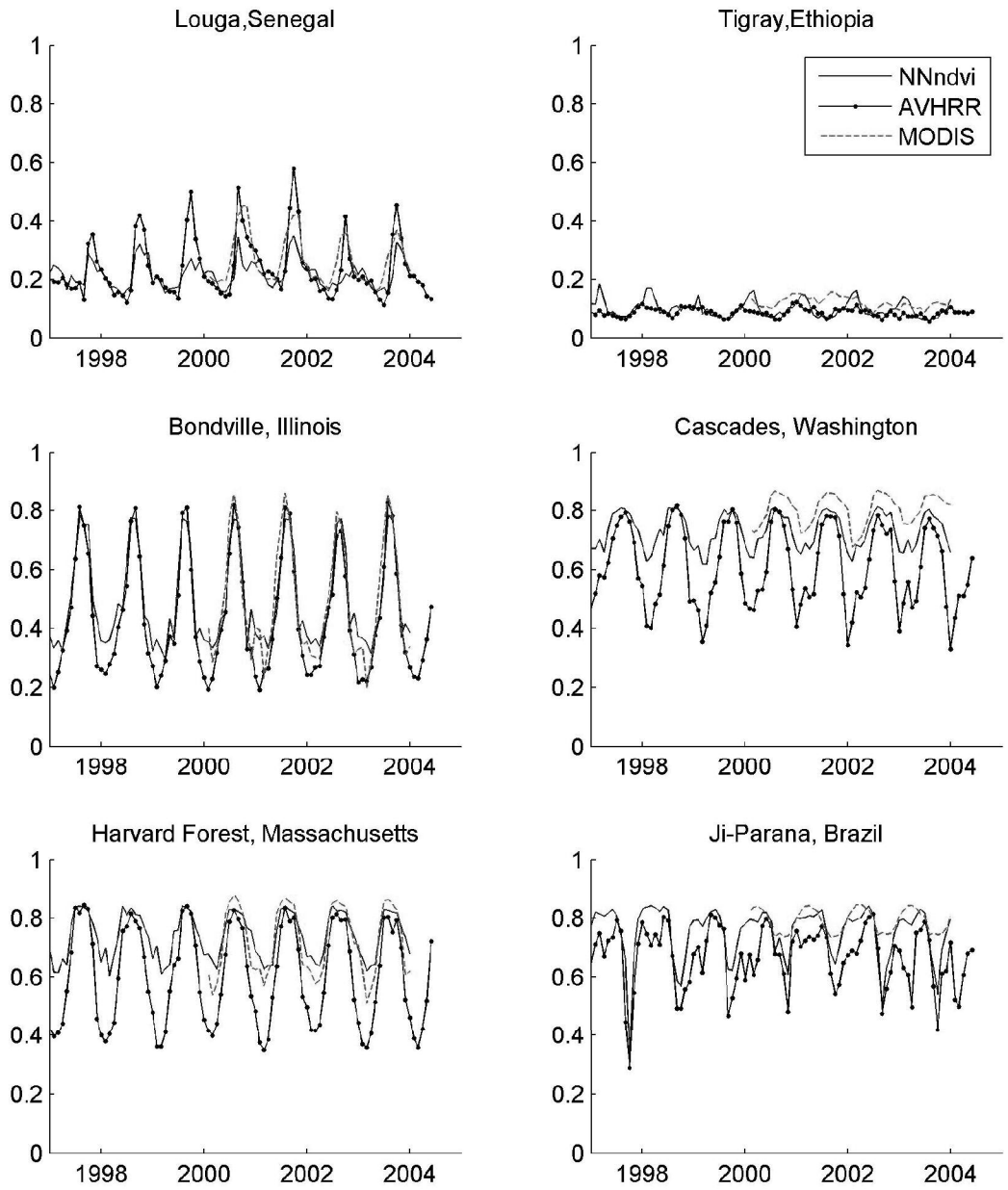
Louga,Senegal
Tigray,Ethiopia
NNndvi
AVHRR
MODIS
Bondville, Illinois
Cascades, Washington
Harvard Forest, Massachusetts
Ji-Parana, Brazil
0
0.2
0.4
0.6
0.8
1
1998
2000
2002
2004

Figure 7.

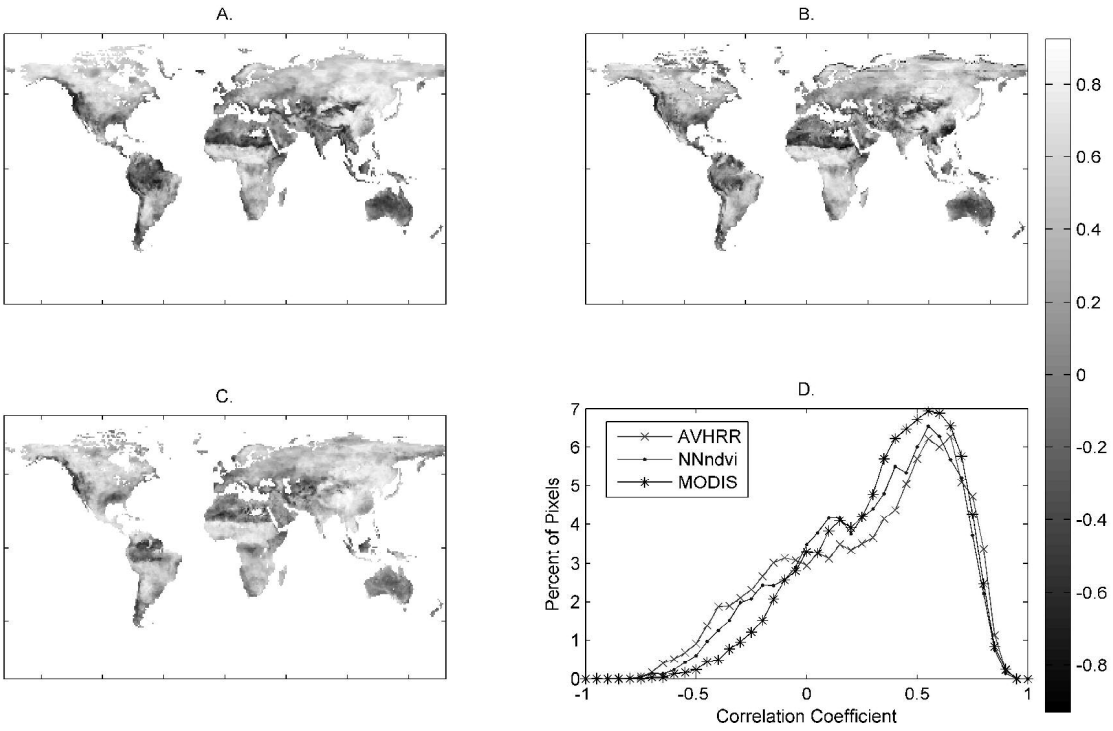

Figure 8.

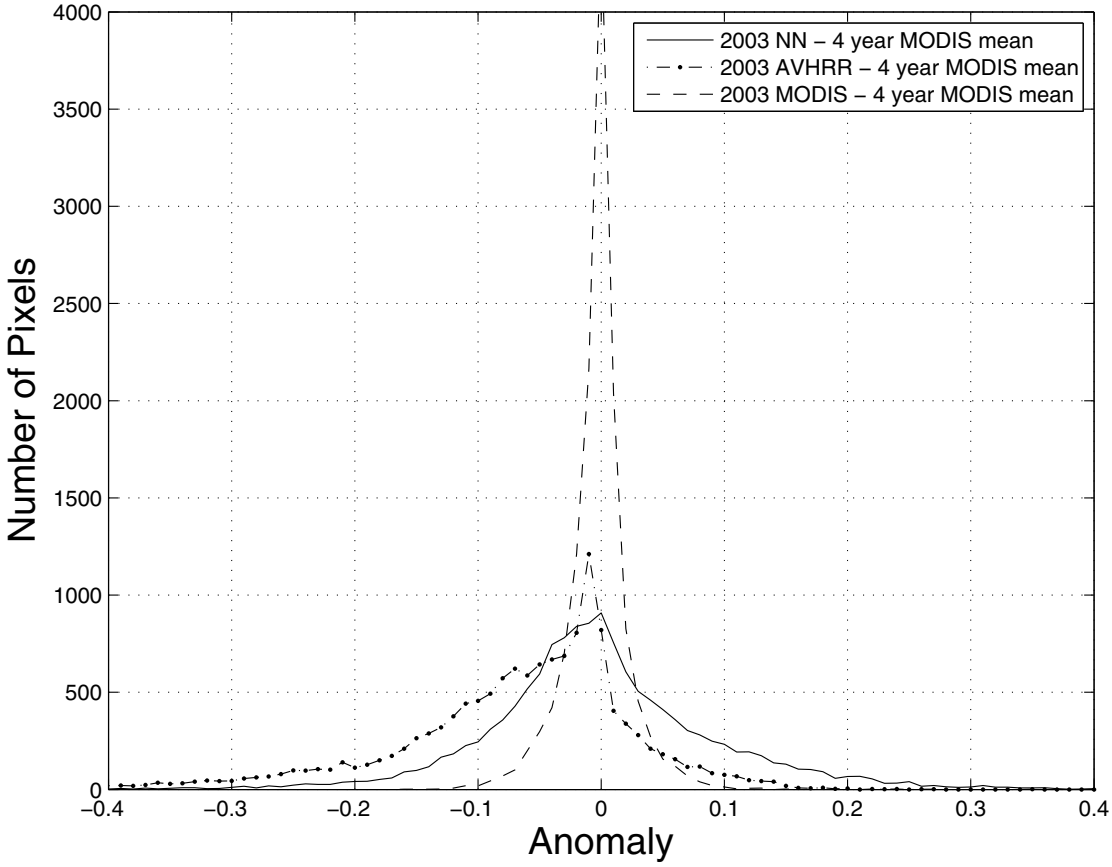

Figure 9.