Chapter 8
Main Global NDVI Datasets, Databases, and Software

Coarse spatial resolution datasets are invaluable at the global scale, but they lack the thematic and spatial detail required for habitat assessments at the country level and for finer-resolution assessments such as vegetation species distribution or high-quality forest-change monitoring. Mapping, monitoring, and assessments at the national and subnational level are performed using moderate-resolution sensors such as Landsat, ASTER, SPOT HRV, and IRS with spatial resolutions from 15 to 60 m. Newer, high-resolution optical sensors (5 m or better) provide enough spatial and spectral detail to discriminate between individual trees and, in some cases, species, but high-resolution imagery is prohibitively costly (see Annex 7) for many national governments and research institutions (Strittholt and Steininger 2007).

8.1 Main NDVI Datasets

Significant research effort has been invested in processing satellite sensor data into NDVI. The most common sensor used in these initiatives is the AVHRR sensor on board the NOAA satellites which enables assembly of global NDVI datasets. The development of these datasets by different research groups involves diverse schemes, protocols, and algorithms for corrections and processing (Scheftic et al. 2014). As a result, the environmental change community currently has a range of datasets that may be used for a variety of applications (Table 8.1).

The most widely used global NDVI datasets are the Global Inventory for Mapping and Modeling Studies (GIMMS); NOAA/NASA Pathfinder (PAL); the Long-Term Data Record (LTD); and the Fourier-Adjusted, Sensor and Solar zenith angle corrected, Interpolated, Reconstructed (FASIR) adjusted (NDVI) (see Annex 1).

The *Global Inventory for Mapping and Modeling Studies (GIMMS)* dataset is the most updated global time-series NDVI product (Fensholt and Proud 2012). It has a temporal resolution of 2 weeks (24 scenes/year) and a spatial resolution

G.T. Yengoh et al., *Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales*, SpringerBriefs in Environmental Science, DOI 10.1007/978-3-319-24112-8_8

Table 8.1 Commonly utilized normalized difference vegetation index (NDVI) datasets (modified from Higginbottom and Symeonakis 2014)

Name	Sensor	Time span	Time step	Resolution
Pathfinder (PAL)	AVHRR	1981–2001	10-day	8 km
Global Vegetation Index (GVI)	AVHRR	1981–2009	7-day	4 km
Land Long-Term Data Record (LTDR)	AVHRR	1981–2013	Daily	5 km
Fourier-Adjusted, Sensor and Solar zenith angle corrected, Interpolated, Reconstructed (FASIR)	AVHRR	1982–1998	10-day	0.125°
GIMMS	AVHRR	1981–2006	15-day	8 km
GIMMS3g	AVHRR	1981–2015	15-day	8 km
S10	SPOT-vegetation	1998+	10-day	1 km
EM10	ENVISAT-MERIS	2002–2012	10-day	1/1.2 km
SeaWiFS	SeaWiFS	1997–2010	Monthly	4 km
MOD (MYD)13 A1/A2	Terra (Aqua)	2000+	16-day	500 m/1 km
MOD13 (MYD) A3			Monthly	1 km
MOD13 (MYD) C1/C2			16-day/monthly	5.6 km
MOD13 (MYD) Q1	MODIS		16-day	250 m
MEDOKADS	AVHRR	1989+	Daily	1 km

of approximately 8 km. The GIMMS NDVI3g dataset (Pinzon and Tucker 2014) now comprises more than 33 years of data corrected for instrument calibration, variations in solar angle and view zenith angle, stratospheric aerosols from major volcanic eruptions, and other effects not related to vegetation change. Cloud and haze effects are minimized by taking the highest fortnightly value within composite 8 km blocks of pixels (Holben 1986).

The *NOAA/NASA Pathfinder (PAL) NDVI* dataset was created from the PAL 8 km daily product (Green and Hay 2002; James and Kalluri 1994). The PAL 8 km daily data were spatially re-sampled, based on maximum NDVI values from AVHRR Global Area Coverage data which have a minimal resolution of 4 km (De Beurs and Henebry 2005; James and Kalluri 1994). The PAL Global 10-day composite NDVI product is part of the Pathfinder Land dataset archived at the Goddard Earth Sciences, Distributed Active Archive Center (GES-DAAC) (Green and Hay 2002). The data have been corrected for changes in sensor calibration, ozone absorption, Rayleigh scattering, and sensor degradation after prelaunch calibration and has been normalized for changes in solar zenith angle (James and Kalluri 1994). The dataset is not continually being processed, and data after 1999 are not accessible online at the GES-DAAC website of the Goddard Space Flight Center.

Long-Term Data Record (LTD) is a global daily dataset of 0.05° (about 5 km ground spatial distance) developed by the NASA-funded LTDR Project. The dataset is currently at its fourth version and available for the period 1981–2013 from the reprocessing of the N07-N18 AVHRR data (Pedelty et al. 2007). The current version includes records from the processing of data from NOAA-16 and NOAA-17 lengthening the LTDR records from AVHRR to 2013.

The Fourier-Adjusted, Sensor and Solar zenith angle corrected, Interpolated, Reconstructed (FASIR) adjusted NDVI datasets are products of the International Satellite Land-Surface Climatology Project, Initiative II (ISLSCP II) data collection, developed to provide a 17-year satellite record of monthly changes in the photosynthetic activity of terrestrial vegetation for use in general circulation climate models and biogeochemical models (Sellers et al. 1994; Sietse 2010). The NDVI collections are provided in data files at spatial resolutions of 0.25, 0.5, and 1.0° latitude/longitude. FASIR adjustments concentrated on reducing NDVI variations arising from atmospheric, calibration, view, and illumination geometries and other effects not related to actual vegetation change (Sietse 2010).

The *Moderate-Resolution Imaging Spectrometer (MODIS)* is an extensive program using sensors on both the Terra and Aqua satellites, each of which provide complete daily coverage of the Earth. Started in January 2000, the MODIS sensor provides vegetation indices (NDVI and EVI) produced globally on 16-day intervals at three resolutions (250, 500, and 1000 m). The MODIS NDVI data are fully consistent spatially and temporally with AVHRR-NDVI products (Tucker et al. 2005). Comparisons between AVHRR and MODIS NDVI products over a wide range of vegetation types have shown a very high correlation, $r > 0.9$ according to (Gallo et al. 2005). See also Fig. 9.1. A complete collection of MODIS Land products can be accessed freely online either from USGS or NASA sites (see Annex 7).

Before deciding on the data to be used for NDVI-related analysis, users should reflect on a number of questions. According to Khorram et al. (2012), some of the most important of these questions include:

- What kind of remotely sensed data do I need? More specifically, what types of data are available from today's remote sensing instruments, and what are their strengths and limitations?
- How must my chosen data be prepared prior to analysis? What are the appropriate processing and/or analytical methods?
- What is the accuracy of the output products I have created? Is that accuracy sufficient for my ultimate objectives?

8.2 Quality-Related Considerations

The potential for using free data for assessment and monitoring of environmental change (principally forest cover change) at the global level has been most clearly demonstrated for Landsat products. The key challenges for creating global products on forest cover and cover change are the processes and tools for atmospheric correction, proper calibration coefficients, working with different phenologies between compilations, terrain correction, accuracy assessment, and the automation of land cover characterization and change detection (Townshend et al. 2012). Most of the commonly used datasets mentioned above (such as the PAL, GIMMS, LTDR, and the FASIR) have undergone many evaluations and intercomparisons on a range of criteria (Beck et al. 2011; Fensholt and Proud 2012).

Beck et al. (2011) undertook a global intercomparison of the four AVHRR-NDVI datasets (PAL, GIMMS, LTDR, FASIR) against Landsat imagery for the period 1982–1999, finding significant differences in trends for almost half of the total land surface. The PAL and the LTDR (Version 3) datasets lacked calibration; GIMMS had the best calibration and was the most accurate in terms of temporal change. In a study investigating whether vegetation trends derived from NDVI and phenological parameters are consistent across products, Yin et al. (2012) compared GIMMS and SPOT–VGT-derived NDVI. Strong similarities were found in interannual trends and, also, in trends of the seasonal amplitude and annual sum NDVI. But there were significant discrepancies between NDVI-derived trends based on phenological parameters such as amplitude (maximum increase in canopy photosynthetic activity above the baseline) and integral of NDVI (canopy photosynthetic activity across the entire growing season) (Yin et al. 2012). These correspond to seasonal vegetation cycles revealed by GIMMS and SPOT VGT. The study attributed these discrepancies to variables such as land cover and vegetation density. Such discrepancies highlight the need for appropriate and rigorous preprocessing when working with data from different remote sensing systems.

8.3 Precipitation Datasets

Various precipitation datasets are used in combination with NDVI data in many earth science applications. Among the most widely used of these datasets are the Modern-Era Retrospective Reanalysis for Research and Applications (MERRA), Interim Reanalysis (or ERA-Interim Reanalysis), Global Precipitation Climatology Project (GPCP), Africa Rainfall Climatology, and the VASClimO (Table 8.2).

The *Modern-Era Retrospective Reanalysis for Research and Applications (MERRA)* is a NASA reanalysis for the satellite era using the Goddard Earth Observing System Data Assimilation System Version 5 numerical weather and climate model. The Project focuses on historical analyses of the hydrological cycle on

Table 8.2 Commonly used precipitation datasets for earth science and environmental applications

Precipitation data	Time span and scale	Reference
NASA MERRA	1979–present at 0.5° × 0.5°	Rienecker et al. (2011)
ERA-Interim	1979–present at 80 km	Dee et al. (2011)
GPCP	1979–2012 at 1.0° × 1.0°	Huffman et al. (2009)
African Rainfall Climatology	1983–2013 at 0.1° × 0.1°	Novella and Thiaw (2013)
VASClimO	1951–2000 at 0.5°, 1.0°, and 2.5°	Beck et al. (2005), Schneider et al. (2008)
TRMM	1997–present at 0.25° × 0.25°	Gentemann et al. (2004)

a broad range of weather and climate time scales and places the NASA EOS suite of observations in a climate context. This dataset has a spatial resolution of 0.5° × 0.5° from 1979 to the present (Rienecker et al. 2011).

The *Interim Reanalysis* (or ERA-Interim Reanalysis) output comes from the European Centre for Medium Range Weather Forecasts. It is a global atmospheric reanalysis from 1979, continuously updated in real time through the present with a spatial resolution of 80 km (Dee et al. 2011).

The *Global Precipitation Climatology Project (GPCP)* Version 2.2 is a blend of precipitation gauge data and satellite data taking advantage of the strengths of each data type. These data are 1° × 1° and run from 1979 to 2012 (Huffman et al. 2009). Tropical Rainfall Measuring Mission or TRMM data form the basis of the GPCP dataset and are blended with station data to improve the rainfall accuracies.

The *Africa Rainfall Climatology* Version 2 is a gridded, daily 30-year (1983–2013) precipitation dataset at 0.1° × 0.1° spatial resolution produced by NOAA's Climate Prediction Center (Novella and Thiaw 2013), produced using an operational rainfall estimation algorithm now updated to Rainfall Estimates Version 2 (Novella and Thiaw 2013).

The *VASClimO* is a global dataset of station-observed precipitation produced by gridding 9343 homogeneity-checked station time series of precipitation for the period 1951–2000 (Rudolf et al. 2005). It provides a globally gridded total monthly precipitation from January 1951 to December 2000 at three resolutions, 0.5° × 0.5°, 1.0° × 1.0°, and 2.5° × 2.5°, and is updated by the GPCC full-data reanalysis product version 4 (Schneider et al. 2008).

Tropical Rainfall Measuring Mission (TRMM) dataset is obtained by active and passive microwave measurements derived by instruments on board the Tropical Rainfall Measuring Mission's (TRMM) Microwave Imager (TMI). Besides measuring rain rates, the TMI can also measure sea surface temperature (SST), ocean surface wind speed, columnar water vapor, and cloud liquid water. TRMM is a joint program between NASA and the National Space Development Agency of Japan (Gentemann et al. 2004).

8.4 NDVI Software

Besides the data requirement, the creation of NDVI products requires software. There are many software products that can be used to create NDVI products and they can be categorized according to the tool sets that they offer, as well as the tasks that each of them can accomplish (Steiniger and Hunter 2013). The general demands for software for geospatial analysis include capabilities for data capture and representation, data visualization and exploration, data editing, data storage, integration of data from different sources, data queries to select a subset of the data, data analysis, creation of new data from existing or available input data, data transformation, and elements of cartographic representation. Not all software have full out of the box capabilities of accomplishing these different tasks. Many software depend on extensions, plug-ins, and application program interfaces (sets of routines, protocols, and tools which specify how software components should interact and are used when programming graphical user interface components) to accomplish some of these tasks (Steiniger and Hunter 2013). The creation of NDVI products demands additional capabilities for the modification, as well as spectral transformation of aerial and satellite image data. The appropriate geospatial software for working with remote sensing products and the creation of NDVI products therefore need capabilities for image radiometric and geometric correction, filtering, georeferencing and ortho-rectification, mosaiking, vectorization, and image object extraction (Steiniger and Hunter 2013; Jensen 2007; Mather and Koch 2011).

Geospatial software may be commercial or free products. There are many commercial software products with different strengths and levels of specialization in the delivery of remote sensing services and products (see Annex 8A for some of the commonly used desktop software). There are also free or open-source software which offer a range of possibilities for geospatial analyses, including capabilities for the creation of NDVI products (see Annex 8B). Over the last decade, there has been a notable increase in the availability of free and open-source software projects for geographic data collection, storage, analysis, and visualization (Steiniger and Hunter 2013). The benefits of open-source software are well known and include cost savings, vendor independence, and open standards (Steiniger and Hunter 2013). There have been several initiatives to compile and make inventories of existing geospatial software products.[1]

Besides the desktop applications presented in Annex 8, there are other platforms through which NDVI products can be created. Steiniger and Hunter (2013) point to the importance of Server GIS and WebSphere Process (WPS) Servers which host software that expose GIS and remote sensing functionality typically found in desktop geographic information systems or remote sensing software. Such functionality does not require direct interaction from a user via a user interface (Cepicky and Becchi 2007).

[1] Initiatives such as those by Michael de Smith and Paul Longley of University College London, and Mike Goodchild of University of California, Santa Barbara, provide information on the main activity for which each software product is designed, their license status (commercial or free), as well as links to their access and further information: http://www.spatialanalysisonline.com/software.html. The Wikipedia page that compares geographic information systems and remote sensing software in terms of license status, operating system, and other operational specifications can be found at: http://en.wikipedia.org/wiki/Comparison_of_geographic_information_systems_software.